

Artículo

Hacia un cambio paradigmático para la evaluación de la actividad científica en la Educación Superior

To a paradigm shift to the evaluation of scientific activity into Higher Education

Vicente Manzano-Arrondo

Departamento de Psicología Experimental. Universidad de Sevilla, Sevilla, España

Recibido el 20 de agosto de 2016; aceptado el 24 de agosto de 2017

Disponible en Internet el 19 de septiembre de 2017

Resumen

La evaluación actual de la actividad científica se lleva a cabo mediante un mismo patrón internacional. Esta circunstancia ha estimulado un notable número de investigaciones críticas. Sin embargo, la abundancia de críticas contrasta con la falta de propuestas alternativas. En este trabajo se presenta una síntesis de los inconvenientes observados en las prácticas de evaluación, identificando consecuencias negativas para la propia ciencia, sus miembros y su utilidad pública respecto a las Instituciones de Educación Superior. También se añade una revisión de voces alternativas. Finalmente, se propone un conjunto de ocho principios para ayudar a promover un cambio de paradigma.

© 2016 Asociación Nacional de Universidades e Instituciones de Educación Superior A.C. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Palabras clave: Evaluación científica; Factor de impacto; Utilidad de la ciencia

Abstract

The current assessment of scientific activity is performed by the same international pattern. This has stimulated a remarkable number of critical researches. However, the abundance of critical publications contrasts with the lack of alternative proposals. This paper presents a synthesis of the drawbacks observed in assessment practices, identifying negative consequences for science itself, its members and its public utility,

Correo electrónico: vmanzano@us.es

in reference to Higher Education Institutions. A review of alternative voices is also added. Finally, this paper proposes a set of eight principles to assist in promoting a paradigm shift.

© 2016 Asociación Nacional de Universidades e Instituciones de Educación Superior A.C. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Scientific evaluation; Impact factor; Usefulness of science

Introducción. La práctica controvertida de las evaluaciones

Varios movimientos en el ámbito internacional han propiciado un elevado interés por obtener indicadores compartidos para evaluar la producción científica. Entre ellos, algunos afectan especialmente a las Instituciones de Educación Superior (IES): 1) el modelo de la Nueva Gestión Pública, que establece, entre sus características definitorias, una cultura de rendición de cuentas desde las instituciones públicas (Christensen, 2011); 2) el desarrollo y la extensión de procedimientos de medición cuantitativa del uso documental de las publicaciones científicas, entre las que sobresale el factor de impacto divulgado a través del *Journal Citation Reports* de Thomson y Reuters (Granda, 2003); 3) el afianzamiento de la meritocracia como inspiración para la distribución de refuerzos o recursos (Liu, 2011), y 4) la creciente exigencia de garantías normativas, mediante la utilización de criterios claros, públicos y equitativos, así como de instituciones que los respalden (Manzano-Arrondo, 2015a).

En consonancia con los cuatro puntos anteriores, desde la década de 1980 la mayoría de los países ha implementado procesos de evaluación de la actividad científica mediante organismos y normativas específicas; éstas —y las prácticas que derivan de ellas— han provocado reacciones muy diversas dentro del mundo científico y académico (Mateo, 2012; Watermeyer, 2014). Dentro de estas reacciones, las publicaciones han mostrado preferentemente argumentos y resultados de investigación muy críticos respecto a las normativas específicas en cada región o disciplina, y con mayor abundancia respecto a los principios rectores que justifican el panorama internacional. A pesar de ello, no se observa un esfuerzo equivalente en la propuesta de modelos alternativos, ni parece que las críticas se traduzcan en prácticas nuevas o, en su defecto, mejoradas (Palafox, 2009).

Este trabajo realiza una labor de síntesis, tanto de las críticas como de las propuestas, con el objetivo último de ofrecer una herramienta de trabajo para la emergencia de un cambio paradigmático. Si consideramos la envergadura del objetivo, resulta fundamental precisarlo en los siguientes términos: un foco centrado en la Educación Superior (ES) a nivel global, sin distinguir entre ámbitos disciplinares y mediante la articulación de principios rectores. Los siguientes párrafos de este epígrafe justifican la acotación y la precisan.

En primer lugar, el foco de la síntesis lo constituye la evaluación de la actividad investigadora dentro de las IES. Si bien la actividad científica del personal docente e investigador de las universidades guarda múltiples similitudes con la que se realiza en otros tipos de instituciones y organizaciones, la ES mantiene importantes señas de identidad que requieren un tratamiento específico. Tanto el autor del trabajo como el grueso de las fuentes bibliográficas consultadas y las experiencias internacionales sobre las que se asienta esta síntesis coinciden en transitar dentro de la ES. La aplicabilidad de estos análisis y propuestas fuera del ámbito de tales instituciones excede los objetivos de este trabajo.

En segundo lugar, el ámbito de referencia es global. Muchos trabajos sobre las reformas en materia de ES versan sobre el contenido, las motivaciones y consecuencias en ámbitos geográficos de nivel regional o estatal. Así, por ejemplo, Mollis (2014) analiza el caso latinoamericano, como Venieris y Cohen (2004) hacen respecto a Grecia; Spinoza (2008) en Chile; Susanti (2011) en Indonesia; Yogev y Michaeli (2011) en Israel; Musial (2010) en los países nórdicos; Yamamoto (2004) en Japón; Kennedy, Senses y Ayan (2011) en Turquía; Araujo (2014) en Argentina, y un largo etcétera. No obstante, no sólo estos trabajos enmarcan sus análisis desde una mirada global, enlazando las concreciones locales con los movimientos internacionales, sino que remarcan con claridad la unidireccionalidad de la relación: la principal causa de las reformas locales en materia de evaluación y de cotidianidad de la ES se sitúa fuera del ámbito local y preferentemente en motivaciones no educativas ni científicas sino económicas. En otras palabras, las publicaciones que analizan los casos locales constatan una fuerte convergencia a escala global (Codd, 2005; Filippakou y Tapper, 2008; González, 2004; Manzano-Arrondo, 2012; Puelles, 2009; Saainen, 2009; Worthington y Higgs, 2011). En consonancia, las síntesis de críticas y de propuestas se inscriben en el ámbito geográfico global y procuran evitar particularidades que sólo afectan a unas concreciones espaciales y no a la mayoría.

En tercer lugar, no existe en este documento un interés específico por centrar la síntesis en un ámbito disciplinar concreto sino en términos generales para la ES. No obstante, tal y como se muestra en epígrafes posteriores, el modelo dominante para la evaluación de la actividad científica en la ES está inspirado en las disciplinas denominadas *duras*, *experimentales* o de *ciencias naturales*, con un claro detrimento de las ciencias sociales, por lo que muchas publicaciones versan en específico sobre los prejuicios generados en éstas. Como consecuencia, parte de las propuestas surgen de los ámbitos disciplinares con menos peso actual como fuente de inspiración de las evaluaciones. Así, por ejemplo, las dificultades para medir la aplicabilidad de ciencias con un alto contenido especulativo y abstracto han propiciado sistemas de medida endogámicos, en el sentido de que versan sobre indicadores del efecto que las publicaciones poseen sobre otras publicaciones (ver más adelante). Como consecuencia, las áreas científicas cuya principal función es la aplicación social observan que su mayor fortaleza tiene un papel relativamente nulo o mínimo en el monto de las operaciones de evaluación. Al incluir estas críticas y considerar sus propuestas no se mantiene un interés específico por unas disciplinas frente a otras, sino por un panorama general equilibrado.

Por último, el objetivo de este trabajo no es resolver a nivel operativo un sistema de concreciones prácticas para la evaluación. Esta pretensión resultaría ilusa no sólo por ubicarse en una sola publicación, sino aun en las dimensiones de un artículo científico. El motivo específico radica en organizar un conjunto de *principios rectores* que sirvan para guiar concreciones operativas.

En definitiva, el objetivo es organizar las abundantes críticas vertidas sobre el sistema de evaluación de la actividad científica en las instituciones de ES a nivel global, así como proponer un conjunto de principios rectores. Los siguientes epígrafes se disponen y siguen la misma lógica. El primero organiza la síntesis de críticas. El siguiente hace otro tanto desde las propuestas, mucho menos abundantes en la práctica. Por último, el proceso facilita la propuesta final de principios.

Críticas vertidas sobre las actuales prácticas evaluadoras

Existe una amplia gama de trabajos científicos que critican el modo de evaluación que converge a nivel mundial. Comprensiblemente, cada trabajo se centra en el aspecto que más motiva a quienes lo publican. El resultado final es un *puzzle* de grandes dimensiones que podría ser organizado de formas diversas, todas ellas bien ajustadas a la experiencia acumulada. De entre

tales posibilidades, este documento ha puesto en práctica el siguiente discurso argumentativo. En el primer apartado de esta sección se explicitan las insuficiencias, imprecisiones, reducciones y arbitrariedades que subyacen al uso que se realiza del llamado *factor de impacto*; su extensa utilización permite comprender buena parte de las críticas globales y enlazar con los epígrafes que le siguen; de este modo, si tanto el factor de impacto como los productos derivados —los *rankings* de universidades— se encuentran a la vez muy aceptados y criticados, debe existir un potente argumento que justifique recurrir a estos instrumentos. De la respuesta a esta inquietud se encarga el segundo apartado, que aborda la función que la cuantificación y la ordenación cumplen en un sistema de derecho organizado en torno a la escasez. En el siguiente epígrafe se describe la denuncia de acientificidad y funcionamiento tautológico de los mecanismos de las prácticas evaluadoras, cuyos efectos se entienden a la luz de sí mismos y de los epígrafes previos; un funcionamiento de estas características deja desprotegidos institucionalmente a los individuos, que reaccionan adaptándose sin convicción. Los dos últimos apartados de la sección tratan respectivamente ambas lecturas de la esfera individual: la desprotección y la adaptación.

Limitaciones del factor de impacto

Se denomina factor de impacto a un indicador de uso documental generado por *Thomson y Reuters' Institute for Scientific Information* (ISI). La lista de inconvenientes asociados al factor de impacto comienza con su propia denominación. Un título más descriptivo —menos sensacionalista y más acorde con el espíritu científico— sería *tasa restringida de citación* (TRC). La tasa se calcula al dividir el número de citas recibidas por la revista, en los dos últimos años, entre el número de artículos computables publicados por la revista en el mismo periodo; es restringida en varios sentidos, como son el conjunto de revistas considerado para realizar los cálculos —a su vez covariante con regiones, disciplinas e idiomas— o el significado que se otorga a una cita.

Dentro de los recursos que ofrece la empresa para considerar la TRC se encuentra el informe anual denominado *Journal Citation Reports* (JCR), en forma de *rankings* de revistas por grandes áreas de interés, disponible a través de su utilidad *Web of Science*.

La TRC, los criterios de clasificación de los JCR, el monto de revistas consideradas y no consideradas, así como el uso que las administraciones realizan del conjunto han sido objeto de una intensa labor crítica. La mayor parte de esta producción se ha centrado en alguno de los supuestos que subyacen al cálculo y uso de la TRC, cuestionando la validez del conjunto. Tales supuestos y las críticas asociadas pueden sintetizarse del siguiente modo.

Supuesto 1. *Toda cita tiene un único significado: representa una unidad constante de importancia positiva. De este modo, toda cita es computada con +1.*

Este valor constante simplifica en extremo la complejidad del comportamiento de citación (Bornmann y Daniel, 2008). La redacción de un artículo científico adquiere, en cierta medida, un carácter ritual (Kreimer, 2011) que afecta, entre otros aspectos, a las muy diferentes funciones que cumple la cita en un texto. Dos citas contenidas en el mismo documento son computadas del mismo modo, aun cuando una apunte a un texto fundamental en el artículo mientras que la otra sólo cumpla una misión ceremonial (como es dar a entender que los autores del artículo están al día en sus lecturas o que conocen un texto muy citado, aunque no haya tenido efecto alguno en su trabajo). Fontela y Genaro (2004) llegan más allá, señalan que una cita puede justificarse incluso en la pésima calidad de la publicación citada, de tal manera que su peso debería ser negativo.

Estas limitaciones dificultan interpretar la importancia de una publicación concreta de acuerdo con el número de citas que ha recibido, puesto que tal número no suministra información sobre el

carácter fundamental o tangencial del trabajo, su uso directo o ceremonial, o su influencia positiva o negativa.

Supuesto 2. *Un trabajo muy citado influye bastante en la práctica, debido a que genera más impacto.*

Las críticas vertidas sobre el supuesto anterior ponen en duda este segundo punto. A ello se añaden tres argumentos específicos. En primer lugar, la relación «cita = influencia» no tiene en cuenta que buena parte del impacto real no se traduce en citas, como ocurre con los profesionales que afectan sus prácticas sin que ello derive en publicaciones (Vito, 2006), o en los casos en que los textos influyen sobre las políticas públicas o decisiones empresariales. En segundo lugar, para que un trabajo influyente se traduzca en citas, debe afectar concretamente a una línea específica de publicaciones; de este modo, trabajos que versan sobre aspectos de interés general en ciencia o en ES pueden influir en alto grado pero no generar citas; pongamos por caso una publicación que suministra la clave del éxito para que los trabajos de investigación sean aceptados en las revistas: quizá llegue a ser leído por todo el mundo investigador, pero prácticamente nadie citará ese texto. En tercer lugar, la velocidad y el modo de un trabajo que germina en citas también muestran una amplia diversidad entre campos de estudio y disciplinas, de tal modo que sólo una observación muy dilatada puede arrojar luz sobre el peso real de un trabajo en un área concretos. Al respecto, Krell (2002) destaca que en algunas áreas de investigación la influencia real de un trabajo sólo se observa pasados tantos años que en ocasiones los autores ya perecieron cuando su publicación alcanza una cota estable y elevada de citas.

Supuesto 3. *Las revistas pueden ordenarse según un criterio de calidad: una revista es tanto mejor cuanto más se cite.*

Diversas investigaciones ponen en tela de juicio la relación entre la calidad de una revista y el número de citas que recibe, a la vez que suministran evidencias empíricas sobre el efecto autoalimentado de la TRC: este índice, en cierto grado, parece medirse a sí mismo más que a la calidad de la publicación a la que se refiere. Así, por ejemplo, Callaham, Wears y Weber (2002) consideraron los 204 trabajos de un congreso de medicina que fueron publicados en revistas con TRC; mediante grupos Delphi, analizaron la calidad de los trabajos y obtuvieron también medidas objetivas metodológicas; a la hora de predecir el número de citas que recibieron las 204 publicaciones, más de tres años después, encontraron que la única variable explicativa relevante fue la TRC de la revista donde se publicó, una vez extraída la influencia de los otros dos factores. Luego, citar un trabajo no parece relacionado con su calidad, sino con la TRC de la revista donde se publica.

En la misma línea, Lange (2002) investigó las citas de dos revistas sobre educación: la más leída de ambas no estaba contemplada en JCR; la otra, sí; la autora descubrió que las referencias a la revista no-JCR eran «confundidas» y asignadas a la sí-JCR por el ISI, debido a la similitud de sus nombres, error que incrementó el número de citas indebidamente asignadas y también las correctamente recibidas.

Supuesto 4. *El criterio de calidad de una revista es trasladable a sus unidades: un artículo es mejor que otro si se ha publicado en una revista con más citas.*

Como señala, entre otros, Verma (2015), la mayor parte de la TRC de una revista proviene de un número muy reducido de artículos, mientras que el grueso tiene un monto bajo de citas, de tal forma que no es posible calificar una contribución por la TRC del soporte donde se publica.

Supuesto 5. *La calidad del curriculum investigador de una persona concreta es proporcional al número de artículos que ha publicado en revistas muy citadas.*

Este supuesto se organiza desde la acumulación de los anteriores, de tal forma que los sesgos también se acumulan, generando un margen de error excesivo como para asumir la relación entre las variables «calidad de la trayectoria individual de investigación» y «monto de artículos con TRC».

