

Mesa Temática N°7: Evaluación académica y aseguramiento de la calidad en el sistema universitario: políticas, prácticas y culturas evaluativas.

Título: El debate de la evaluación académica y la viabilidad de producir cambios en las universidades argentinas

Autores:

Fernanda Beigel INCIHUSA-CONICET, CECIC-UNCuyo, fernandabeigel@gmail.com

Aníbal Bar. Universidad Nacional del Nordeste. anibalrbar@hum.unne.edu.ar

Jaquelina E. Noriega Universidad Nacional de San Luis, jenoriega@email.unsl.edu.ar

Introducción

Este trabajo se enmarca en un proyecto en ejecución financiado a partir de la Convocatoria PISAC 2022 cuyo objetivo general orienta hacia la producción de un diagnóstico de la evaluación científico tecnológica en la Argentina con la finalidad de identificar las posibilidades existentes, en el contexto actual de sus universidades, para promover un cambio sustancial e informado que permita avanzar en prácticas responsables de investigación.

Esta ponencia ofrece una sistematización y análisis de la literatura existente acerca de la evaluación de la ciencia, dando cuenta del debate actual. No sólo se retoman discusiones presentes en la bibliografía disponible sino, además, se revisan estudios empíricos realizados recientemente en el país. Este estado de la cuestión, reseña las discusiones enmarcadas en la idea de excelencia académica, las diferencias entre impacto académico y relevancia social de la investigación, así como los principios de evaluación responsable de la investigación.

Está bien documentado que la evaluación de la ciencia, en el sentido de una evaluación de pares, es tan antigua como las revistas científicas y los congresos académicos pero fue especialmente a partir de la década de 1990 cuando se instauró una práctica de evaluación externa que fue condicionando el financiamiento de las instituciones y los proyectos, así como la promoción de la carrera académica de las personas. El principal instrumento de este giro evaluador en América Latina fue el establecimiento de normas de acreditación externa de las universidades e incentivos salariales para fomentar la

investigación. Una injerencia heterónoma que en gran medida fue eficaz, aunque con matices diferenciales en cada país, según la fuerza de la autonomía institucional de las universidades, el papel jugado por los consejos rectorales, el peso del movimiento estudiantil y la incidencia de los gremios docentes. En sintonía con lo que ocurrió en otras latitudes, las políticas de evaluación de la calidad en nuestra región modificaron las prácticas de investigación y tuvieron efectos directos en el ingreso o permanencia en la docencia, organizando segmentos profesionales con diferente jerarquía (Araujo, 2003; Naidorf y Perrota, 2016). La particular forma en que cada país desplegó esos incentivos y el sistema de evaluación que implantó para distribuirlos marca situaciones diferenciales. En diez países de nuestro continente estos sistemas de evaluación derivaron en la construcción de *sistemas nacionales de categorización* que Vasen et al (2021) han comparado recientemente.

Este y otros estudios que han observado las experiencias de evaluación de la ciencia y la tecnología en América Latina, han demostrado que uno de los principales problemas es que, en la práctica, los esquemas de ponderación de las trayectorias académicas se reducen finalmente a la producción científica, entendida ésta casi exclusivamente en términos de publicaciones. Esto significa, primeramente, desconocer el peso de las actividades de extensión, transferencia, gestión, docencia y formación de recursos humanos (Echeverry-Mejía, 2021; Schimanski y Alperin, 2018). Pero, además, tiene otras implicancias, porque ponen el foco en la productividad y esto refuerza las asimetrías de género. Hay estudios que muestran que estas manifestaciones de la brecha de género están enraizadas en factores estructurales, pero también en dinámicas endógenas del campo académico, como el reparto del capital social que circula en el proceso de formación doctoral, en las redes y proyectos de los equipos de investigación, todos marcados por una acumulación o des-acumulación sexuada (Beigel y Gallardo, 2021; Beigel et al 2023).