A los efectos anteriores se añaden otros, específicos para este supuesto. Uno de los que debería resultar más sorprendente desde la perspectiva de las IES, en especial para las públicas, es el alto coste que acompaña con frecuencia la publicación de un trabajo en revistas con TRC. Parte del dinero que los Estados dedican a la realización de investigaciones en las IES termina en las cuentas de empresas dedicadas al negocio de las publicaciones científicas. Como se denuncia en [Manzano-Arrondo \(2015a\)](#), los Estados pagan no sólo para formar a personal investigador y para mantener el edificio científico de las universidades, sino para tareas que podrían desempeñar los miembros de la institución si fueran estimuladas por los estándares de evaluación de la actividad científica, pero que ahora desempeñan empresas externas. Así, el Estado paga también para que su personal académico pueda consultar bases de datos de resúmenes, para leer revistas donde se publican los trabajos realizados por el propio personal, para traducir al inglés los documentos que optan a ser publicados e, incluso, para publicar en revistas con TRC. Según el estudio de [Björk y Solomon \(2014\)](#), el pago por publicar artículos generó un monto de 182 millones de dólares en 2012, con un crecimiento anual del 34%, y la estimación de que cubrirá el 20% de las publicaciones científicas antes de 2020; el gasto medio por publicar un artículo en abierto, dentro de una revista de acceso restringido —es decir, que se mantiene por suscripciones—, se acerca a los 3,000 dólares, según el mismo estudio. Según [Fuentes \(2014\)](#), existe una fuerte tensión entre la tendencia ingenua de la ciencia —compartir sin trabas los resultados de investigación— y el interés privado por obtener beneficios que protagonizan las empresas del sector. A ello se suma la tendencia de los Estados por controlar la producción científica a través de políticas públicas ([Jarab, 2008](#)). No obstante y a la luz de la cifras de negocio del sector de las publicaciones y de las normas en política científica que llevan al personal de investigación a centrarse en las publicaciones privadas, la lucha entre las tres fuerzas parece estar decantándose hacia el aprovechamiento capitalista de la producción científica ([Dávila, 2014](#)). Si bien es comprensible que el interés privado busque vías para obtener beneficios, el interés público debería proteger los fundamentos del progreso social, entre los que se encuentra generar y compartir conocimiento sin desequilibrios fundamentados en las posibilidades de renta.

Las denuncias sobre los costes de publicación no implican que necesariamente todo *curriculum* de éxito se ha construido mediante el gasto de importantes sumas de dinero; sin embargo, afecta de manera directa a los supuestos de un modelo que se edifica desde el único criterio teórico de la valía investigadora, sin considerar las posibilidades económicas de quien investiga o de la institución en la que se investiga. Incluso no sólo habría que considerar que «no todo el mundo puede permitirse pagar los gastos de poner en abierto su publicación», sino que «no todo el mundo considera moral este comportamiento» y, por ende, no lo lleva a cabo. Por otro lado, la constancia de estos gastos hace tambalear el supuesto de escasez que se aborda en un apartado posterior: el sistema actual de evaluación se sustenta en el interés de los Estados por administrar de forma óptima recursos económicos siempre escasos.

A pesar de la debilidad de este quinto supuesto, y como denuncian [Gómez y Jódar \(2013\)](#) respecto a un número creciente de países, ocurre que el acceso a puestos de docencia o de investigación, la promoción laboral, la obtención de recursos de financiamiento para proyectos

académicos, la oportunidad para dirigir tesis doctorales o de contar con becas adscritas de personal investigador, los sueldos, las posibilidades de gestión académica, la participación en programas de doctorado, en tribunales, etc., son algunos de los asuntos que se ven afectados por el número de publicaciones que los miembros de las instituciones de investigación realizan en la base ISI, sobre todo si las revistas se encuentran en los primeros puestos de alguno de los *rankings* considerados.

En definitiva, la TRC mide una cosa distinta a lo que supone su utilización como materia de evaluación, pues más que calidad expresa visibilidad (Granda, 2003), difusión (Buela-Casal, 2003) o pertinencia de inversión bibliotecaria (Garfield, 2007), todo ello mediante un importante sesgo de cobertura que afecta a revistas, disciplinas, idiomas, regiones e intereses de investigación de forma muy dispar (Aguado, Rogel, Garduño y Zúñiga, 2008; Fernández-Quijada, Masip y Bergillos, 2013; Fischman, 2013; Matías, 2013; Pérez, 2006). A pesar de ello, ha llegado a convertirse en el capital académico máspreciado para las instituciones de todo el mundo (Beigel, 2014).

Reducción métrica y rankings

Diversas críticas se centran en lo que Gary Hall —en entrevista publicada por Broekhuizen, Dawes, Mikelli y Wilde (2015)— denomina «metricidación académica»: la obsesión por recurrir a métricas cuantitativas como fundamento para la evaluación y la interpretación de la actividad del personal académico.

En términos generales, esta obsesión lleva con facilidad a simplificar lo complejo, con la justificación de que esa operación resulta viable, aunque se cuantifiquen aspectos de poca relevancia o queden sin abordar otros, tan difícilmente cuantificables como relevantes (Báez, 2012). Robinson (2014) concluye que los sistemas de evaluación pueden ser descritos como una fuerte restricción cualitativa (reducción de un universo amplio a una categoría casi exclusiva), sobre la que se aplica una fuerte motivación cuantitativa (protagonismo del volumen de publicaciones). Como resultado, procesos apenas tangibles se concretan en estrategias de medida objetivables y homogeneizadoras (Egido, 2005), abandonando las dificultades de valorar el ideal para concentrarse en las posibilidades prácticas (Urcelay y Galetto, 2011). De este modo, el objetivo de traducir en números algo tan complejo como la actividad investigadora mueve a sumar notables concesiones e imprecisiones en todas las fases del proceso (Manzano-Arrondo, 2010). Entonces, si el modelo es o parece ser tan defectuoso como señalan las críticas acumuladas, ¿por qué está en uso? El éxito no se apoya en la validez del procedimiento para estimar los aportes sino en otras variables.

Cuando el contexto de referencia puede definirse como situación de escasez, en el sentido de que la demanda de recursos es superior a su disponibilidad u oferta, se crea un problema de justificación del éxito, pues no todo el mundo —individuos e instituciones— puede saciar el objetivo de acceder a los recursos. Si este contexto se inserta en una sociedad de derecho, las decisiones que adopta el sistema deben quedar muy justificadas, no pueden ser arbitrarias y mucho menos subjetivas. El uso de *rankings* es una salida sencilla para esta situación, con la triple condición de que el procedimiento sea viable, objetivo y adaptable a los recursos.

La viabilidad del *ranking* se sitúa en su carácter aritmético: opera con cálculos sencillos a partir de la cuantificación de lo que se evalúa. Si esta cuantificación es más o menos adecuada, acertada o pertinente, o si lo que se evalúa se corresponde o no con lo que debería evaluarse, son aspectos que no afectan al hecho de que el procedimiento consiga sumar las cuantías de resultados de medidas para cada elemento y provea un listado ordenado de éstos, según la suma obtenida.

La objetividad se sitúa en el uso del instrumento: una vez asumido que el *ranking* es una buena herramienta para tomar decisiones, cualquier individuo, organización o agencia puede operar con

él, obteniendo los mismos resultados. De este modo, la agencia queda protegida ante la posible acusación de parcialidad o de subjetividad en las decisiones.

La adaptación a los recursos disponibles implica que primero se disponen los recursos y después se van repartiendo bajo el orden en que se acomodan los elementos del *ranking*. Un sistema donde cualquier individuo o institución que cumpliera unos criterios tuviera acceso a los recursos dejaría al agente distribuidor desprotegido y sin capacidad de previsión respecto al monto final de recursos requerido.

Así pues, las administraciones públicas y demás entes sociales que cuentan con un conjunto escaso de recursos, a repartir entre un número superior de demandantes, agradecen contar con instrumentos que, con independencia de su validez científica, resuelvan el problema de la repartición con sencillez y objetividad. En este discurso no participa el grado en que la sencillez ha requerido reducir una realidad muy extensa, el grado en que la objetividad ha necesitado homogeneizar en extremo una realidad compleja, ni el grado en que la escasez pueda ser provocada, que tuviera solución por otras vías, o que incluso se vea incrementada debido a los gastos asociados al procedimiento.

En el epígrafe anterior se ha mencionado a los *rankings* de revistas que se publican vía JCR, en función de la TRC. Los *rankings* de investigadores por áreas específicas o países protagonizan también un hábito cada vez más frecuente. Sin embargo, el tipo de ordenación más conocido incluso por el público en general es el que opera con las IES.

Un *ranking* de universidades es una clasificación unidimensional de IES que se genera como resultado de aplicar una combinación de variables, según criterios diversos (Bernasconi, 2006). La combinación considera por lo regular criterios objetivos —como el presupuesto institucional— y subjetivos —como el prestigio— (Martínez, 2011); este mismo autor denuncia que los *rankings* más utilizados —especialmente el de Shangai— son los más pobres, es decir, los que cuentan con procedimientos de peor calidad para establecer sus cálculos, muy basados en las actividades de investigación y, dentro de ellas, en la TRC. Por ello, es habitual referirse a estos instrumentos como herramientas demasiado burdas, imperfectas, simples, a-científicas, poco fundamentadas e incluso, en ocasiones, fraudulentas, como para considerarlas herramientas serias en la toma de decisiones sobre la distribución de fondos o el prestigio de individuos e instituciones (Fernández, Martínez y Velasco, 2006; Greenwood, 2012; López, 2013; Morley y Aynsley, 2007; Sanz-Casado, García-Zorita, Serrano-López, Efraín-García y de Filippo, 2013; Usher y Savino, 2006). A pesar de ello, se anuncia que han llegado para quedarse (Villaseñor, Moreno y Flores, 2015).

Debido al creciente protagonismo que adquieren, generan poco a poco más efectos presupuestarios, que a su vez afectan a la cotidianidad de las instituciones. Como consecuencia, y de acuerdo con una expresión de Enders (2015), se observa ya una verdadera «carrera armamentística» de alto coste, en la cual las universidades desvían esfuerzos humanos, temporales y monetarios hacia estrategias de marca y de posicionamiento que deterioran o se mantienen al margen de sus funciones básicas.

Lo más relevante de los *rankings* de universidades, a los objetivos de este documento, es el modo como son reforzados. Según describe Stack (2016), lo habitual es que las universidades desprecien la validez de los *rankings* a la vez que se insertan en ellos, afianzándolos, por el temor a las consecuencias que podrían derivarse de no ser visibles en estos instrumentos demasiado mediatizados. De nuevo, el instrumento se explica a sí mismo: el simple intento de las universidades para salvarse de los *rankings*, procurándose una buena posición, es parte fundamental del refuerzo que reciben estas ordenaciones, otorgándoles un poder que no desean quienes participan.

Centralismo y lógica cíclica acientífica

El actual sistema de evaluación tiene origen en el mundo anglosajón. Se ha denunciado en múltiples ocasiones (véanse, por ejemplo, los trabajos de [Aguado-López, 2013](#); [Arnett, 2008](#); [Benítez, 2012](#); [Deulofeu, 2009](#); [Gasca-Pliego y Olvera-García, 2011](#); [Madorrán, 2012](#); [Martín-Baró, 1986](#); [Quirós, 2014](#), o [Vázquez, Escabí, Quiñones y Pacheco, 2012](#)) el modo como se asumen y refuerzan, a escala global, los estándares, criterios, medios y motivaciones de investigación generadas y gestionadas desde el centro del primer mundo. No resulta extraño, por tanto, que el planeta científico utilice conceptos, hallazgos y revistas con un mismo origen. [Kreimer \(2011\)](#) se refiere a este fenómeno con la expresión *integración subordinada*, y subraya que la subordinación es una condición para no quedar atrás respecto al resto del mundo científico. Esta dinámica, además de suponer serios problemas para el ámbito regional, implica un grave costo de oportunidad para la ciencia, al prescindir en igualdad de condiciones de aportes generados mediante estándares, criterios, medios y motivaciones de naturaleza más diversa.

Un apartado específico del centralismo lo constituye el idioma, inserto en lo que [Alatas \(2003\)](#) denomina *neocolonialismo académico*. Pareciera que el personal académico de todo el planeta debe comunicarse utilizando el inglés porque el inglés *es el idioma de la ciencia*. De este modo, el uso del idioma entra en un mecanismo cíclico: se utiliza porque se parte del supuesto de que todo el mundo lo utiliza, lo que aumenta su uso. Es más, como afirman [Meneghini y Packer \(2007\)](#), publicar en inglés ha terminado por convertirse en un valor por sí mismo, con independencia de la calidad de lo que se publica.

Sin embargo, resulta pertinente hacer algunas acotaciones. En primer lugar, el supuesto de que el inglés es ya un idioma ampliamente compartido no es cierto; por ejemplo, [Gazzola \(2014\)](#), a partir de encuestas europeas, evidencia un bajo nivel de uso del idioma en la población general europea, de tal forma que su consideración oficial como lengua franca excluiría a un amplio sector de la población; a su vez, [Ansele \(2017\)](#) alerta sobre el alto porcentaje de trabajos relevantes que no son publicados en inglés y a los que el personal científico que sólo lee en este idioma no tiene acceso; en un campo inmerso en este idioma, como es la biología, se calcula en un 36% los trabajos no accesibles para anglolectores, porcentaje muy superior en ciencias sociales. En segundo lugar, la apuesta por el inglés como lengua franca de la ciencia tiene efectos perjudiciales en diversos ámbitos, uno de ellos es el perjuicio local generado paradójicamente por la intención manifiesta de promover el progreso local; así, [Hanafi \(2011\)](#) muestra que las universidades árabes terminan por optar hacia la proyección internacional o la implicación local, utilizando respectivamente el inglés o el árabe, y observando la incompatibilidad práctica entre ambas opciones; en un estudio similar respecto a los países asiáticos, [Majhanovich \(2014\)](#) describe el aumento de la desigualdad en universidades y ciudadanía, tras la apuesta del inglés como idioma preferente para la educación en materias de economía y ciencia, creándose una nueva élite asociada a lo que denomina «el idioma de la Reina»; otra línea de perjuicios se asocia con las dificultades inherentes al inglés como segunda lengua debido a la abundancia de irregularidades, la conexión ambigua entre el lenguaje escrito y el hablado, la dificultad de su sistema de sonidos en el panorama internacional de lenguas, y a la fuerte asimilación cultural con que va aparejado su aprendizaje ([Koutny, 2005](#)). En tercer lugar, co-existe un problema de justicia lingüística de enorme peso en la elección del inglés como idioma internacional fáctico ([Alcalde, 2016a, 2016b](#)) al generar una notable desigualdad de oportunidades en función del idioma materno. Por último, la implantación del inglés como idioma de la ciencia no se respalda mediante argumentos científicos sino, como afirman [Lacey \(2015\)](#) y otros autores mencionados en su trabajo, la justificación se inserta en los procesos de hegemonía económica, política y militar. De este modo, el inglés se promociona como el idioma internacional

a partir de diversas operaciones en las tres dimensiones mencionadas, muy en particular desde la Segunda Guerra Mundial (Chorg-Shing, 2003) y como referente capital en los procesos de globalización intensificados cuatro décadas más tarde (Majhanovich, 2014).

El centralismo científico se comporta como un agujero negro que aumenta su capacidad de atracción conforme engrosa su masa. Las normativas de los diferentes Estados en materia de universidades, de ciencia y de evaluación se ven absorbidas por este proceso de atracción, bajo la denominada *convergencia internacional*. La búsqueda de esta convergencia se centra entonces más en reproducir lo que se va consolidando que en el contenido en sí de lo que se consolida (Amigot y Martínez, 2013; Manzano-Arrondo, 2015a). Por ello, no es la sustancia de la propuesta lo que motiva la adopción de las medidas, sino los comportamientos observados en terceros o en referentes externos (Gracia, 2014; Urcelay y Galetto, 2011). Como resultado, el mecanismo se vuelve cíclico.

El modelo parte de una creencia implícita en la mano invisible académica (Manzano-Arrondo, 2012), basada a su vez en el individualismo metodológico que considera suficiente al sujeto como la explicación del conjunto (Noguera, 2003; Evers y Lakomski, 2013): la ciencia mejora en la medida en que se faciliten las vías para que sus miembros sacien de modo individual su ambición de éxito, de promoción o de supervivencia, adaptándose a los sistemas de premios y de castigos que se derivan de los procesos de evaluación. El motor de la ciencia no es entonces la búsqueda de la verdad o el interés por incrementar desinteresadamente el conocimiento, sino el auto-gerencialismo de las trayectorias individuales, centradas no en el saber, sino en la producción. La suerte de cada individuo queda sujeta al grado en que consiga obedecer a estos criterios.

A la luz de lo anterior, resulta obvio que la impronta de los sistemas de evaluación no es científica. El corazón de la ciencia se define en términos de colaboración, no de competición; de amor honesto al saber y no a los réditos individuales; de interés por incrementar el conocimiento y no por la producción de publicaciones. La ciencia tiene mecanismos de calidad, que llamamos más bien *validez*, que le han permitido desempeñar una función fundamental en la sociedad y gozar de altas cotas de prestigio y de credibilidad. El interés, y más aún la presión, por evaluar la producción científica no queda justificado por la motivación de mejorar la ciencia, sino que proviene y aterriza fuera de ésta, inserto en los procesos de centralismo y de convergencia mencionados.