Este conjunto de asimetrías y desigualdades que atraviesan la producción científica a nivel intra e inter-nacional también ponen sobre el tapete las bases de datos que nutren los estudios e indicadores cuantitativos. Nos referimos a los sesgos ya suficientemente estudiados de sistemas de indexación producidos por empresas comerciales, como Scopus y Web of Science. Los artículos incluidos en esos sistemas fueron identificados como los únicos bienes de calidad internacional “exportados” por los países, relegando cada vez más la producción publicada en revistas nacionales, con el menosprecio que

trajo aparejada su desvalorización en los procesos de evaluación. Esos indicadores de producción científica fueron determinantes para la evolución de las prácticas científicas y tuvieron una incidencia decisiva para la deriva de las agendas de investigación (Invernizzi, 2022).

La búsqueda de la excelencia a través de indicadores de producción publicada, adoptado de manera generalizada, como un estándar transnacional, afecta de modo directo la promoción de la producción tecnológica. Se trata de una actividad diferenciada, que es llevada a cabo por personal académico en las universidades y el CONICET, pero además constituye el objeto principal de organismos públicos como el INTI, el INTA, la CNEA, entre otros. Sin embargo, los indicadores actualmente utilizados para evaluar carreras, proyectos e instituciones no reflejan todavía esa diversidad de perfiles y requieren de mayores esfuerzos y discusión.

Son muchas las voces institucionales que se alzan a nivel global demandando una transformación urgente de los sistemas de evaluación académica, propugnando una atención mayor a la calidad de los resultados de la investigación que a la cantidad, el uso contextualizado de los indicadores y una evaluación responsable, basada en procesos diversificados y adaptados a distintos perfiles y etapas de la trayectoria académica (DORA, 2012; FOLEC, 2020). En lo que sigue, analizaremos los principales nudos conceptuales que rodean, a nivel global, regional y nacional, la cuestión de la evaluación responsable de la investigación, para finalmente focalizar en las posibilidades de un cambio en las prácticas de evaluación en las universidades nacionales.

¿Qué es la excelencia académica?

Arocena y Sutz (2020) se pregunta por qué las universidades del Sur Global corren la misma carrera que las que están en el top de los rankings universitarios, si éstas no participan por lo general en ese mercado de servicios de educación superior. La respuesta principal es que esas universidades y su performance se constituyen en una suerte de faro indicando hacia dónde todo el mundo debería llegar. Ese faro indica una forma de validación de la calidad académica típica de esas instituciones del Norte acostumbradas a publicar en revistas editadas también en países hegemónicos que son las que definen el reconocimiento internacional. Son los sistemas de evaluación los que

se orientaron por ese camino y fueron ampliando esos criterios en universidades con contextos totalmente diferentes.

Este paradigma estandarizado de excelencia, supuestamente global, es problemático para evaluar la investigación producida en los países del Sur porque no es allí donde se originaron esos estándares y, por lo tanto, tienen una parte de extrañeza y otra parte del “deber ser” que los propios investigadores se auto-imponen. Hay algunos sesgos de esa noción de excelencia que se han vuelto cada vez más evidentes. Por ejemplo, la ventaja acumulativa de ciertos países e instituciones que pueden considerarse excelentes, dada la forma en que los recursos de investigación se distribuyen, incluida la financiación directa y el acceso a la infraestructura (equipo, suscripciones a bibliotecas, etc.) que genera asimetrías estructurales en términos de formación de equipos, redes y publicación científica. Además, los académicos y científicos de los países de bajos ingresos tienden a enfrentar obstáculos adicionales en el desarrollo de su carrera (falta de movilidad, mayor carga docente) que restringen su capacidad para publicar prolíficamente y promover sus publicaciones. Cada vez hay mayor consenso sobre la idea de que la identificación entre excelencia académica y productivismo ha tenido consecuencias negativas para la calidad intelectual, humana y laboral de las universidades (Faber, 2021).