Un aspecto más que fortalece esta conclusión es la ausencia de meta-evaluaciones. En la impronta científica existe un marcado interés por comprobar empíricamente modelos y teorías. Por el contrario, se ha denunciado en numerosas ocasiones que los modelos de evaluación no aparecen como objeto de este mismo tratamiento. No se observa una preocupación relevante por valorar las consecuencias de estas dinámicas. No hay una cultura de meta-evaluación (Alderman y Brown, 2005; Miguel y Apodaca, 2009), ni meta-evaluación significativa (Gaytán, 2012; Martín, Conde y Baute, 2012), sino más bien ímpetu evaluador basado en prácticas que se imitan sin ánimo de comprobación (Arrabal y Fernández, 2012). Los conatos de meta-evaluación tienden a la tautología: sea cual fuere el destino D al que se pretende llegar, se establece que D es alcanzable mediante la productividad científica P , medida en los términos descritos antes; el modo de comprobar si la relación « $P \rightarrow D$ » es correcta consiste en afirmar que P está incrementándose realmente (véase, por ejemplo, Moreno-Pulido, López-González, Rubio-Garay, Saúl y Sánchez-Elvira-Paniagua, 2013). El autor del presente trabajo no ha encontrado ninguna publicación académica, ni informe de administraciones públicas, que defina D con la misma precisión métrica utilizada en la evaluación y muestre con evidencias que está mejorando debido al incremento en P .

La acientificidad de los procesos de centralismo y de convergencia que subyacen a la evaluación estimula diversos ciclos de inclusión y de exclusión que afectan a las revistas, a los motivos de investigación, a los individuos y a las instituciones. Respecto a las revistas, los trabajos

mencionados de Callaham et al. (2002) y Lange (2002) muestran que la TRC se explica a sí misma. Los individuos leen de preferencia las revistas con altos puntajes TRC con el objetivo de publicar en ellas, citando sus trabajos no tanto por la calidad de los mismos como por el interés de familiarizarse con el medio en el que desean publicar. Este efecto tiene consecuencias también sobre las líneas y motivos de investigación, y priman aquellas que tienen mayor probabilidad de ser publicadas en revistas de elevado TRC.

Respecto a las trayectorias individuales, la publicación de trabajos con TRC facilita la obtención de subvenciones a la investigación, con las que no sólo se generan investigaciones publicables, sino que proveen fondos para cubrir los costes de publicación en revistas con TRC, como los ya mencionados pagos por publicar y los derivados de las traducciones. Contar con ese tipo de publicaciones y las subvenciones asociadas genera un efecto «llamada» para investigadores en formación, quienes llevan a cabo sus tesis doctorales en tales equipos, incrementando las posibilidades de nuevas publicaciones con TRC y la obtención de subvenciones. En algunos países, como ocurre en España, si un miembro del personal docente e investigador se encuentra inserto en este ciclo, se le *descarga* parcialmente de la docencia, de tal forma que puede dedicar más energías y tiempo a la tarea de producir artículos. En definitiva, las trayectorias individuales divergen en dos ciclos: quienes entraron en la mecánica descrita y quienes, quedando fuera, les resulta difícil entrar, pues tautológicamente sólo pueden entrar quienes ya están dentro.

Esta lógica cíclica afecta también a las instituciones insertas en los *rankings* de universidades. Las que están mejor posicionadas actúan como atractores de individuos y de recursos que potencian su posición, mientras los mecanismos puestos en marcha para generar los *rankings* tienden a perpetuar los resultados. Así, por ejemplo, se ha denunciado que en la elección sobre los criterios que van a tenerse en cuenta en la elaboración de un *ranking* tiene un peso fundamental que los resultados no se alejen de lo que proponen otros productos de ordenación (Usher y Savino, 2006). De este modo, las primeras universidades son siempre las primeras.

En definitiva, los mecanismos de evaluación muestran un marcado comportamiento cíclico, tautológico y no coherente con los criterios que han caracterizado el desarrollo de la ciencia hasta la fecha.

Desprotección de los individuos

Los apartados previos dibujan una situación poco halagüeña: en el corazón del sistema de evaluación se encuentra el protagonismo de la TRC con abundantes limitaciones y supuestos que no se cumplen, a lo que siguen reducciones métricas centradas en la viabilidad más que en la coherencia con los postulados de la ciencia, de tal forma que el resultado final tiene poco que ver con ella. En este contexto, los individuos se ven presionados para adaptarse bajo la amenaza de ingresar en los ciclos de exclusión.

Aunque parezca trivial que una fuerte presión derive en consecuencias individuales negativas, la literatura científica se ha ocupado también de este asunto, por ejemplo: 1) las entrevistas al personal académico muestran un creciente desencanto y pérdida de la confianza en sus instituciones (Shields y McGinn, 2011); 2) los individuos que no cumplen con las crecientes exigencias son devaluados (Malecki, 2000); 3) aumenta el individualismo y el credencialismo en el mundo académico (Shumar, 2004), donde los individuos se transforman en «fundamentalistas del éxito» (Balestena, 2001), lo que deriva en una notable reducción de las investigaciones colaborativas (Greenwood, 2012) y de la cohesión (Macfarlane, 2005); 4) se denuncian efectos en términos de amargura, sensación de persecución pública, frustración en la consecución de objetivos cada vez más difícilmente alcanzables, cansancio e inseguridad (Ball, 2003); 5) derivación frustrante de

energías hacia tareas de gestión de las evaluaciones internas y externas poco o nada reconocidas (Galán, González y Román, 2012); 6) presión por auto-gerenciarse, mutando continuamente sin posibilidades de estabilidad psicológica (Gómez y Jódar, 2013; Grinberg, 2009), y 7) incluso se describe una creciente incidencia académica de alcoholismo, depresión y suicidio (Shaw y Ward, 2014).

En el otro extremo y frente a los racionalismos universalistas, la ética del cuidado se presenta como defensora a ultranza de las circunstancias particulares, como garantía de que ninguna norma orientada a una ética generalista pueda dañar a los individuos en su cotidianidad (León-Correa, 2008). Se trata de preocuparse por el otro, no en abstracto, sino en concreto; de ofrecerle ayuda y de atender su petición de ayuda (Albarado, 2004). La ética del cuidado se inserta en visiones más generales, como la ética de la salud (Jurdi y Amiralian, 2013) o la planetaria (Boff, 2001), sin que por ello deje de apuntar con claridad hacia la protección de las personas concretas. De ello se deriva que no es de recibo presionar a unos individuos —el personal académico— para un supuesto beneficio social del que no forman parte y que ha sido definido desde una visión abstracta de la realidad. Cuidar a los propios miembros de la institución, sin confundir esta práctica con el corporativismo ni el privilegio, constituye uno de los pilares del poder académico (Manzano-Arrondo, 2015a), sin el cual es difícil garantizar que la institución científica pueda ejercer un trabajo útil a la sociedad. McArthur (2011) es muy claro al respecto: para que la universidad pueda ejercer su labor de progreso de la sociedad, es necesario que sus miembros sientan en ella un contexto de protección desde el que puedan ejercer observación, reflexión, crítica y ensayos sin que se vean obligados a protegerse más a sí mismos que a cumplir la misión que tienen encomendada.

Empobrecimiento adaptativo

La universidad no es cualquier institución. En su seno trabajan, supuestamente, individuos con una alta capacidad intelectual, acceso privilegiado al conocimiento y posibilidades para elaborar juicios bien fundamentados. Los apartados previos muestran un proceso unidireccional de reforma, donde las directrices globales son implementadas a escala local y los miembros de la academia sienten con frecuencia que han quedado a merced de criterios que no parecen justificados por argumentos científicos o académicos. La claridad de esos criterios facilita la adaptación, mientras que la capacidad de juicio y la unidireccional estimulan la falta de convicción. El resultado es un personal que se adapta pero que no cree.

Se han propuesto diversas acotaciones para referirse a este fenómeno. Así, en una investigación con miembros de la universidad de diversos países, Teelken (2012) encuentra que existe el acuerdo generalizado en la necesidad de evaluar, tanto como desacuerdo respecto al modo en que se está llevando a cabo; por lo general, se asumen los cambios pero no la creencia de que ello implique una mejora de la calidad; la adaptación tiene lugar a un «nivel cosmético o superficial» (Teelken, 2012: 278). En la misma línea y también a través de un estudio empírico, esta vez sobre el comportamiento de investigadores de éxito, Cheng (2011) encuentra en las narrativas de las entrevistas que estas personas no creen en las reglas del juego pero se adaptan a los criterios de evaluación para sobrevivir, en lo que denomina un «jugar el juego». Temple (2005) interpreta lo que está ocurriendo en términos de *moda*; describe el modelo de excelencia académica al uso como fruto de procesos emocionales y cíclicos, donde concurren imitación, entusiasmo inicial y frustración posterior. Por último, Kreimer (2011) habla de *racionalidad burocrática* para referirse al fundamento de las actividades de evaluación de la ciencia y recalcar, con ello, que los individuos

no actúan con el ánimo de contribuir al conocimiento científico, sino con el de responder de forma pertinente al funcionamiento del sistema.

Sea burocracia, moda, juego, adaptación cosmética o supervivencia, las descripciones sobre el modo como se implementan las reformas a partir de los sistemas de evaluación atacan los supuestos de un modelo de excelencia. Es ya una constante que los manuales sobre promoción de calidad adviertan que las prácticas no han de quedar reducidas a un sistema de premios y castigos, sino que deben obedecer a un principio fundamental: afianzarse mediante una sólida cultura organizacional que contagie a todos los miembros de la institución (Winn y Green, 1998). Una cultura de calidad no se nutre de nuevas leyes o normas, sino de la implicación sincera de cada miembro del colectivo (Jáuregui, 2004). La extendida falta de convicción, que ha sido descrita, se encuentra en las antípodas de este principio.

Esta situación no afecta sólo a la falta de convicción, sino que menoscaba la honestidad. Fundamentar la evaluación de los individuos en el conteo de artículos implica partir de la equivalencia «investigación = artículo». Esta suposición reduce el significado de ambos lados de la igualdad y estimula estrategias para incrementar de manera artificial el número de aportes. Algunas de ellas son practicar auto-plagio (Robinson, 2014), trocear los resultados (Cerezo, 2007) o intercambiar autorías regaladas (Collado y Vázquez, 2006). En esta línea, no es raro que la literatura sobre los comportamientos fraudulentos asocie las conductas deshonestas, en ciencia, con las presiones que se viven para conseguir fondos y para la adaptación a estándares que requieren una alta velocidad de producción (Fernández-Quijada et al., 2013; Robinson, 2014). Entre las presiones y el fraude, se indica como variable intermedia la reducción de la garantía que supone participar del espíritu colectivo de construcción de conocimiento (Irigoyen, 2011), a cambio de la búsqueda de buenos resultados individuales que permitan superar con éxito las presiones (Cerezo, 2007).

Finalmente, y sin llegar al extremo de los comportamientos fraudulentos, se denuncia también que los procesos de adaptación a las evaluaciones positivas están reduciendo la calidad de las investigaciones científicas mediante la pérdida de creatividad, de innovación y de riesgo en las propuestas (Argullol, 2014; García-Quero, 2014). El mencionado trabajo de Cheng (2011), a través de entrevistas focales a investigadores con altas puntuaciones en los procesos de auditoría y de acreditación, encuentra que buena parte de las personas confiesan que han renunciado a la creatividad a cambio de publicaciones fáciles y rápidas en revistas. Hallazgos similares encuentran Valero, Jiménez y Coca (2013) al analizar los niveles de innovación en publicaciones científicas de investigadores en función de su consolidación. Alderman y Brown (2005) sostienen que las prácticas de evaluación lastiman la propia calidad de las instituciones de investigación, al moldear el comportamiento de éstas hacia los aspectos competitivos más que de construcción del conocimiento. Manzano-Arrondo (2012) describe que las aportaciones más trascendentes provienen de procesos lentos, creativos e innovadores, con regularidad en manos de personas con un perfil multidisciplinar y, por tanto, menos productivo que el basado en una alta especialización. Es más, buena parte de estos ejercicios propositivos adoptan la forma de libro o de conferencias, formatos que se debilitan ante el protagonismo de la revista (Fernández-Quijada et al., 2013).

Algunas propuestas de mejora

La abundancia de críticas sobre las prácticas de evaluación no debe interpretarse como una imagen proporcional a los defectos de los estándares. Es posible que los errores estimulen más que los aciertos. A pesar de esta salvedad, las publicaciones evidencian que los sistemas actuales son mejorables y conviene establecer nuevas prácticas. Las argumentaciones en torno a que se

trata de un modelo imperfecto pero no tenemos nada mejor (Hoeffel, 1998) ni tan fácilmente disponible (Granda, 2003) tienen poca relevancia científica.

En lo que sigue, sin resolver de forma exhaustiva el asunto, se exponen algunos movimientos que, de algún modo, sugieren cambios; quedan organizados en dos apartados, en función del grado de transformación con que pretenden afectar a las prácticas actuales, y ponen mayor énfasis en el segundo, que aborda las propuestas de cambio paradigmático.

Matizaciones al modelo

Buena parte de las propuestas se encuentran insertas en el modelo de referencia, es decir, el grueso de la evaluación considera: 1) la influencia que unas publicaciones tienen sobre otras; 2) la medida a través del uso de citas, y 3) el interés, principal o secundario, de juzgar trayectorias individuales. Las propuestas se ofrecen para corregir aspectos relacionados con la cobertura de publicaciones, el modo de computar citas, el procedimiento para calcular impactos, las bases de referencia a utilizar, etc. En ningún caso se pone en duda el modelo en sí.

Desde el ámbito mercantil han surgido diversas iniciativas que aspiran a copar parte del mercado de las publicaciones científicas y de sus productos derivados, especialmente el acceso a bases de datos y la obtención de métricas de uso documental. La empresa Elsevier ofrece *Scopus*, un producto similar a la *Web of Science* de Thomson y Reuters, pero con mayor cobertura en la consideración de revistas revisadas por pares, muy en particular en los mercados menos atendidos por la empresa rival (Santa y Herrero-Solana, 2010). *Google Scholar* ofrece dos cambios más significativos: el acceso gratuito a múltiples contenidos y la aplicación de su motor de búsqueda para localizar los documentos y para calcular métricas de citas, si bien todavía cuenta con imprecisiones importantes (Torres-Salinas, Ruiz-Pérez y Delgado-López-Cózar, 2009).

A las iniciativas privadas se le suma una amplia oferta de proyectos públicos exitosos, en su mayoría originados en Latinoamérica, que ofrecen tres soluciones: 1) corregir sesgos de cobertura regional, de idioma y de disciplina; 2) fomentar el acceso abierto al contenido íntegro de las revistas consideradas, y 3) manejar un conjunto extenso y transparente de criterios científicos de inclusión de revistas en la base de datos. En esta línea destacan Redalyc (Becerril-García, Aguado-López, Rogel-Salazar, Garduño-Oropeza y Zúñiga-Roca, 2012) con acceso libre a 535,290 artículos completos¹; SCIELO (Packer, Luccisano, Montarani, Santos y Meneghini, 2006) con 573,525 artículos y más de 13 millones de citas²; Latindex (Flores, Penkova y Román, 2009), con un directorio de 25,353 revistas —de las que 8,949 se incluyen en el catálogo de calidad— y 36 países participantes en el proyecto³; o Dialnet (León y Magriñá, 2004), que permite el acceso a más de 10,000 revistas y 5 millones de documentos, entre los que se encuentran 68,442 tesis doctorales⁴.

Cambio de perspectiva: declaraciones de principios DORA e iniciativas previas

En diciembre de 2012, en San Francisco, un grupo de editores reunidos en la cita anual de la Asociación Americana de Biología Celular redacta un documento crítico sobre la evaluación de la

¹ Consulta realizada el 6 de abril de 2017 en la web www.redalyc.org.

² Consulta realizada el 6 de abril de 2017 en la web www.scielo.org.

³ Consulta realizada el 6 de abril de 2017 en la web www.latindex.org.

⁴ Consulta realizada el 6 de abril de 2017 en la <https://dialnet.unirioja.es>.

actividad científica, documento conocido como DORA⁵ (*San Francisco Declaration of Research Assessment*), provocando una repercusión que ha excedido con mucho el marco en el que se generó. A modo de síntesis específica para este trabajo, la declaración establece que, para evaluar la calidad:

1. De un artículo, hay que prescindir de la TRC de la revista y de cualquier otra estrategia similar.
2. De una revista, hay que acudir a estrategias multidimensionales, combinando índices diversos.
3. De la actividad científica a través de sus resultados, no hay que ceñirse sólo a las publicaciones. Debe suponerse que:
 - a. Es necesario considerar otros productos, como las bases de datos o el software.
 - b. Hay que superar el corsé del interés de la propia comunidad científica y utilizar también indicadores de impacto político y social

La recomendación 1 es habitual en la literatura, como ya se ha mencionado; la 2 se encuentra en la línea de otros trabajos que plantean la combinación de recursos diversos para contar con una imagen más compleja y real de la situación (por ejemplo, [Filippakou y Tapper, 2008](#); [Irigoyen, 2011](#); [Morley y Aynsley, 2007](#)); la 3a es considerada ya en algunas pautas de evaluación, donde se han incluido la dirección de tesis doctorales, la generación de patentes, etc., a lo que puede añadirse la consideración de formatos abiertos como documentos pdf, páginas web, blogs o vídeos ([Vázquez-Cano, 2013](#)); la recomendación 3b constituye, sin lugar a dudas, el mayor reto en el campo de la evaluación de la actividad científica. Tampoco es nueva. Son continuas las voces que reclaman el necesario protagonismo de los indicadores de impacto social para la ciencia y la academia ([Aguado et al., 2008](#); [Boni y Gasper, 2011](#); [De la Corte, 2004](#); [Ellacuría, 1999](#); [Liu, 2011](#); [Martin, 1993](#); [Scobey, 2011](#); [Sívori, 2012](#); [Speck, 2001](#); [Vallaeys, 2010](#)).

La principal virtud de DORA no es aportar críticas y recomendaciones originales o novedosas, sino hacerlo desde el centro del centro, es decir, desde el ámbito científico anglosajón, por parte de una disciplina con amplios réditos derivados de los actuales usos de la TRC y avalada por referentes de éxito en el mismo ámbito. Numerosas iniciativas previas habían ya señalado la urgencia de abordar sistemas de reconocimiento del impacto político y social; así ocurre, por ejemplo, con la declaración con que se inició la Red de Talloires⁶ de universidades en 2005, en la que se indica que el compromiso social de la actividad académica ha de tener una importancia en las evaluaciones no inferior a la que se atribuye a los procesos de comunicación interna mediante publicaciones. Ocurre también con ejemplos generados fuera del Norte, como es la declaración del Consejo de Decanos de las Facultades de Ciencias Sociales y Humanidades de Argentina⁷, que destaca la prioridad en el establecimiento de medidas del impacto social que superen el modelo imperante, inspirado en el impacto interno de las ciencias duras; algunos ejemplos que sugiere son la implicación en debates públicos a partir del conocimiento experto, la realización de consultorías en el desarrollo de políticas públicas, la resolución de problemas locales, o las publicaciones que incluyen orientaciones prácticas para la solución de problemas reales en contextos reales. Otro ejemplo destacado es la Declaración Mundial sobre la Educación Superior de la UNESCO ([UNESCO, 1998](#)), en la que se insiste en el carácter social y comprometido de la actividad académica (docente,

⁵ <http://www.ascb.org/dora/>.

⁶ <http://talloiresnetwork.tufts.edu/> En consulta del 8 de abril de 2016, se cuenta con 363 universidades miembro, procedentes de 77 países.

⁷ <http://www.fhuc.unl.edu.ar/decanosociales/2012/f.pdf>.

científica y de gestión externa e interna) con los problemas más relevantes de la sociedad planetaria, mediante el posicionamiento y la intervención protagonista de valores éticos. Por último y en la misma línea, en un estudio reciente e independiente encargado por la *Higher Education Funding Council for England* (Wilsdon, 2015), las principales recomendaciones vuelven a ser la necesidad de dar forma a indicadores para el impacto social y político, la transversalidad de la transparencia en todo el proceso—incluyendo los indicadores y su gestión—y la consideración de la diversidad.

Dos declaraciones recientes

En enero de 2014, tras varios años de reflexiones y de acciones colectivas, en especial del profesorado francés y belga, se publica la Carta de la Desexcelencia, que es divulgada desde la página *Salvemos la Universidad*⁸. Se trata de un documento muy crítico que relata las consecuencias de los discursos —y las consiguientes normas— sobre la excelencia que afectan a la universidad. La Carta invita a todos los individuos que forman parte de la institución a sumarse al movimiento de la desexcelencia, entendiendo éste como una defensa del amor por la ciencia y por la docencia, articulado a través de la autonomía y de la promoción de las condiciones idóneas para el trabajo disfrutado y bien hecho. Es un listado amplio de concreciones a tener en cuenta en docencia, investigación y gestión. La Carta reivindica, por ejemplo, reforzar las investigaciones de largo recorrido, estimular evaluaciones no cuantitativas, publicar en cada idioma y sobre los asuntos de interés local, fomentar estructuras democráticas o establecer discusiones abiertas sobre los resultados de investigación, entre otros puntos.

En septiembre del mismo año, en la ciudad de Leiden (Países Bajos), en el seno de la 19.^a Conferencia sobre Indicadores de Ciencia y Tecnología, se fragua una declaración que sería justificada y publicada por Hicks, Wouters, Waltman, Rijcke y Rafols (2015) en la revista *Nature*. La declaración es conocida como el Manifiesto de Leiden⁹, el cual tiene un valor especial, pues proviene de las personas que mejor conocen el mundo de la cienciometría. Su postura es opuesta a los hábitos evaluadores y globalizados del momento. En los diez principios del manifiesto se afirma en síntesis que la evaluación debe ser: 1) cualitativa con apoyo cuantitativo y, por tanto, nunca exclusivamente cuantitativa ni potenciando una precisión irreal; 2) no aplicada mediante un único patrón o estándar, sino estableciendo juicios adaptados a los objetivos que persiguen las investigaciones particulares o de grupos e instituciones, así como a las preocupaciones locales y los diferentes hábitos disciplinares, y 3) implementada y mejorada mediante métodos y resultados transparentes y sujetos a discusión continua con los grupos evaluados, a la vez que preocupada por estudiar los efectos sistémicos que se provoca.

Una visión de conjunto

Frente a la acumulación de críticas sobre el modelo globalizado actual de evaluación de la actividad científica en la ES, las reacciones se han instalado en dos compartimentos no conectados. Por un lado, se encuentran las sugerencias de mejora que, sin abandonar el modelo vigente, plantean solucionar algunos inconvenientes concretos. Por otro lado, están las propuestas de un cambio general que parta de otros principios rectores. En esta segunda línea, las sugerencias convergen en cinco puntos:

⁸ <http://www.sauvonsluniversite.com/spip.php?article6751>.

⁹ Además de la cita en *Nature*, puede encontrarse una traducción al castellano en <http://www.ingenio.upv.es/es/manifiesto#.V1KbDL7NJxI>.

1. Utilizar de preferencia métodos cualitativos.
2. Generar sistemas muy sensibles a la amplia diversidad de situaciones y contextos de investigación.
3. Considerar el amplio abanico de actividades investigadoras, buena parte de las cuales no se traduce en publicaciones.
4. Incluir como criterio las repercusiones que la actividad científica genera fuera del ámbito expreso de la ciencia, como ocurre con el impacto social.
5. Practicar una transparencia absoluta, que incluye la participación de los agentes implicados; la transparencia y la participación, a su vez, garantizan que el método no se anquilose.

Resulta lamentable la obsesión por generar una evaluación precisa (aunque no sea válida) en cierto grado incompatible con los criterios que parecen tener más peso en la impronta científica, como es el amor al conocimiento, elemento que destaca la Carta de la Desexcelexencia. ¿Cómo medirlo? Aunque no puede ser traducido a números, el amor al conocimiento es la mayor de las garantías para un trabajo científico bien hecho. No se puede medir, pero puede favorecerse o garantizarse de algún modo. En otras palabras, ningún cambio paradigmático puede proveer un grado tan elevado de precisión en la medida como observamos hoy. Esta pretensión es del todo ilusa. En términos generales, el cambio debe ir orientado a garantizar unas buenas condiciones contextuales para que el trabajo de investigación cumpla con la misión social que se le supone, es decir, que sirva para ayudar en el progreso social, sin atender a plazos ni a caminos prefijados. La ciencia o es libre o no es ciencia.

Una propuesta para principios rectores

Con base en las críticas mencionadas, en las declaraciones descritas y en las fuentes académicas que se mencionarán en cada caso, este apartado propone un esquema desde el cual realizar un esfuerzo de concreción, en la forma de principios rectores para un estilo de evaluación de la actividad científica que redunde en beneficio de la ciencia, de sus miembros y de la sociedad en su conjunto. El esquema se organiza a través de los siguientes ocho epígrafes. No obstante, es necesario prevenir antes sobre la comparabilidad de las prácticas actuales y el cambio paradigmático.

Si el objetivo es realizar una buena evaluación, hay que sopesar si se cuenta con tiempo y recursos suficientes para ello. Es preferible no evaluar a realizarlo mal, pues las consecuencias erróneas pueden resultar perjudiciales para individuos e instituciones. En este sentido, el sistema globalizado actual es, como ya se ha remarcado, tan viable como inadecuado. Su relativa sencillez parte de focalizar la medida en aquello que se deja medir, aunque no forme parte del núcleo de lo que se puede considerar un buen trabajo científico inserto en la ES. Apuntar a ese núcleo requiere más tiempo, personal especializado y métodos de investigación que no van a suministrar números sino información más rica y compleja. En definitiva, un mejor sistema es más complejo. No es posible sustituir lo que tenemos hoy por algo que, a la vez, sea mejor e igualmente sencillo de aplicar y simple de manejar. No obstante, hablamos de la institución universitaria, donde el nivel supuesto en el manejo de conocimientos debería menospreciar este inconveniente.

Descentralizar la evaluación

Tal y como denuncian Arrabal y Fernández (2012), el hábito establece encargar a empresas externas las tareas de evaluación, bajo los criterios de independencia y juicio de experto. No se

espera de quienes juzgan que sepan algo sobre lo que juzgan, sino que conozcan cómo se llevan a cabo las evaluaciones en los procesos de convergencia. Una entidad de tamaño limitado, que recibe un pago limitado para realizar una tarea en un tiempo limitado, no puede enfrentarse a un reto que parecería ilimitado, entrando en el contenido expreso de cada objeto de evaluación, sino que aplicará los métodos de conteo al uso que necesariamente homogeneizan lo que evalúan.

Al hilo de las duras críticas vertidas sobre la homogeneización y la estandarización, una solución que equipara la envergadura de la tarea con la del equipo que la realiza radica en descentralizar o abrir la evaluación, de tal forma que el número de personas dedicadas a ello sea en apariencia ilimitado. Una amplia participación garantiza que puedan aplicarse juicios específicos sobre los asuntos sustantivos de las investigaciones y demás objetos que se evalúan. En esencia, este asunto ha sido abordado por el modelo de la Responsabilidad Social Universitaria, que insiste en la transparencia de los procesos y en la participación efectiva de los grupos de interés (Vallaey, 2010, 2011), como garantías del compromiso social académico en la solución de problemas reales (De la Calle y Giménez, 2011; Gasca-Pliego y Olvera-García, 2011; Naval y Ruiz-Corbella, 2012; Pelekais y Aguirre, 2010). Iborra (2014) amplía el concepto de *grupo de interés* hasta el de *grupos afectados*, lo que incluye todo sector poblacional que pueda recibir las consecuencias de la actividad.

Descentralizar la evaluación implica que todo individuo u organización es un agente evaluador en potencia.

El desarrollo compartido de *software* constituye un buen modelo para inspirar cómo puede compartirse un sistema descentralizado de evaluación. Entre un número cada vez más elevado de iniciativas, el entorno de programación estadística R constituye un buen ejemplo¹⁰. Está desarrollado por una comunidad internacional de individuos y grupos que diseñan soluciones concretas operativas para todo tipo de situaciones donde se requiere el auxilio de técnicas estadísticas desde 1993. En la actualidad no sólo las revistas de todo tipo publican trabajos sobre R y, más aún, investigaciones analizadas mediante R, sino que esta iniciativa cuenta también con su propia revista¹¹. Los equipos trabajan sin ánimo de lucro; el código y la documentación de apoyo se disponen en modo de acceso abierto; cualquier individuo o grupo puede inspeccionar el código o la documentación y mejorar el resultado. Una de las consecuencias de esta actividad es la disponibilidad de más de 10,000 paquetes específicos de análisis, con acceso abierto y documentación concreta. Es difícil imaginar que esta intensa actividad de producción, transparente, efectiva y gratuita pueda hacerse realidad de otro modo que no sea mediante el desarrollo compartido y generado desde el ánimo honesto de contribuir.

Un sistema abierto de evaluación no sólo garantiza la participación de un número enorme de individuos, grupos e instituciones, sino también la adaptación del sistema a cada región, disciplina o campo de estudio; avala la discusión abierta y la consiguiente mejora de los procedimientos, y certifica las cuotas más elevadas de transparencia.

A la solidez de la experiencia acumulada en torno al desarrollo compartido cabe añadir una garantía fundamental. No es lo mismo compartir la creación y la mejora de un programa de ordenador, que poner en marcha evaluaciones que afectan la vida de individuos y de organizaciones. Para aumentar las garantías es necesario distinguir entre cuatro categorías de participación.

Cualquier agente —individuo o grupo— (categoría 1) puede realizar aportes al sistema, pero sólo los agentes *acreditados para la participación* (categoría 2) generan aportes que tienen con-

¹⁰ <https://www.r-project.org/>.

¹¹ <https://journal.r-project.org/>.

secuencias o efectos en los resultados de evaluación. Por su parte, los agentes *acreditados para el diseño* (categoría 3) afectan el modo como se comporta el sistema, se ponderan los aportes, se consiguen las acreditaciones y, en general, influyen en las características de funcionamiento de la herramienta. Por último, cabe un estilo de participación en esencia científico. Todo movimiento (aporte, decisión, cambio de diseño) que se ejerce sobre el sistema queda registrado en las bases de datos, y éstas son de acceso abierto. Tal circunstancia permite que un número elevado de equipos de investigación (categoría 4) basen su actividad precisamente en el estudio de tales bases de datos, aportando a la comunidad científica y evaluadora los resultados de sus análisis, mismos que poseen un valor notable para continuar en la mejora del sistema en su conjunto o en aspectos específicos que repercuten sobre determinadas disciplinas, trayectorias, regiones, etc.

Un modelo de estas características requiere pasar un periodo dilatado de prueba hasta alcanzar la madurez que permita constituirse como un sistema real de evaluación. En cierta medida, el resto de los principios rectores que configuran esta propuesta pueden ser considerados puntos de arranque del sistema. Es difícil imaginar que alguna otra alternativa consiga mejorar el número de personas preparadas que participan, la transparencia, la capacidad de adaptación a toda situación o el dinamismo en la auto-mejora.

Evaluar la evaluación

Este punto implica someter a evaluación los criterios, los modelos, los procedimientos, los agentes que participan, las estrategias de trabajo, las garantías, los resultados y cuantos elementos participen en la actividad evaluadora. Se trata de un requisito indispensable (Gaytán, 2012) que debería formar parte de la cultura organizacional (Alderman y Brown, 2005). La evaluación debería ser al menos tan exigente consigo misma como lo es respecto a lo que evalúa.

Un sistema descentralizado, como el descrito en el principio anterior, garantiza la meta-evaluación a través de tres fuentes. Por un lado, los agentes acreditados para el diseño establecen las rutinas de meta-evaluación que, como todo el sistema, son transparentes; del mismo modo, las discusiones y decisiones de estos agentes constan también en las bases de datos de acceso abierto. Por otro lado, se encuentra el mencionado recurso del trabajo libre y externo al sistema, realizado por parte de los equipos de investigación que acceden a las bases y publican sus resultados de análisis. Además de ambas fuentes, cabe esperar que las IES establezcan prácticas específicas de meta-evaluación, estimulando investigaciones de las bases del sistema y generando datos específicos relativos a su propia organización, por ejemplo, mediante entrevistas focales o encuestas a su personal y a los grupos de interés o afectados.

Desprotagonizar la evaluación a favor de las garantías

El mundo empresarial tiene la relativa facilidad de medir cuantitativamente los efectos a corto plazo de su gestión, en términos de monto de beneficios, por ejemplo. No obstante, esta traslación al mundo universitario —considerando a esta entidad como un agente productor que debe rendir cuentas en términos de resultados cuantificables— se vuelve antinatural como para generar múltiples problemas de implementación (García, 2008; Jarab, 2008; Montalbá, 2015; Stensaker, Frolich, Gornitzka y Maasen, 2008).

No es ya nada original afirmar que los resultados que provee la ciencia son difusos, complejos y a largo plazo, además de difícilmente atribuibles a individuos aislados (Manzo, 2014). Lo mismo ocurre con la educación (Egido, 2005), el otro gran componente de la actividad universitaria. En lugar de centrar la preocupación en evaluar resultados, conviene mucho más establecer garantías

y velar por su presencia efectiva. Luego, el objetivo principal de un nuevo modelo de evaluación es centrar el foco en el grado en que las garantías están en pleno vigor y no en someter a juicio a individuos e instituciones a partir de su labor de producción.