La excelencia implica determinación por comparación y, por tanto, competencia (por la financiación de la investigación, por las publicaciones en las mejores revistas, etc.). No es sorprendente que a menudo se entienda que la excelencia tiene que ver con la ciencia de élite. Dentro de las principales recompensas de estos sistemas de evaluación se encuentran los incentivos salariales basados en una dotación monetaria como recompensa a cierta performance individual. Estos incentivos generaron un lazo indisoluble entre excelencia y productividad cuyos efectos nocivos han sido ampliamente analizados, tanto por las distorsiones que produjeron estas presiones sobre el profesorado universitario como por el malestar que se extendió cada vez más por los pasillos de las Facultades (Giménez Toledo, 2016). Esos "mejores" investigadores no solo son maestros en campos especializados, sino que también son creativos y originales. Están bien posicionados para determinar lo que se necesita hacer en ciencia y se les debe ofrecer financiamiento para sus propuestas de investigación.

En América Latina y el Caribe hace varias décadas que se viene planteando una discusión acerca de la diferencia entre “excelencia” y “calidad académica”, por cuanto la primera se terminó construyendo en base a indicadores de citación de las revistas, alejándose cada vez más de la idea de calidad de cada contribución científica y, mucho más lejos aún, respecto de su relevancia social (Latindex, 2011; Vessuri, Guédon y Cetto, 2014). Las recompensas ofrecidas por los sistemas de evaluación basados en los índices de impacto tuvieron una incidencia cada vez más significativa en la construcción de agendas científicas institucionales y nacionales, forzándolas hacia temáticas internacionalizada para asegurar las publicaciones. En el contexto latinoamericano, esto ha desestimulado la interdisciplinaridad y la creatividad, impulsando la primacía del inglés y reforzando una tendencia persistente a producir una ciencia poco aplicable (Invernizzi 2022). Estos efectos nocivos de la imposición de un concepto unívoco de excelencia académica plantean la necesidad de revisar la noción de “impacto” de la ciencia, preguntándonos para qué los gobiernos hacen importantes esfuerzos para financiar el avance científico y tecnológico.

En los últimos años se ha revitalizado la preocupación por la utilidad de la ciencia y la duda acerca de si la inversión científica de los países en verdad trae algún beneficio concreto para el desarrollo, la expansión productiva o la calidad de vida de la población. En definitiva, apunta a determinar si tal inversión tiene retorno y cómo medirlo (Piovani, 2019). Este es un asunto de antigua data que acompañó el proceso de institucionalización de la ciencia, convirtiéndose en alter ego de la demanda de libertad académica para determinar agendas de investigación. Y se desplegó también como una tensión que formó parte del proceso de desarrollo de las funciones de transferencia y extensión en las universidades de la región.

Por supuesto, la pregunta por la utilidad o la relevancia social de la ciencia tiene particularidades y riesgos para las ciencias sociales y las humanidades, porque pueden ser objeto de un análisis también productivista, esta vez orientado a un conocimiento capaz de generar productos o servicios de interés mercantil. De hecho, la difusión de la idea de ciencia aplicada como prerrogativa de las ingenierías, las ciencias exactas y naturales, tuvo como efecto directo el desconocimiento del potencial que las ciencias sociales y humanas tienen para intervenir en los problemas de la sociedad. Efectivamente, hay un gigantesco stock de información sobre los temas más diversos y una producción considerable que se nutre de un flujo que muchas veces queda dentro de

las fronteras académicas. Eso se debe a la conjunción de la falta de hábito por parte del mundo científico de dialogar con otros saberes y otras audiencias, y de la dificultad del mundo de la gestión gubernamental para tomar decisiones informadas.

La evaluación de pares en discusión: evolución, nuevas tendencias y buenas prácticas.