Al aceptar que la actividad científica, en esencia, provee resultados que no tienen por qué ser inmediatos, ni fácilmente acotables, ni claramente atribuibles, es obvio que el fruto de las garantías puede medirse en parte y permanecerá sujeto a discusión, por lo que un sistema coherente de evaluación no puede centrarse en esos frutos. Puede darse el caso que una institución garante facilite que, con el paso del tiempo, sea el personal de otra institución quien coseche los resultados en forma de publicaciones. Asignar a este segundo grupo el mérito exclusivo sería una incorrección ética y un error de estrategia. No puede cosecharse si antes no se ha sembrado. Dar crédito en exclusiva al monto de la cosecha y a quien la realiza potencia el oportunismo científico y pone en peligro la tarea de sembrar, propia del amor al conocimiento que debería seguir siendo constitutivo del hacer científico.

No obstante, *poner el foco* no implica *abandonar* por completo el interés por valorar en qué medida los agentes implicados generan resultados a corto o a medio plazo. Estos resultados son indicadores imperfectos que apuntan al constructo *garantía*. Su imperfección no les anula como *parte* del conjunto de indicadores a considerar para establecer juicios sobre el modo como las garantías funcionan en la práctica.

No obstante, ¿de qué garantías hablamos? Por un lado, la propia ciencia se construye sobre un edificio de garantías científicas, que [Frumento \(1998\)](#) enumera en términos de la búsqueda de objetividad, la estructura conceptual, el método de conocimiento, la precisión del lenguaje, el manejo probabilístico de la verdad, la refutabilidad de las afirmaciones, la evolución de los campos de conocimiento, el interés por la coherencia interna y la cantidad de logros que ha provisto. A estas garantías propias del método científico habría que añadir otras relativas a la gestión institucional.

[Manzano-Arrondo \(2016\)](#) sugiere seis para garantizar que:

- El personal que accede a la institución lo haga con vocación y conocimientos previos.
- El clima institucional sea el adecuado, con medios suficientes disponibles, condiciones laborales dignas, participación en los órganos de decisión, etc.
- Exista una clara relación dialógica entre la institución y el contexto global en el que se inserta (abordado con más detenimiento en un epígrafe específico sobre la globalidad).
- Se observe la transparencia de las metas organizacionales.
- El sistema tenga interés por medir cómo la actividad investigadora afecta a los individuos, a la propia institución, al conocimiento científico y al entorno político y social.
- Se pongan en marcha prácticas de evaluación de todo lo mencionado.

Considerar las garantías en cierta medida implica dar la vuelta al enfoque. Ya no se trata tanto de juzgar a los individuos sobre el grado en que están produciendo unidades, sino de interrogarles como jueces que reportan sobre el modo y el grado en que la institución en la que trabajan y el cuadro normativo de las legislaciones públicas que les afectan cumplen con las garantías que permiten y facilitan un buen trabajo científico. Desde esta perspectiva, la responsabilidad es repartida entre todos los agentes implicados.

El procedimiento idóneo para esta investigación de garantías no es la encuesta, sino la combinación de entrevistas focales con investigación documental, cuya justificación se corresponde con un principio rector abordado más adelante. La universidad hoy está sujeta a una proliferación de encuestas de satisfacción que busca evaluar los servicios institucionales y permitir altas

puntuaciones en los estándares de excelencia al uso (Alderman y Brown, 2005; Jáuregui, 2004). Sin embargo, la consecuencia más criticada es la inutilidad práctica para distinguir entre las instituciones, la superficialidad de las medidas, así como el cansancio y la saturación de agendas que provocan (Manzano-Arrondo, 2011; Ordine, 2013). Por el contrario, resulta preferible realizar una selección previa de informantes clave —considerando disciplinas, situaciones laborales, perfiles de investigación, etc.— y llevar a cabo entrevistas en profundidad que permiten obtener riqueza de información en torno a qué ocurre con las garantías. Del mismo modo, la investigación documental permite observar en las normativas, en las estructuras de funcionamiento y en las memorias de las instituciones indicadores para valorar el mismo asunto. Los sujetos de evaluación no se ciñen sólo a las IES sino, por encima de ellas, a las administraciones públicas, agencias de acreditación y legislaciones regionales o estatales; y, por debajo, a centros, departamentos y equipos de investigación. En definitiva, el interés radica en aplicar la lógica y los procedimientos de la investigación científica —en este caso de carácter cualitativo— para identificar los puntos fuertes y débiles en el estado de las garantías, elaborando un informe periódico final que incluye a las partículas descriptivas y que concreta un fuerte interés propositivo.

Ampliar el concepto de productividad científica

Considerar la producción de la actividad científica en las IES no sólo debe ser desprotagonizada, en los términos estimados en el apartado anterior, sino también notablemente ampliada. La intensa focalización en la publicación de trabajos en revistas científicas favorece concebir a la ciencia como un sistema de producción de «información registrada en formatos permanentes» (Spinak, 1998: 141). En epígrafes previos se ha destacado la necesidad de considerar otras unidades de evaluación, también propias de la actividad científica. La identificación de estas unidades depende de cada disciplina y campo de aplicación, por lo que su categorización es viable pero no simple. El sistema descentralizado descrito en el primer principio enunciado, aunque parta de un esquema como el que se ejemplifica en este epígrafe, debe tener como uno de los cometidos iniciales articular mediante el debate abierto un esquema mucho mejor adaptado a la amplia diversidad de las actividades reales.

El apartado más difícil de concretar, que ha recibido ya amplios esfuerzos pero que se encuentra lejos de generar unanimidad, es la medida de los impactos políticos y sociales. Esta dificultad no merma su pertinencia. Muy al contrario, constituye la preocupación más creciente y relevante. Se han propuesto criterios concretos para establecer sistemas de categorías de valoración, como son las centradas en la co-participación de la sociedad civil (Sívori, 2012), el compromiso social (Manzano-Arrondo, 2012), o los derechos humanos (Boni y Gasper, 2011), entre otros. Lo más prometedor en este sentido es la amplia experiencia de modelos de investigación que consideran el impacto social como el centro de interés. En esta línea se encuentran la investigación-acción participativa (Brydon-Miller, Greenwood y Maguire, 2003), así como las propuestas de investigación y de acción colectiva (Botero, 2012), aprendizaje-servicio (Speck, 2001), unidades de acción comprometida (Manzano-Arrondo y Suárez, 2015), investigación militante (Conti, 2004), metodología comunicativa crítica (Gómez, Latorre, Sánchez y Flecha, 2006) o coproducción investigativa (Bialakowsky, Costa, Patrouilleau, Martínez y López, 2007). Estos modelos incluyen, en su definición, las garantías de impacto que pueden servir de inspiración para la construcción de indicadores específicos.

En definitiva, superar la actual obsesión por la productividad científica centrada en el conteo de artículos implica redefinir lo que puede ser considerado un *aporte* científico. En términos generales, hay que asumir que la ciencia tiene o puede tener utilidades internas y externas. En

cada caso, el aporte puede tomar múltiples formas sin perder su naturaleza de unidad generada a partir de la investigación y el saber científico. Lo que sigue es un boceto que puede servir de inspiración para la resolución de este principio.

Actividad científica:

- I. De utilidad interna
 - A. Para un campo científico
 - i. En formato permanente
 - a. Publicaciones clásicas (artículo, libro, informe. . .).
 - b. Nuevos formatos (vídeo, página web, blog. . .).
 - c. Útiles (programas de ordenador, bancos de recursos, protocolos y guías. . .).
 - ii. En formato transitorio (conferencia, taller, curso, mesa redonda. . .).
 - B. Para la propia institución
 - i. Aportes para la mejora del funcionamiento interno.
 - ii. Labores de gestión de la investigación (dirección de revistas, gestión de ediciones científicas, organización de eventos, revisión de artículos y proyectos, etc.).
- II. De utilidad externa
 - A. Para la población general
 - i. Divulgación de la ciencia.
 - ii. Labores de interpretación sobre acontecimientos de actualidad y crítica fundamentada.
 - B. Para poblaciones específicas
 - i. Resolución de problemas concretos.
 - ii. Elaboración de guías en formatos diversos para la resolución práctica de problemas o para estilos de afrontamiento.
 - iii. Capacitación de agentes para la solución de problemas locales.
 - iv. Formación en metodología de la investigación e intervención, adaptada a situaciones concretas.
 - C. Para las administraciones públicas
 - i. Transferencia de conocimiento en formato de informes, directrices, vías de actuación o solución, etcétera
 - ii. Labores de asesoría
 - D. Para organizaciones sin ánimo de lucro y empresas
 - i. Transferencia de conocimiento.
 - ii. Colaboraciones de investigación, acción y desarrollo.
 - iii. Labores de asesoría.

Escoger unidades y perfiles pertinentes

La estandarización actual no sólo se aplica sobre el procedimiento, sino sobre el individuo-tipo que moldea. Es obvio que ninguna institución científica, y menos académica, puede mantenerse en exclusiva con personas cuya especialidad sea la publicación de artículos en revistas científicas. La ciencia implica la complicitad de especialidades diversas, en términos de qué se investiga y qué función se cumple dentro de cada cosa que se investiga. Con tal punto de partida, [Manzano-Arrondo \(2015a\)](#) describe algunos ejemplos de perfiles que concurren en la efectividad de las instituciones académicas. Así, además de publicar en revistas científicas, otras funciones fundamentales son, por ejemplo: 1) la divulgación, que implica *traducir* la jerga disciplinar a un lenguaje accesible para la población general e incluso para colegas y estudiantes sin una formación

específica; 2) el estímulo a la reflexión, que permite poner en duda los supuestos conocimientos indiscutibles y que se expresa con frecuencia fuera del estándar de los artículos, como ocurre con la actividad docente, las conferencias, los blogs y otros formatos; 3) el trabajo en las fronteras disciplinares, mediante el que algunos individuos hacen de puente entre diferentes tradiciones, líneas de investigación, intereses de conocimiento, etc., con una baja productividad en términos de publicaciones o adquiriendo formatos lentos y no bien reconocidos, como ocurre con los libros; 4) la creatividad casi aleatoria, que suministra resultados inesperados, generados mucho más por la curiosidad honesta que por el interés de rentabilizar los esfuerzos, por lo regular en manos de personas que priorizan la pasión científica; 5) la conexión con los entornos locales, con frecuencia en manos de personas más cercanas a las *trincheras* de barrios y movimientos que al corsé de la comunicación estándar de la ciencia, o 6) la organización de colectivos y de eventos a través de los cuales se estimula el intercambio de elementos de conocimiento científico y se crean cohesión, identidad y fuerza colectiva. No todo individuo tiene por qué cumplir con todas las funciones. Lo que se observa hoy es que un ejército amplio de miembros de la academia siembra las condiciones para que un grupo más reducido termine rentabilizando al final el esfuerzo colectivo en publicaciones que, a la postre, constituyen la actividad protagonista, tal y como ha sido descrito más atrás. Un nuevo paradigma de evaluación debe equilibrar la situación, reconociendo la importancia que los diferentes perfiles juegan en el resultado final. Es más, la cuestión no pasa por contemplar perfiles diversos, sino exigirlos. Una institución monocroma no es deseable. La diversidad de perfiles constituye una excelente garantía para esperar resultados innovadores y robustos.

En esta misma línea se concluye que, si la ciencia es una aventura que concierne a individuos de perfiles muy variados, entonces la focalización en la unidad-individuo quizá sea errónea o, al menos, muy incompleta. El individuo perfecto no existe, vive en la misma ficción que los superhéroes. Nadie es capaz, de forma autosuficiente o individual, de realizar una lectura exhaustiva de su campo de especialización; de establecer juicios pausados y reflexiones sobre el estado de la cuestión que investiga; de conocer las convocatorias, leerlas, redactar los proyectos y rellenar los formularios requeridos; de realizar el trabajo de campo en la calle o en el laboratorio, cara a cara con las personas entrevistadas, los animales observados, o los materiales de taller; de dominar todos los entresijos metodológicos, tanto de diseño como de técnicas de análisis de datos, que requiere cada investigación que aborda; de diseñar, redactar e incluso traducir para conseguir artículos de éxito; de dar forma a otros informes o memorias, para administraciones y otros agentes; de mantener al día su propio *curriculum*; de realizar labores de mentoría de personal investigador en formación; de dirigir por completo las tesis doctorales asociadas, las becas de investigación, el estudiantado en prácticas; de interactuar con su propia institución, ocupando cargos académicos, liderando normativas e iniciativas de mejora; etc. Aunque toda regla tiene su excepción, y quizá pudiera encontrarse algún caso aislado, no se puede organizar un sistema de evaluación sobre un individuo inexistente.

Además de la concurrencia de perfiles diversos de individuos que cumplen funciones necesarias en el progreso del conocimiento científico, concurre otra razón de peso para variar la unidad de análisis o de evaluación. Como destacan Muñoz, García y López (2014) en su estudio sobre los efectos de las investigaciones sobre la calidad de las IES y de la actividad científica, «el trabajo intelectual no es individual, es colectivo» (p. 194). Rara vez las personas trabajan de forma aislada. Firman quienes firman las contribuciones, el saber científico es una tarea de grupo. Las aportaciones que se realizan en ciencia se nutren de un amplio bagaje heredado, donde es más fácil observar líneas colectivas de investigación y de construcción de conocimiento que trayectorias individuales puras. De hecho, la impronta colectivista y cooperativista del conocimiento constituye un componente capital del imaginario científico, al menos desde el siglo XVIII (Manzo, 2014).

En tales circunstancias, resulta más coherente valorar el trabajo de grupos o de equipos que de personas aisladas. ¿Cómo realizar esta tarea? La concreción última es una tarea compleja, propia de un sistema como el descrito en el primer principio. No obstante, algunas orientaciones son: 1) considerar la productividad de los equipos en los que se participa cuando se evalúan las trayectorias individuales y no sólo la de los individuos tomados de manera aislada; 2) someter a valoración a los propios equipos, no sólo en términos de productividad, sino de garantías, considerando por ejemplo criterios de diversidad interna de sus miembros en torno a los perfiles mencionados; 3) valorar los equipos en los términos con que [Paloma y Manzano-Arrondo \(2011\)](#) definen las organizaciones que generan beneficios sociales fuera de sí, al tiempo que hacen otro tanto dentro de sí, y 4) valorar el funcionamiento y los resultados de los equipos en función de los medios de que disponen para realizar su labor.

En definitiva, las valoraciones de las trayectorias individuales y de los equipos de investigación deben ponderarse en resultados comunes, de tal forma que un individuo no recibe una valoración final aislada, sino que depende de los grupos en los que ha trabajado, del mismo modo que los grupos beben de los aportes de sus miembros. De este modo se relativizan los protagonismos individuales y queda reconocida la colectividad del trabajo y del espíritu científico.

Introducir el enfoque cualitativo

El protagonismo de los enfoques cuantitativos de investigación —que llegan a confundir ciencia con ciencia cuantitativa ([Ruiz e Ispizúa, 1989](#))— dificulta la visibilidad de criterios de calidad que se encuentran ya muy asentados en la metodología cualitativa científica (véase, por ejemplo, [Kidd, 2002](#); [Márquez, 2007](#); [Martínez, 2006](#); [Pérez, 1994](#); [Quintana, 2006](#)). Es más, la aplicación de los estándares al uso para la evaluación dificulta la consideración y la valoración positiva de las investigaciones cualitativas ([Noreña, Alcaraz-Moreno, Rojas y Rebollo-Malpica, 2012](#)). A pesar de ello, este enfoque puede enriquecer los indicadores de calidad y solucionar algunos de los problemas de concreción del impacto social, en aspectos que afectan valores, compromiso o implicación ([Sandín, 2000](#)), así como la consideración de procesos, datos y resultados complejos ([Cornejo y Salas, 2011](#); [Jorge, 2003](#)). Por ello, iniciativas ya descritas, como el Manifiesto de Leiden, generado desde el corazón de la cienciometría, aconsejan utilizar enfoques cualitativos, complementados pero no comandados por técnicas cuantitativas.

La consideración de la epistemología cualitativa en la evaluación de la actividad científica puede tener varias implicaciones prácticas. Con el ánimo de ejemplificar indicaciones menos generalistas:

- I. Para la valoración de un aporte concreto:
 - A. Auto-reflexividad y transparencia: consideración positiva de las *confesiones* de los equipos de investigación en torno a los valores que subyacen a sus decisiones.
 - B. Referencias de utilidad: indicaciones concretas para la aplicabilidad de los resultados de investigación en problemáticas específicas sociales, ambientales, tecnológicas, económicas, etcétera.
 - C. Visión sistémica: referencias al marco lo más amplio posible en el que se inserta la investigación y a su papel en ello.
 - D. Participación y diversidad: valoración positiva de la diversidad interna del equipo de investigación y de la participación de otros agentes (población afectada, organización de interés, administración pública. . .).

II. Para la tarea propia de evaluación:

- A. Triangulación: utilización de varios métodos paralelos para someter a juicio los propios resultados de evaluación.
- B. Categorizaciones: recurso de un sistema de categorías mutuamente no excluyentes en cuyas celdas puede ubicarse la investigación o las investigaciones que se valoran. El sistema puede considerar, por ejemplo, categorías necesarias, recomendables y opcionales.