La evaluación de pares ha sido concebida por lo general como la piedra fundacional del sistema de acreditación académica y de valoración de la calidad de la producción científica. La evolución de esa particular tecnología de la dictaminación es compleja, registra diferencias disciplinares, geográficas e institucionales. Existen evaluaciones de pares realizadas por especialistas externos a la institución y otras que exigen externalidad respecto del país de origen del objeto analizado. En lo concerniente al anonimato de un referato Pontille y Thorny (2020) mencionan cuatro modalidades: doble ciego, evaluación ciega, ciega unilateral y la más reciente evaluación abierta. Ahora bien, conviene analizar el papel de los pares especialistas en sus distintos ámbitos de actuación que son muy variados, pero nos ocuparemos aquí de tres principales por su relación con las dimensiones de la cartografía propuesta en este proyecto PISAC: a) referatos de revistas y libros, b) comités evaluadores de concursos docentes o carreras de investigación y c) pares convocados para evaluar proyectos de investigación.

Históricamente, ser convocado como experto para la evaluación de un artículo en una revista especializada era considerado un reconocimiento y una obligación, en un ámbito que se reproduce en base al prestigio simbólico. Avanço et al (2021) hicieron una consulta a referees de revistas y la mayoría de las veces contestan que siguen haciendo esta tarea porque les permite conocer novedades de su campo de estudio y sienten que colaboran con el progreso científico, antes que por las recompensas monetarias que pueda reportar esa tarea (Avanço et al, 2021). Sin embargo, con el paso del tiempo, la recarga de tareas y la creciente mercantilización de las revistas, la evaluación de artículos encuentra crecientes dificultades, tanto en la perspectiva de los investigadores como para los editores.

La tarea de gestión de la evaluación es más compleja para los editores de las revistas universitarias que para aquellas que son gestionadas por editoriales comerciales. Esto no sólo obedece a que se trata de equipos editoriales menos profesionalizados sino a que la mayoría usa el sistema de doble ciego que es informado en el artículo, mientras otras

sólo recurren a un dictamen de evaluación y no siempre ciega. Aguado López y Becerril García (2021) señala que las revistas diamante de América latina, sostenidas por las instituciones universitarias, toda la gestión es un reto importante, porque cuentan con personal académico con cargo en la institución, pero normalmente la gestión administrativa y la infraestructura que utilizan se comparte con muchas otras actividades propias de la institución

Analizando los resultados de la encuesta aplicada por LatinREV a equipos editoriales de 219 revistas universitarias afiliadas a la Red, Ruiz del Corda, y Ferrier (2020) observan que los docentes e investigadores prestan parte de su tiempo para participar como directivos o editores de las revistas y en sus comités científicos, sin que esta actividad sea suficientemente reconocida en los sistemas de evaluación de las instituciones a las que pertenecen. Esto genera frecuentemente cambios en los equipos editoriales que pueden producir irregularidades en la publicación y continuidad de los procesos.

En las revistas gestionadas por editoriales comerciales la carga de trabajo obedece a un flujo mayor de manuscritos para ser evaluados y la necesidad de contar con más evaluadores, asimismo, más eficaces. Esto impulsó diferentes formas de recompensa, descuentos, vouchers y otras formas de incentivo para asegurar las dictaminaciones esperadas.

Si pasamos ahora al libro académico, la problemática de la evaluación de pares cambia significativamente, porque la indexación de este formato de publicación es mucho más reciente a nivel internacional y regional. En América Latina, existe una tradición editorial de largo aliento que ha favorecido la publicación de libros y, consecuentemente, el desarrollo de formas de prestigio impulsadas por redes, centros regionales, editoriales especializadas. Contribuyen decididamente también las editoriales universitarias y el dinamismo de la producción de las ciencias sociales y humanas en la región.

En relación con la evaluación de carreras académicas y de investigación, en América Latina existe un modelo regional de sistemas de categorización de investigadores que, según Vasen et al (2021), existe en once países de la región <https://impactoabierto.org/mapa/> Estos sistemas tienen un papel estandarizador muy importante, tanto entre disciplinas como dentro de ellas y cumplen tres objetivos principales. El primero está ligado a la sistematización de información respecto a las

personas que realizan investigación en el país, su inscripción institucional y orientación disciplinar. Una segunda finalidad consiste en diferenciar dentro de un conjunto más amplio de personas asociadas a la ciencia y la educación superior, aquellos que hacen verdaderamente investigación, por sobre aquellos que se supone que la hacen a medias, o simulan hacerla. En tercer lugar, existe un componente de reconocimiento. Aquellos que ingresan al sistema, obtienen distintas retribuciones, que pueden ser simbólicas (ocupar segmentos de prestigio o habilitación para dirigir proyectos) o económicas (complementos salariales).

Para producir estas clasificaciones, estos sistemas de categorización de investigadores aplican cuatro procesos comunes a todos los países donde se implementa. Hay un requisito de entrada que consiste en la afiliación a una universidad o institución que realiza investigación. El proceso de evaluación de las postulaciones se realiza de forma centralizada, atendiendo a particularidades disciplinares antes que institucionales o regionales. Por lo general, reina el uso de los indicadores de impacto de las revistas indexadas en Web of Science o Scopus, según el país, y esas métricas constituyen un aspecto determinante del orden de méritos establecido por las comisiones asesoras. El sistema otorga a los postulantes admitidos un reconocimiento simbólico, una “categoría” que se materializa en una forma de prestigio, una habilitación y/o mayor visibilidad dentro de la comunidad científica. El foco de las actividades evaluadas es la investigación, aunque las personas desarrollen además otras tareas como la enseñanza, extensión, transferencia, gestión (Vasen et al, 2021).

La evaluación responsable de la investigación y la controversia sobre los indicadores cuantitativos de producción

La reforma de los sistemas tradicionales de evaluación de la investigación son una herramienta fundamental para corregir los efectos nocivos de los indicadores de impacto de las publicaciones que terminaron por distorsionar las bases a la que apuntaba la evaluación académica, que era determinar la calidad y originalidad de una contribución científica (Larivière, y Gingras, 2010; Piro y Sivertsen, 2016; Ràfols, 2019; Molas-Gallart, Boni Aristizábal, Giachi y Schot; 2021). Una de las principales consecuencias de este sistema de publicaciones y sus rankings fue la consolidación de una industria del prestigio que comercializó el acceso a la ciencia por parte de la propia comunidad científica y la cercenó a la sociedad, con la consecuente aparición de prácticas de publicación cuestionables y revistas predatorias (Gingras, 2016; Biagioli y Lippman,

2020; Pölönen y Sivertsen, 2021). Confluyen en dirección crítica iniciativas internacionales como la Declaración de San Francisco sobre la Evaluación de la Investigación (DORA, 2012), la Iniciativa Helsinki sobre el Multilingüismo (2019)¹, el Manifiesto de Leiden (2015) y muchas otras.

Más allá de que mantiene un fuerte anclaje en Estados Unidos, DORA se convirtió en una organización internacional muy activa en el proceso de reflexión sobre la evaluación académica y sobre todo a nivel propositivo para producir cambios institucionales concretos. La idea de la “evaluación responsable” de la investigación es precisamente uno de los ejes principales del advocacy propuesto por la organización y consiste en poner una atención mayor a la calidad de los resultados de la investigación que a la cantidad. También se trata de la utilización de manera adaptada de indicadores y procesos diversificados, prescindiendo del factor de impacto de la publicación como una medida sustituta de la calidad de los artículos de investigación individuales, para evaluar las contribuciones de un científico individual, o en las decisiones de contratación, promoción o financiación. La declaración efectúa recomendaciones específicas para editores, organizaciones encargadas de elaborar métricas y otros agentes del mundo universitario. A las agencias de financiación les recomienda explicitar claramente los criterios utilizados para evaluar la productividad científica de los solicitantes de fondos de investigación, especialmente atendiendo a los investigadores que están iniciando su carrera.