Desde la impronta del enfoque cualitativo, aunque no confinado exclusivamente en él, se inserta un modelo que presenta excelentes similitudes con las características del cambio paradigmático de evaluación: la Investigación-Acción Participativa (IAP). Como puede sintetizarse a partir de la literatura sobre el tema (véase, por ejemplo, los trabajos de [Balcazar, 2003](#); [Botero, 2012](#); [Brydon-Miller, 1997](#); [Kalliola, 2009](#); [Leal, 2009](#), o [Thiollent, 2011](#)), la IAP busca: 1) investigar, saber qué ocurre; 2) actuar, solucionar o mejorar lo que ocurre, y 3) garantizar la relevancia de lo que se hace mediante la participación de la población afectada. Las tres características se encuentran en plena sintonía con una motivación evaluadora —saber qué ocurre— cuyo objetivo último es incrementar la calidad de la actividad científica —mejorar lo que ocurre— mediante la condición indispensable de contar con una cultura organizacional afín —complicidad de los miembros de la institución— que ya señalaron [Winn y Green \(1998\)](#).

La marcada unidireccionalidad de los actuales sistemas globalizados de evaluación facilita reacciones contraproducentes, caso de que el objetivo fuera la mejora de la calidad científica. Ya se ha descrito más atrás que las respuestas por parte *del otro lado* de la evaluación unidireccional, es decir, individuos, grupos e instituciones bajo juicio, se inscriben más bien en la resistencia, la marginalidad o la frecuente adaptación acrítica y sin convicción. Por el contrario, un sistema inspirado en los principios de la IAP favorece el compromiso y la co-responsabilidad. Individuos, grupos e instituciones son las entidades con mayor motivación para que las evaluaciones resulten pertinentes respecto al espíritu científico y a la amplia diversidad de concreciones que se derivan de él en la práctica cotidiana.

Utilizar criterios globales

Aunque en su origen la glocalización se refiere al modo en que un contexto local se adapta a una dinámica global ([Giulianoti, 2007](#); [Saainen, 2009](#)) o se resiste a una presión global ([Acosta, 2003](#)), el significado que considero para el concepto es el de combinación dialógica entre lo local y lo global ([Manzano-Arrondo, 2012](#)).

En buena medida, la cualidad glocal de los aportes científicos pueden garantizarse mediante el seguimiento de otros principios de esta propuesta, como la participación de sectores afectados o la consideración de criterios cualitativos. A su vez, cabe esperar que la propia naturaleza del saber científico garantice la motivación intrínseca por la conexión global. No obstante, estos precedentes son insuficientes para justificar el valor glocal de las aportaciones científicas. En este sentido, algunas concreciones podrían ser:

1. *Garantías de acceso.* Los datos, el método, los resultados y las publicaciones u otros aportes se encontrarán en modo de acceso abierto. En el caso que se generen publicaciones en un idioma diferente al local, serán traducidas. En el caso que las aportaciones requieran una alta preparación para ser comprendidas, se realizarán adaptaciones con carácter divulgativo.

2. *Diversidad de ámbito.* En la trayectoria de publicaciones, se combinarán revistas —u otros formatos— de ámbito global, con otras especializadas en el contexto local del individuo o equipo.
3. *Diversidad de origen.* Las publicaciones referenciadas, los marcos teóricos y las concreciones prácticas tendrán orígenes diversos. Se valorará positivamente superar la categoría «esta investigación es un caso de» (Kreimer, 2011) con que las publicaciones de la periferia refuerzan los modelos teóricos occidentales.
4. *Participación en redes.* El individuo o equipo formará parte activa de redes de ámbito local y de ámbito global, ambas sobre los temas de investigación que aborda.

El objetivo específico en la evaluación de la glocalidad es valorar en qué medida y modo los aportes generados por el individuo, el grupo o la institución cuentan con dos características fundamentales. La primera de ellas es que los resultados de investigación se encuentran al alcance de todo individuo u organización con interés potencial, siguiendo la consideración preferente de «grupo afectado» a que se refiere Iborra (2014). Ello implica a todo sector poblacional que pueda recibir el impacto positivo o negativo de las investigaciones, aunque no ocupe una posición de poder suficiente como para ejercer algún tipo de presión o visibilidad. La accesibilidad se refiere al acceso sin barreras monetarias y a garantizar que tampoco existen barreras lingüísticas o intelectuales, de tal forma que todo grupo local pueda comprender los hallazgos sin necesidad de contar con una cualificación específica. En este apartado se consideran también las administraciones públicas que pudieran afectar sus políticas mediante los resultados de la actividad científica; por su parte, la evaluación de la glocalidad valora la relevancia o la pertinencia locales de las temáticas que se abordan cuando la investigación no es fundamentalmente básica, de tal forma que existan garantías de que el estudio no se reduce a la replicación de intereses propios de los países o disciplinas centrales, tal y como denuncia el mencionado trabajo de Kreimer (2011).

Usar referentes y no ordenaciones

Los criticados *rankings* de universidades y de revistas según su TRC comparten la misma combinación de dos supuestos: *calidad es posición* y *progreso es competición*; el primer enunciado ha sido ya abordado en epígrafes previos, requiere simplificar la complejidad en una dimensión única cuantificable y realizar un ejercicio de fe en torno a la pertinencia de los resultados; el segundo enunciado supone que la sociedad en general o la ciencia en particular progresarán gracias a que sus miembros compiten entre sí por ocupar posiciones que son limitadas en número.

Este octavo principio establece la pertinencia de utilizar referentes de buena ejecución en lugar de ordenaciones. De este modo, por ejemplo, una buena revista lo será no porque ocupe una posición relativa en un listado, sino porque cumple con requisitos establecidos desde criterios científicos. Así se considera, por ejemplo, en las iniciativas mencionadas de Redalyc y Latindex. Las ordenaciones garantizan la insatisfacción de las unidades no contempladas en los listados, y de las unidades contempladas pero mal posicionadas. La utilización de referentes —establecidos con rigor y el nivel de exigencia que se considere— permite contar con un horizonte preciso y accesible por méritos propios. Todo equipo de investigación, todo miembro de la academia, toda revista, todo aporte concreto, pueden aspirar a cumplir los referentes que sean de aplicación en cada caso, sin que el éxito en este empeño dependa del fracaso comparativo de otros individuos, grupos o iniciativas.

No existen argumentos de validez científica que aconsejen acudir a ordenaciones. Imagine que una investigación se considerara válida no por criterios absolutos, sino por su posición en un

ranking unidimensional que escogiera algún criterio objetivo y cuantificable, como el tamaño de las muestras utilizadas, el número de citas que contiene, el monto de dinero que ha consumido o la cantidad de personas que firman el trabajo. Utilizar cualesquiera de estos criterios para generar ordenaciones de investigaciones, ofrecidas como *posiciones de calidad*, resulta tan aberrante y ridículo desde la impronta científica como lo que se hace hoy con los *rankings* al uso. Una reflexión similar puede aplicarse a revistas, a las IES, a equipos de investigación o a trayectorias individuales. Las ordenaciones constituyen una drástica reducción que debería resultar inadmisibles desde la perspectiva de la ciencia. Todo sistema de evaluación que desee resultar relevante a la actividad científica y la utilidad social pertinentes debe centrarse en referentes, nunca en ordenaciones.

Discusión y conclusiones

El enfoque de rendición de cuentas y la preocupación de las administraciones públicas por encontrar criterios objetivos que guíen una buena distribución de los recursos son argumentos de peso para asumir tareas de evaluación de la actividad científica. Sin embargo, en los apartados previos hemos observado un amplio listado de inconvenientes prácticos derivados de los sistemas de evaluación utilizados en el presente. La literatura sobre la impronta globalizada ha mostrado: 1) la debilidad de los supuestos en los que se asienta el uso de la TRC y de los *rankings*; 2) las consecuencias indeseables en términos de adaptación estética e incluso de fraude en torno a la producción científica; 3) el deterioro de los principios rectores que se le suponen a la ciencia; 4) un profundo desequilibrio de trato entre regiones, idiomas, disciplinas y enfoques, y 5) efectos psicológicos y laborales negativos derivados de la instrumentalización del personal universitario.

Estos inconvenientes son evitables. Existen propuestas que abordan mejoras o incluso un cambio de paradigma. Este trabajo se ha situado en la segunda opción, fundamentando un relevo sustancial del modo como se evalúa la actividad científica, desde dentro y fuera del ámbito académico. En su esencia, la propuesta pretende ubicarse en la ética del cuidado, librando al personal académico de una situación cuasi-patológica que no se justifica por el razonable interés en una buena gestión de los recursos públicos, ni mucho menos por la utilidad social de la ciencia. La propuesta es, además, afín a la mencionada Carta de la Desexcelencia, desde el criterio general de que no hay mayor garantía de un trabajo bien hecho que el que se ha generado desde la motivación intrínseca, concretada mediante la pasión por el aprendizaje y el descubrimiento, así como por el espíritu crítico y desinteresado de la construcción colectiva.

En esta línea, el documento ha propuesto una serie de ocho principios para inspirar un cambio paradigmático: descentralizar la actividad evaluadora estimulando una participación abierta y acreditada, capaz de adaptarse a las abundantes particularidades del trabajo investigador; realizar meta-evaluación con los mismos niveles de precisión y de exigencia que en la evaluación; centrar el interés en las garantías estructurales más que en la producción de resultados; ampliar la visión de esa producción, contemplando una imagen más real de la actividad científica; potenciar la diversidad de perfiles individuales, insertos en equipos de investigación, cuya unidad es más coherente para el modo de funcionamiento de la ciencia; incrementar el enfoque cualitativo en la evaluación, matizado pero no sustituido por mediciones cuantitativas; valorar positivamente la imbricación local de la ciencia, en una conversación equilibrada con su impronta universal; y usar referentes de buena ejecución en lugar de establecer ordenaciones como criterio para establecer juicios de validez. La amplitud y la ambición de la propuesta son inevitables en un trabajo que sugiere un cambio de paradigma, pero añaden una potente limitación práctica: no es viable llegar a un nivel de concreción que resulte suficiente para aplicar el modelo de inmediato en situaciones cotidianas. Esta limitación abre la puerta de necesarios desarrollos futuros, cuyo objetivo sea no

sólo someter al habitual debate el conjunto de la propuesta, sino profundizar en sus posibilidades de concreción hasta niveles operativos.

No obstante, bajo el supuesto de que dicha propuesta fuera inhabitual y viable, la bondad de un paradigma alternativo no es variable suficiente para augurar éxito en el empeño. El principal inconveniente es el notable grado de implantación del sistema actual, lo que atañe no sólo a las dinámicas internacionales y a las legislaciones estatales, sino de forma muy especial al adiestramiento y moldeamiento de las trayectorias individuales. En una situación tan homogeneizada y arraigada como ésta, todo cambio tiene un pronóstico difícil.

Además de los hábitos establecidos, hay que asumir que las normativas se generan y se aplican mediante estructuras administrativas dirigidas por personal que sabe hacer muy bien su trabajo. El cometido específico es aplicar la legislación vigente en la cotidianidad de la gestión institucional. La propuesta de cualquier modelo alternativo debe contar con un nivel máximo de concreción para contemplarse por la gerencia administrativa, puesto que cualquier laguna de implementación implicaría una importante indefensión a la hora de tomar decisiones cotidianas. En otras palabras: para un gestor académico sujeto a una realidad cotidiana con frecuencia estresante, prescindir de un modelo de evaluación es impensable si no cuenta con una alternativa definida al menos con el mismo nivel de detalle y que sea respaldada políticamente. En coherencia con ello y a partir de una investigación sobre perfiles académicos, [Manzano-Arrondo \(2015b\)](#) muestra que el profesorado universitario con responsabilidad de gestión tiende a reproducir y reforzar los modelos actuales, con independencia de que los considere o no apropiados o pertinentes. El principal argumento es que son modelos operativos y compartidos. Este hábito gerencial casa bien con las investigaciones que muestran la resistencia de los individuos para abandonar conductas que al menos aparentan ser compartidas e incluso tradicionales, para evitar las sensaciones negativas derivadas de posibles errores o funcionamientos imperfectos ([Frith, 2014](#)). La profundización en estos mecanismos ha propiciado que algunas iniciativas no hegemónicas hayan ido desapareciendo por falta de inversión pública ([Giménez-Toledo, 2015](#)).

Los hábitos establecidos, las normativas afines y las exigencias operativas de la gerencia constituyen, por tanto, barreras serias para el éxito de un cambio paradigmático de la evaluación. Pero la actualidad no parece estable y los numerosos inconvenientes señalados resultan cada vez más visibles y compartidos, a juzgar por la abundancia de la literatura que se ocupa de ello y de la proliferación de iniciativas internacionales en los últimos años. Luego, el cambio es necesario. La solución al conflicto entre necesidad y *statu quo* sólo puede consistir en ensayar un modelo como el propuesto, desarrollando todos los detalles de funcionamiento cotidiano hasta alcanzar la madurez que le permita ser utilizado como paradigma mejorado. Lo contrario implicaría seguir reforzando dinámicas que han mostrado de sobra su carácter contraproducente. La evaluación de la ciencia debería ser científica, no acomodaticia.

El primero de los principios inspiradores ha intentado describir de forma muy somera las características de un sistema de evaluación descentralizado, organizado a partir del resto de los principios, y estimulando un debate especializado que permita eclosionar en una versión inicial de ensayo. La amplia experiencia acumulada en campos participativos de desarrollo, de investigación y de mejora, como el software participativo o la IAP, permite prevenir errores de diseño y de ejecución, así como afinar en la definición de esa versión inicial.

Por último, es necesario tener presente que los miembros de la ciencia y de la academia merecen ser considerados fines en sí mismos y no instrumentos cuyo uso permita un hipotético beneficio social a partir del incremento de la calidad de las investigaciones científicas. En una sociedad sabia, todas sus instituciones deberían diseñarse y funcionar como criaturas promotoras de justicia y bienestar. Las IES no tienen por qué ser una excepción a esta norma. En coherencia,

sea cual fuere el paradigma de evaluación utilizado, debería resultar evidente que la actividad evaluadora se dirige hacia la comprobación de las garantías y de los resultados que promueven justicia y bienestar, no solo fuera de los muros de la academia, sino también dentro.

Referencias

- Acosta, Alberto. (2003). En la encrucijada de la glocalización. Algunas reflexiones desde el ámbito local, nacional y global. *Polis*, 1, 1–18.
- Aguado, Eduardo, Rogel, Rosario, Garduño, Gustavo y Zúñiga, María. (2008). Redalyc: una alternativa a las asimetrías en la distribución del conocimiento científico. *Ciencia, Docencia y Tecnología*, 19(37), 11–30.
- Aguado-López, Eduardo. (2013). ¿Son los promotores del acceso abierto los alejandrinos del siglo XXI? *Cuadernos del Pensamiento Crítico Latinoamericano*, 6, 6.
- Alatas, Syed Farid. (2003). Academy dependency and de global division of labour in the Social Sciences. *Current Sociology*, 51(6), 599–633. <http://dx.doi.org/10.1177/00113921030516003>
- Albarado, Alejandra. (2004). La ética del cuidado. *Revista Aquichan*, 4(4), 30–39.
- Alcalde, J. (2016). Naux demandoj pri lingva justeco en la fakliteraturo. En J. A. Vergara y A. André (Eds.), *Internacia Kongresa Universitato*. Rotterdam: UEA.
- Alcalde, Xavier. (2016). La justicia lingüística segons la filosofía política. *Kataluna Esperantisto - Llengua Internacional i Drets Lingüístics*, 365(5), 12–37.
- Alderman, Geoffrey y Brown, Roger. (2005). Can quality assurance survive the market? Accreditation and audit at the crossroads. *Higher Education Quarterly*, 59(4), 313–328. <http://dx.doi.org/10.1111/j.1468-2273.2005.00300.x>
- Amigot, Patricia y Martínez, Laureano. (2013). Gubernamentalidad neoliberal, subjetividad y transformación de la universidad. La evaluación del profesorado como técnica de normalización. *Athenea Digital*, 13(1), 99–120.
- Ansele, Manuel. (2017). Amenazas para el planeta de las que nadie se entera porque están escritas en chino. *El País*, 2 de enero.
- Araujo, Sonia M. (2014). La evaluación y la universidad en Argentina: políticas, enfoques y prácticas. *Revista de la Educación Superior*, 43(4), 57–77.
- Argullol, Rafael. (2014). La cultura enclaustrada. *El País*, 5 de abril.
- Arnett, Jeffrey J. (2008). The neglected 95%. Why American psychology needs to become less American. *American Psychologist*, 63(7), 602–614. <http://dx.doi.org/10.1037/0003-066X.63.7.602>
- Arrabal, Juan José y Fernández, Rodrigo. (2012). Calidad y evaluación: una perspectiva sindical. En A. Castro (Ed.), *Calidad, evaluación y encuestas de la docencia universitaria* (pp. 15–26). Murcia: Ediciones Laborum.
- Báez, René. (2012). Ecuador: homo sapiens y homo economicus confrontan en las universidades. En R. Báez (Ed.), *Universidades: entre la mercadofilia y la alterglobalización*. (pp. 11–18). Quito: Gallo Rojo. Centro de Pensamiento Político.
- Balcazar, Fabricio E. (2003). Investigación Acción Participativa (IAP): Aspectos conceptuales y dificultades de implementación. *Fundamentos en Humanidades*, 4, 59–77.
- Balestena, Eduardo. (2001). Ética del saber y de las instituciones. En N. Kisnerman (Ed.), *Ética, ¿un discurso o una práctica social?* (pp. 45–62). Buenos Aires: Paidós.
- Ball, Stephen J. (2003). Profesionalismo, gerencialismo y performatividad. *Revista Educación y Pedagogía*, 15(37), 87–104.
- Becerril-García, Arianna, Aguado-López, Eduardo, Rogel-Salazar, Rosario, Garduño-Oropeza, Gustavo y Zúñiga-Roca, María. (2012). De un análisis centrado en la revista a un modelo centrado en entidades: la publicación y producción científica en la nueva plataforma Redalyc.org. *Aula Abierta*, 40(2), 53–64.
- Beigel, Fernanda. (2014). Publishing from the periphery: Structural heterogeneity and segmented circuits. The evaluation of scientific publications for tenure in Argentina's CONICET. *Current Sociology*, 62(5), 743–765. <http://dx.doi.org/10.1177/0011392114533977>
- Benítez, M. (2012). Propuesta para una academia alternativa. En R. Báez (Ed.), *Universidades: entre la mercadofilia y la alterglobalización* (pp. 32–36). Quito: Gallo Rojo. Centro de Pensamiento Político.
- Bernasconi, Andrés. (2006). La difícil tarea de clasificar universidades. *Calidad en la Educación*, 25, 81–96.
- Bialakowsky, Alberto, Costa, María I., Patrouilleau, Mercedes, Martínez, Rocío S. y López, Ana I. (2007). Capitalismo y método. Alternativas de la coproducción investigativa. *Revista Laboratorio-on line. Estudios sobre Cambio Estructural y Desigualdad Social*, 8(19).
- Björk, Bo Christer y Solomon, David. (2014). *Developing and Effective Market for Open Access Article Processing Charges*. London: Wellcome Trust.

- Boff, Leonardo. (2001). *Ética planetaria desde el Gran Sur*. Madrid: Trotta.
- Boni, Alejandra y Gasper, Des. (2011). La universidad como debiera ser. Propuestas desde el desarrollo humano para repensar la calidad de la universidad. *Sistema*, 220-221, 93-109.
- Bormmann, Lutz y Daniel, Hans Dieter. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80. <http://dx.doi.org/10.1108/00220410810844150>
- Botero, Patricia. (2012). Investigación y acción colectiva [IAC]. Una experiencia de investigación militante. *Utopía y Praxis Latinoamericana*, 17(57), 31–47.
- Broekhuizen, Francien, Dawes, Simon, Mikelli, Danai y Wilde, Poppy. (2015). Just because you write about posthumanism doesn't mean you aren't a liberal humanist: An interview with Gary Hall. *MECCSA*, 9(1), 1–11.
- Brydon-Miller, Mary. (1997). Participatory action research: Psychology and social change. *Journal of Social Issues*, 53(4), 657–666. <http://dx.doi.org/10.1111/j.1540-4560.1997.tb02454.x>
- Brydon-Miller, Mary, Greenwood, Davydd y Maguire, Patricia. (2003). Why action research? *Action Research*, 1(1), 9–28. <http://dx.doi.org/10.1177/14767503030011002>
- Buela-Casal, Gualberto. (2003). Evaluación de la calidad de los artículos y de las revistas científicas: propuesta del factor de impacto ponderado y de un índice de calidad. *Psicothema*, 15(1), 23–35.
- Callahan, Michael, Wears, Robert L. y Weber, Ellen. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *Journal of American Medical Association*, 287(21), 2847–2850. <http://dx.doi.org/10.1001/jama.287.21.2847>
- Cerezo, Héctor. (2007). ¿Publicar o morir? *Elementos*, 66, 21–25.
- Cheng, Ming. (2011). The perceived impact of quality audit on the work of academics. *Higher Education Research & Development*, 30, 179–191. <http://dx.doi.org/10.1080/07294360.2010.509764>
- Chong-Shing, D. (2003). Between English and Esperanto: What does it take to be a world language? *International Journal of the Sociology of Language*, 164(33), 33–63. <http://dx.doi.org/10.1515/ijsl.2003.055>
- Christensen, Tom. (2011). University governance reforms: Potential problems of more autonomy? *Higher Education*, 62, 503–517. <http://dx.doi.org/10.1007/s10734-010-9401-z>
- Codd, John. (2005). Teachers as 'managed professionals' in the global education industry: The New Zealand experience. *Educational Review*, 57, 193–206. <http://dx.doi.org/10.1080/0013191042000308369>
- Collado, Susana y Vázquez, Mariana Asunción. (2006). Ética en las publicaciones científicas. *Biociencias*, 4.
- Conti, Antonio. (2004). La encuesta como método político. En Marta Malo (Ed.), *Nociones Comunes* (pp. 55–65). Madrid: Traficantes de Sueños.
- Cornejo, Marcela y Salas, Natalia. (2011). Rigor y calidad metodológicos: un reto a la investigación social cualitativa. *Perspectiva. Individuo y Sociedad*, 10(2), 12–34. <http://dx.doi.org/10.5027/psicoperspectivas-vol10-issue2-fulltext-144>
- Dávila, Lisha Pamela. (2014). Políticas públicas en ciencia y tecnología. Participación pública para la construcción de una nueva república democrática de la ciencia. *Revista Cuestiones de Población y Sociedad*, 6(6), 63–73.
- De la Calle, Carmen y Giménez, Pilar. (2011). Aproximación al concepto de Responsabilidad Social del Universitario. *Revista Comunicación y Hombre*, 7, 236–247.
- De la Corte, Luis. (2004). Apuntes sobre crítica y ciencia social: a vueltas con la cuestión del compromiso (I). *Realidad*, 99, 261–297.
- Deulofeu, Evarina Victoria. (2009). Acerca del conocimiento. Una aproximación a la epistemología. En R. Valera y E. V. Deulofeu (Eds.), *Introducción a los debates filosóficos actuales. Selección de lecturas*. (pp. 93–110). La Habana: Editorial Félix Varela.
- Egido, Inmaculada. (2005). Reflexiones en torno a la evaluación de la calidad educativa. *Tendencias Pedagógicas*, 10, 17–28.
- Ellacuría, Ignacio. (1999). *Escritos Universitarios*. San Salvador: UCA Editores.
- Enders, Jürg. (2015). Una 'carrera armamentística' en la academia: los rankings internacionales y la competencia global para crear universidades de clase mundial. *Revista de la Educación Superior*, 44(4), 83–109.
- Evers, Colin W. y Lakomski, Gabriele. (2013). Methodological individualism, educational administration, and leadership. *Journal of Educational Administration and History*, 45(2), 159–173. <http://dx.doi.org/10.1080/00220620.2013.768969>
- Fernández, Rodrigo, Martínez, Yerko y Velasco, Nicolás. (2006). Indicadores y estándares internacionales de calidad universitaria. *Calidad en la Educación*, 25, 17–29.
- Fernández-Quijada, David, Masip, Pere y Bergillos, Ignacio. (2013). El precio de la internacionalidad: la dualidad en los patrones de publicación de los investigadores españoles en comunicación. *Revista Española de Documentación Científica*, 36(2) <http://dx.doi.org/10.3989/redc.2013.2.936>
- Filippakou, Ourania y Tapper, Ted. (2008). Quality assurance and quality enhancement in higher education: Contested territories? *Higher Education Quarterly*, 62, 84–100. <http://dx.doi.org/10.1111/j.1468-2273.2008.00379.x>

- Fischman, Gustavo E. (2013). 'Hacerlo bien': acceso, visibilidad e impacto de la investigación latinoamericana. *Cuadernos del Pensamiento Crítico Latinoamericano*, 6.
- Flores, Ana María, Penkova, Snejanka y Román, Adelaida. (2009). Once años de LATINDEX: una experiencia al servicio de las publicaciones científicas iberoamericanas. *Simbiosis*, 6(1), 1–27.
- Fontela, Jorge L. y Genaro, Julio A. (2004). Biocomentarios. Acerca de los impactos. *Coyuco*, 14, 20–21.
- Frith, Chard. (2014). Action, agency and responsibility. *Neuropsychologia*, 55, 137–142. <http://dx.doi.org/10.1016/j.neuropsychologia.2013.09.007>
- Frumento, Antonio S. (1998). ¿Qué garantías nos da la ciencia? *El escéptico*, Invierno, 30–32.
- Fuentes, Miguel. (2014). El boicot a Elsevier y sus implicaciones respecto del acceso a las publicaciones científicas. En M. C. Di Gregori, L. Rueda, y L. Mattarollo (Eds.), *El conocimiento como práctica. Investigación, valoración, ciencia y difusión* (pp. 137–148). Buenos Aires: Universidad Nacional de La Plata.
- Galán, Arturo, González, María de los Ángeles y Román, Marcos. (2012). La irrupción del factor comunitario en el perfil del profesorado universitario. *Bordón*, 64(3), 133–148.
- García, María José. (2008). El impacto de la globalización en la universidad del siglo XXI. *Tendencias Pedagógicas*, 13, 59–78.
- García-Quero, Fernando. (2014). Crisis y Universidad: de intelectuales a hacedores de 'papers'. *eldiario.es*, 30 de junio.
- Garfield, Eugene. (2007). The evolution of the Scienc Citation Index. *International Microbiology*, 10(1), 65–69. <http://dx.doi.org/10.2436/20.1501.01.10>
- Gasca-Pliego, Eduardo y Olivera-García, Julio César. (2011). Construir ciudadanía desde las universidades, responsabilidad social universitaria y desafíos ante el siglo XXI. *Convergencia. Revista de Ciencias Sociales*, 56, 37–58.
- Gaytán, Susana Pilar. (2012). Evaluar la docencia. Estudio meta-evaluador del desarrollo del proceso de encuestas al alumnado y su fiabilidad en el grado en biología de la Universidad de Sevilla. En Alfonso Castro (Ed.), *Calidad, evaluación y encuestas de la docencia universitaria* (pp. 41–53). Murcia: Ediciones Laborum.
- Gazzola, Michele. (2014). *Lingva justeco: kiel taksi gxin? La exemplo de la Euxropa Unio*. Rotterdam: UEA.
- Giménez-Toledo, Elea. (2015). La evaluación de la producción científica: breve análisis crítico. *Relieve*, 21(1), 1–9. <http://dx.doi.org/10.7203/relieve.21.1.5160>
- Giulianoti, Richard. (2007). Forms of glocalization: Globalization and migration strategies of Scottish football fans in Northamerica. *Sociology*, 41, 133–152. <http://dx.doi.org/10.1177/0038038507073044>
- Gómez, Jesús, Latorre, Antonio, Sánchez, Montse y Flecha, Ramón. (2006). *Metodología comunicativa crítica*. Barcelona: El Roure Editorial.
- Gómez, Lucía y Jódar, Francisco. (2013). Ética y política en la universidad española: la evaluación de la investigación como tecnología de la subjetividad. *Athena Digital*, 13(1), 81–98.
- González, Ignacio. (2004). *Calidad en la universidad. Evaluación e indicadores*. Salamanca: Ediciones de la Universidad de Salamanca.
- Gracia, J. (2014). El sistema no son ellos. *El País*, 20 de marzo.
- Granda, José Ignacio. (2003). Algunas reflexiones y consideraciones sobre el factor de impacto. *Archivos de Bronconeumología*, 39(9), 409–417.
- Greenwood, David J. (2012). Doing and learning action research in the neo-liberal world of contemporary higher education. *Action Research*, 10(2), 115–132. <http://dx.doi.org/10.1177/1476750312443573>
- Grinberg, Silvia M. (2009). Tecnología del gobierno de sí en la era del gerenciamiento: la autoayuda entre el narcisismo y la abyección. *Psicoperspectivas*, 8(2), 293–308.
- Hanafi, Sari. (2011). University systems in the Arab East: Publish globally and perish locally vs publish locally and perish globally. *Current Sociology*, 59(3), 291–309. <http://dx.doi.org/10.1177/0011392111400782>
- Hicks, Diana, Wouters, Paul, Waltman, Ludo, de Rijcke, Sarah y Rafols, Ismael. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431.
- Hoeffel, C. (1998). Journal Impact Factors. *Allergy*, 53(12), 1225–1225.
- Iborra, María. (2014). Hacia una teoría ética de identificación y relevancia de los grupos de interés: responsabilidad, intencionalidad y previsión, poder y dependencia, urgencia y vulnerabilidad. *GCG. Globalización, Competitividad y Goberabilidad*, 8(2), 87–101. <http://dx.doi.org/10.3232/GCG.2014.V8.N2.05>
- Irigoyen, Juan. (2011). La refundación de la evaluación. *Laberinto*, 33, 35–44.
- Jarab, Josef. (2008). Reforming systems and institutions of higher education: Towards the creation of a European and global higher education area. *Education, Citizenship and Social Justice*, 3, 85–96. <http://dx.doi.org/10.1177/1746197907086720>
- Jáuregui, Ramón N. (2004). El mito de la excelencia en la universidad. *Educere*, 8, 235–238.
- Jorge, Elena. (2003). Nuevos elementos para la reflexión metodológica en sociología. Del debate cuantitativo/cualitativo al dato complejo. *Papers*, 70, 57–81.

- Jurdi, Andrea Perosa Saigh y Amiralian, María Lucía Toledo Moraes. (2013). *Ética do cuidado: a brinquedoteca como espaço de atenção a crianças em situação de vulnerabilidade*. *Interface (Botucatu)*, 17(45), 275–285. <http://dx.doi.org/10.1590/S1414-32832013005000009>
- Kalliola, Satu. (2009). Learning along with participatory action research — A Finnish perspective. *International Journal of Action Research*, 5(3), 289–321.
- Kennedy, Nilgün Fehim, Senses, Nazlı y Ayan, Pelin. (2011). Grasping the social through movies. *Teaching Higher Education*, 16, 1–14. <http://dx.doi.org/10.1080/13562517.2010.507305>
- Kidd, Sean A. (2002). The role of qualitative research in psychological journals. *Psychological Methods*, 7(1), 126–138. <http://dx.doi.org/10.1037/1082-989X.7.1.126>
- Koutny, Ilona. (2005). Interkultura komunikado en Eŭropo: la angla kaj esperanto kiel alternativaj komunikiloj. En Ch. Kinselman (Ed.), *Simpomio pri interkultura komunikado* (pp. 115–132). Dobrichovice: Kava-Pech.
- Kreimer, Pablo. (2011). La evaluación de la actividad científica: desde la indagación sociológica a la burocratización, Dilemas actuales. *Propuesta Educativa*, 36(2), 59–77.
- Krell, Frank. (2002). Why impact factors don't work for taxonomy. *Nature*, 415, 957–960.
- Lacey, Joseph. (2015). Considerations on English as a Global Lingua Franca. *Political Studies Review*, 13(3), 363–372. <http://dx.doi.org/10.1111/1478-9302.12004>
- Lange, Lidia L. (2002). The impact factor as a phantom. Is there a self-fulfilling prophecy effect of impact? *Journal of Documentation*, 58(2), 175–184. <http://dx.doi.org/10.1108/00220410210425449>
- Leal, Eduardo. (2009). La investigación acción participación, un aporte al conocimiento y a la transformación de Latinoamérica, en permanente movimiento. *Revista de Investigación*, 33(67), 13–34.
- León, Joaquín y Magriñá, Marta. (2004). Dialnet, una herramienta virtual de revistas hispanas sobre la base de la cooperación bibliotecaria. *El Profesional de la Información*, 13(4), 281–283.
- León-Correa, Francisco J. (2008). Ética del cuidado feminista y bioética personalista. *Persona y Bioética*, 12(30), 52–61.
- Liu, Amy. (2011). Unraveling the myth of meritocracy within the context of US higher education. *Higher Education*, 62(4), 383–397. <http://dx.doi.org/10.1007/s10734-010-9394-7>
- López, Santos. (2013). La calidad de las universidades públicas estatales en México desde la perspectiva de un multi-ranking. *Revista de la Educación Superior*, 42(2), 57–80.
- Macfarlane, Bruce. (2005). The disengaged academic: The retreat from citizenship. *Higher Education Quarterly*, 59, 296–312. <http://dx.doi.org/10.1111/j.1468-2273.2005.00299.x>
- Madorrán, Carmen. (2012). Filosofías para la liberación latinoamericana. *Bajo Palabra. Revista de Filosofía*, 7, 505–514.
- Majhanovich, Suzanne. (2014). Neo-liberalism, globalization, language policy and practice issues in the Asia-Pacific region. *Asia Pacific Journal of Education*, 34(2), 168–183. <http://dx.doi.org/10.1017/S0047404513000328>
- Malecki, Edward S. (2000). Teaching/Power. Part II. *Radical Pedagogy*, 2.
- Manzano-Arrondo, Vicente. (2010). La realidad significativa de investigación a través de las bases de datos. *Revista Salvadoreña de Psicología*, 1(2), 20–34.
- Manzano-Arrondo, Vicente. (2011). El papel de la universidad en la sociedad de ignorantes. *Científica*, 12, 29–55.
- Manzano-Arrondo, Vicente. (2012). *La Universidad Comprometida*. Vitoria: Hegoa.
- Manzano-Arrondo, Vicente. (2015a). Academia, evaluación y poder. *RASE*, 8(2), 197–222.
- Manzano-Arrondo, Vicente. (2015b). Activismo frente a norma: ¿quién salva a la universidad? *RIDAS. Revista Iberoamericana de Aprendizaje Servicio*, 1, 28–55.
- Manzano-Arrondo, Vicente. (2016). La calidad de la docencia universitaria ¿exclusiva responsabilidad del personal docente? En A. M. Chocrón (Ed.), *Calidad, docencia universitaria y encuestas: Bolonia a coste cero*. (pp. 45–53). Sevilla: AMLA.
- Manzano-Arrondo, Vicente y Suárez, Enrique. (2015). Unidad de Acción Comprometida: una propuesta de solución ante el problema universitario del servicio a la sociedad. *Hábitat y Sociedad*, 8, 147–166.
- Manzo, Silvia. (2014). Dimensiones colectivas del conocimiento en la Modernidad. En M. C. Di Gregori, L. Rueda, y L. Mattarollo (Eds.), *El conocimiento como práctica. Investigación, valoración, ciencia y difusión* (pp. 169–183). Buenos Aires: Universidad Nacional de La Plata.
- Márquez, María Luz. (2007). Metodología cualitativa o la puerta de entrada de la emoción en la investigación científica. *Liberabit*, 13, 53–56.
- Martín, A., Conde, J. y Baute, M. (2012). Criterios que utiliza el alumnado para la evaluación del profesorado. En A. Castro (Ed.), *Calidad, evaluación y encuestas de la docencia universitaria* (pp. 331–342). Murcia: Ediciones Laborum.
- Martin, Ben. (1993). The critique of science becomes academic. *Science, Technology, & Human Values*, 18(2), 247–259. <http://dx.doi.org/10.1177/016224399301800208>
- Martín-Baró, Ignacio. (1986). Hacia una psicología de la liberación. *Boletín de Psicología de El Salvador*, 22, 219–231.

- Martínez, Felipe. (2011). Los rankings de universidades: una visión crítica. *Revista de la Educación Superior*, 40(1), 77–97.
- Martínez, Miguel. (2006). La investigación cualitativa (síntesis conceptual). *Revista IIPSI*, 9(1), 123–146.
- Mateo, Josep Lluís. (2012). Les revistes d'impacte i el nou ordre científic mundial. *L'Avenç*, 383, 41–49.
- Matías, A. (2013). La evaluación de revistas de psicología: correlación entre el factor de impacto, el índice h y los criterios Latindex. *Investigación Bibliotecológica*, 27(61), 15–27.
- McArthur, Jan. (2011). Exile, sanctuary and diaspora: Mediations between higher education and society. *Teaching Higher Education*, 16, 579–589. <http://dx.doi.org/10.1080/13562517.2011.560264>
- Meneghini, Rogerio y Packer, Abel L. (2007). Is there science beyond English? *EMBO reports*, 8(2), 112–116. <http://dx.doi.org/10.1038/sj.embor.7400906>
- Miguel, Mario y Apodaca, Pedro. (2009). Criterios para evaluar el impacto de los Planes de Evaluación de la Calidad de la Educación Superior. *Revista de Educación*, 349, 295–310.
- Mollis, Marcela. (2014). Administrar la crisis de la educación pública y evaluar la calidad universitaria en América Latina: dos caras de la misma reforma educativa. *Revista de la Educación Superior*, 43(1), 25–45.
- Montalbá, Carmen. (2015). El sueño de la excelencia. *Arxius de Ciències Socials*, (32), 83–98.
- Moreno-Pulido, A., López-González, M. A., Rubio-Garay, F., Saúl, L. A. y Sánchez-Elvira-Paniagua, A. (2013). Evolución de las revistas españolas de Ciencias Sociales en el Journal Citation Reports (2006–2010) y su valoración como indicio de calidad en la normativa evaluadora española. *Revista Española de Documentación Científica*, 36(3) <http://dx.doi.org/10.3989/redc.2013.3.987>
- Morley, Louise y Aynsley, Sarah. (2007). Employers, quality and standards in higher education: Shared values and vocabularies or elitism inequalities? *Higher Education Quarterly*, 61, 229–249. <http://dx.doi.org/10.1111/j.1468-2273.2007.00353.x>
- Muñoz, Humberto, García, Susana y López, Romualdo. (2014). La evaluación de los académicos. *Perfiles Educativos*, 36(146), 188–195.
- Musial, Kazimierz. (2010). Redefining external stakeholders in Nordic higher education. *Tertiary Education and Management*, 16, 45–60. <http://dx.doi.org/10.1080/13583881003629822>
- Naval, Concepción y Ruiz-Corbella, Marta. (2012). Aproximación a la responsabilidad social universitaria: la respuesta de la universidad a la sociedad. *Bordón*, 64(3), 103–115.
- Noguera, José Antonio. (2003). ¿Quién teme al individualismo metodológico? Un análisis de sus implicaciones para la teoría social. *Papers*, 69, 101–132.
- Noreña, Ana Lucía, Alcaraz-Moreno, Noemí, Rojas, Juan Guillermo y Rebolledo-Malpica, Dinora. (2012). Aplicabilidad de los criterios de rigor y éticos en la investigación cualitativa. *Aquichán*, 12(3), 263–274.
- Ordine, Nuccio. (2013). *La utilidad de lo inútil. Manifiesto*. Barcelona: Acantilado.
- Packer, Abel L., Luccisano, A., Montarani, F., Santos, S. y Meneghini, R. (2006). El modelo SciELO de publicación científica de calidad en acceso abierto. En D. Babini y J. Fraga (Eds.), *Edición electrónica, bibliotecas virtuales y portales para las Ciencias Sociales en América Latina y el Caribe* (pp. 191–208). Buenos Aires: CLACSO.
- Palafox, Juan Carlos. (2009). Breves reflexiones sobre calidad educativa. En J. E. Villa y C. Núñez (Eds.), *Compromiso social por la calidad de la educación* (pp. 57–60). México: Instituto Politécnico Nacional.
- Paloma, Virginia y Manzano-Arrondo, Vicente. (2011). The role of organizations in liberation psychology: Applications to the study of migrations. *Psychosocial Intervention*, 20(3), 309–318. <http://dx.doi.org/10.5093/in2011v20n3a7>
- Pelekakis, Cira de y Aguirre, René. (2010). Gestión universitaria socialmente responsable: un camino efectivo hacia el servicio comunitario. *REDHECS. Revista Electrónica de Humanidades, Educación y Comunicación Social*, 5(8), 90–101.
- Pérez, Gloria. (1994). Investigación cualitativa. In *Retos e interrogantes. Vol II. Técnicas y análisis de datos*. Madrid: La Muralla.
- Pérez, Miguel Ángel. (2006). Usos y abusos de la cienciometría. *Cinvestav, Enero-Marzo*, 29–33.
- Puelles, Manuel. (2009). Calidad reformas escolares y equidad social. En M. A. Santos (Ed.), *Políticas educativas y compromiso social. El progreso de la equidad y la calidad*. (pp. 25–39). Madrid: Octaedro Editorial.
- Quintana, Alberto. (2006). Metodología de la investigación científica cualitativa. En A. Quintana y W. Montgomery (Eds.), *Psicología: tópicos de actualidad* (pp. 47–84). Lima: UNMSM.
- Quirós, Fernando. (2014). Prisioneros de la Bibliometría. Contra lo cuantitativo como sistema de evaluación de la actividad investigadora. *Ámbitos*, 25, 1–12.
- Robinson, Susan R. (2014). Self-plagiarism and unfortunate publication: An essay on academic values. *Studies in Higher Education*, 39(2), 265–277. <http://dx.doi.org/10.1080/03075079.2012.655721>
- Ruiz, José Ignacio y Ispizúa, María Antonia. (1989). *La descodificación de la vida cotidiana. Métodos de investigación cualitativa*. Bilbao: Universidad de Deusto.

- Saainen, Taina. (2009). Whose quality? Social actors in the interface of transnational and national higher educational policy. *Discourse: Studies in Cultural Politics of Education*, 29, 179–193. <http://dx.doi.org/10.1080/01596300801966807>
- Sandín, Mary Paz. (2000). Criterios de validez en la investigación cualitativa: de la objetividad a la solidaridad. *Revista de Educación Educativa*, 18(1), 223–242.
- Santa, Samaly y Herrero-Solana, Víctor. (2010). Cobertura de la ciencia de América Latina y el Caribe en Scopus vs Web of Science. *Investigación Bibliotecológica*, 24(52), 13–27.
- Sanz-Casado, Elías, García-Zorita, Juan Carlos, Serrano-López, Antonio Eleazar, Efraín-García, Preiddy y de Filippo, Daniela. (2013). Rankings nacionales elaborados a partir de múltiples indicadores frente a los de índices sintéticos. *Revista Española de Documentación Científica*, 36(3), 1–18.
- Scobey, David. (2011). Journals and journeys: New horizons in public scholarship. *Journal of Public Scholarship in Higher Education*, 1(1), 7–9.
- Shaw, Claire y Ward, Lucy. (2014). Dark thoughts: Why mental illness is on the rise in academia. *The Guardian. Higher Education Network*, 1, 6 de marzo.
- Shields, Carmen y McGinn, Michelle K. (2011). The erosion of academic troth: Disengagement and loss. *Teaching Higher Education*, 16, 471–482. <http://dx.doi.org/10.1080/13562517.2011.570441>
- Shumar, Wesley. (2004). Global pressures, local reactions: Higher education and neo-liberal economic policies. *International Journal of Qualitative Studies in Education*, 17, 823–839. <http://dx.doi.org/10.1080/0951839042000256475>
- Sívori, Luchino. (2012). La evaluación de la calidad académica y el rol de las ciencias sociales. *Revista Científica de Información y Comunicación*, 9, 103–129.
- Speck, Bruce W. (2001). Why Service-Learning? *New Directions for Higher Education*, 114, 3–13. <http://dx.doi.org/10.1002/he.8.abs>
- Spinak, Ernesto. (1998). Indicadores cienciométricos. *Ciência da Informação*, 27(2), 141–148.
- Spinoza, Oscar. (2008). Creating (in) equalities in access to higher education in the context of structural adjustment and post-adjustment policies: The case of Chile. *High Education*, 55, 269–284. <http://dx.doi.org/10.1007/s10734-007-9054-8>
- Stack, Michelle. (2016). *Global University Rankings and the Mediatization of Higher Education*. Hampshire: Palgrave Macmillan.
- Stensaker, Bjorn, Frolich, Noline, Gornitzka, Ase y Maasen, Peter. (2008). Internationalisation of higher education: The gap between national policy-making and institutional needs. *Globalisation, Societies and Education*, 6, 1–11. <http://dx.doi.org/10.1080/14767720701855550>
- Susanti, Dewi. (2011). Privatisation and marketisation of higher education in Indonesia: The challenge for equal access and academic values. *Higher Education*, 62, 209–218. <http://dx.doi.org/10.1007/s10734-010-9333-7>
- Teelken, Christine. (2012). Compliance or pragmatism: How do academics deal with managerialism in higher education? A comparative study in three countries. *Studies in Higher Education*, 37, 271–290. <http://dx.doi.org/10.1080/03075079.2010.511171>
- Temple, P. (2005). The EFQM excellence model: Higher education's latest management fad? *Higher Education Quarterly*, 59, 261–274. <http://dx.doi.org/10.1111/j.1468-2273.2005.00297.x>
- Thiollent, Michel. (2011). Action research and participatory research: An overview. *International Journal of Action Research*, 7(2), 160–174.
- Torres-Salinas, Daniel, Ruiz-Pérez, Rafael y Delgado-López-Cózar, Emilio. (2009). Google Scholar como herramienta para la evaluación científica. *El Profesional de la Información*, 18(5), 501–510. <http://dx.doi.org/10.3145/epi.2009.sep.03>
- UNESCO. (1998). *Conferencia Mundial sobre la Educación Superior. La educación superior en el siglo XXI. Visión y acción*. París: Ediciones de la UNESCO.
- Urcelay, Carlos y Galetto, Leonardo. (2011). ¿Editar o no editar? Reflexiones sobre las revistas científicas regionales y algunas propuestas. *Kurtziana*, 36(1), 3–7.
- Usher, Alex y Savino, Massimo. (2006). Estudio global de los rankings universitarios. *Calidad de la Educación*, 25, 33–53.
- Valero, Jesús Antonio, Jiménez, Elena y Coca, Juan R. (2013). La perspectiva socio-hermenéutica en la evaluación del trabajo científico en España. *Investigación Bibliotecológica*, 27(59), 15–34.
- Vallaey, François. (2010). ¿Qué es la responsabilidad social universitaria? (consultado 6 Abr 2017). Disponible en: <http://creasfile.uahurtado.cl/RSU.pdf>.
- Vallaey, François. (2011). La responsabilidad social universitaria: ¿cómo entenderla para quererla practicar? (consultado 6 Abr 2017). Disponible en: <http://blog.pucp.edu.pe/blog/eticarsu/2007/11/26/como-entender-la-responsabilidad-social-universitaria-para-quererla>.

- Vázquez, Carlos, Escabí, Aracelis, Quiñones, Sylmarie y Pacheco, Wanda. (2012). El refortalecimiento como una herramienta de trabajo comunitario: reflexiones desde la comunidad. En A. Zambrano y H. Berroeta (Eds.), *Teoría y práctica de la acción comunitaria* (pp. 257–276). Santiago (Chile): RIL Editores.
- Vázquez-Cano, Esteban. (2013). El videoartículo: nuevo formato de divulgación en revistas científicas y su integración en MOOCs. *Comunicar*, 21(41), 83–91. <http://dx.doi.org/10.3916/C41-2013-08>
- Venieris, George y Cohen, Sandra. (2004). Accounting reforms in Greek universities: A slow moving process. *Financial Accountability & Management*, 20, 183–204. <http://dx.doi.org/10.1111/j.1468-0408.2004.00192.x>
- Verma, Inder M. (2015). Impact, not impact factor. *PNAS*, 112(26), 7875–7876.
- Villaseñor, Jorge Ignacio, Moreno, Carlos Iván y Flores, Jorge Enrique. (2015). Perspectivas actuales sobre los rankings mundiales de universidades. *Revista de la Educación Superior*, 42(3), 41–67. <http://dx.doi.org/10.1016/j.resu.2015.09.001>
- Vito, Eduardo Luis. (2006). Algunas consideraciones en torno al uso del Factor de Impacto y de la bibliometría como herramienta de evaluación científica. *Revista Argentina de Medicina Respiratoria*, 1, 37–45.
- Watermeyer, Richard. (2014). Issues in the articulation of ‘impact’: The responses of UK academics to ‘impact’ as a new measure of research assessment. *Studies in Higher Education*, 39(2), 359–377. <http://dx.doi.org/10.1080/03075079.2012.709490>
- Wilsdon, James. (2015). *The Metric Tide. Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. London: HEFCE. <http://dx.doi.org/10.13140/RG.2.1.4929.1363>
- Winn, Robert C. y Green, Robert S. (1998). Applying total quality management to the educational process. *International Journal of Engineering Education*, 14, 24–29.
- Worthington, Andrew Charles y Higgs, Helen. (2011). Economies of scale scope in Australian higher education. *Higher Education*, 62, 387–414. <http://dx.doi.org/10.1007/s10734-010-9337-3>
- Yamamoto, Kiyoshi. (2004). Corporatization of national universities in Japan: Revolution for governance or rhetoric for downsizing? *Financial Accountability & Management*, 20, 153–181. <http://dx.doi.org/10.1111/j.1468-0408.2004.00191.x>
- Yogev, Esther y Michaeli, Nir. (2011). Teachers as society-involved “organic intellectuals”: training teachers in a political context. *Journal of Teacher Education*, 62, 312–324. <http://dx.doi.org/10.1177/0022487110397841>