La diatriba contra la tendencia cuantitativa en evaluación plantea su reemplazo por el enfoque cualitativo como si ésta fuese una solución total, mientras que tampoco resuelve todos los problemas. De hecho, los indicadores cuantitativos surgieron para objetivar aquello que las evaluaciones de pares parecían desmerecer, opacar o simplemente ocultar como resultado de conflictos de intereses. Delegar toda la responsabilidad de la evaluación en un comité de pares no asegura una evaluación responsable, porque todo comité funciona con un consenso intersubjetivo que depende de las creencias y trayectorias de esos pocos evaluadores.

Frente a la tendencia al uso de un único indicador de impacto tipo mainstream, surgen enfoques multidimensionales de la evaluación que ofrecen un mayor margen para promover una diversidad de perfiles y proyectos de acuerdo con agendas

¹ Ver <https://www.helsinki-initiative.org/es/info>

contextualizadas en escala nacional o local. Robinson-Garcia, Costas, Van Leeuwen y Nane (2019) proponen evaluar las prácticas científicas en relación con cuatro dimensiones: la trayectoria, el compromiso académico, el compromiso social y la construcción de capacidades. Cada una de esas dimensiones abre posibilidades de considerar variables del contexto, la vida personal o situaciones ligadas con el género.

Otro camino relevante para impulsar nuevas formas de evaluación abierta se aloja en los repositorios institucionales que están teniendo un rápido desarrollo a nivel global y también en América Latina al compás de la aplicación de leyes y políticas de ciencia abierta. Muchos repositorios han avanzado sistemáticamente en la cosecha y curaduría de la producción institucional, pero siguen usando protocolos y tecnologías diseñadas veinte años atrás, antes de la aparición de los grandes buscadores de internet.

La evaluación de la investigación en Argentina

A partir de 1990 Argentina adopta un modelo internacional de profesionalización académica basada fundamentalmente en dos ejes: la formación de posgrado y la investigación. Este modelo se constituyó en un marco regulador y ordenador de prioridades de las distintas prácticas que realizan los/as académicos/as dentro de las universidades argentinas. Sobre todo marcó el rumbo de la actividad de investigación y la formación académica asociada a tal actividad, signada fundamentalmente por la búsqueda de la formación de doctorado, por sus implicancias en la especialización de investigadores/as.

Esta reforma de la Educación Superior trajo aparejado el surgimiento de distintos programas y agencias que, de manera directa o indirecta, incidieron en la regulación de la actividad de investigación de los docentes universitarios (Moore, F., Rovelli, L. I. y Corengia, A., 2021). Algunos de ellos son: el Fondo para el Mejoramiento de la Calidad de la Educación Superior (FOMECE), el Programa de Incentivos a Docentes Investigadores (PROINCE), la creación de la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU), y el surgimiento de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT). El PROINCE fue determinante ya que implicaba la erogación de un plus salarial, que en sus inicios era muy importante, orientando la actividad académica fuertemente hacia la investigación y la producción de conocimientos que luego fueran publicados en revistas de impacto (Beigel y Bekerman, 2019).

Lo cierto es que este mecanismo orientado por el incentivo económico importante al inicio, resultó ser un "anzuelo" muy eficaz para organizar la investigación académica en la universidad con un único sentido, uniforme, unívoco y sin prácticamente fisuras. De hecho la investigación no asociada a los estándares propuestos por el PROINCE fueron casi inexistentes y por supuesto periféricos (Beigel y Bekerman, 2019).

Recientemente, y después de un largo periodo de estancamiento (última convocatoria realizada en 2014), el sistema de categorización de docentes investigadores tuvo una nueva convocatoria no solo al ingreso sino también a la recategorización. El nuevo programa se llamó PRINUAR (Programa para la investigación Universitaria Argentina). Si bien se tejieron debates en distintos ámbitos respecto a los lineamientos de esta Convocatoria, no hubo un cambio sustancial en los requerimientos y criterios de valoración de las trayectorias académicas. En la actualidad se encuentra en proceso de evaluación de postulaciones. La nueva Convocatoria significó una gran movilización al interior de las universidades dedicándose mucho tiempo y esfuerzo de trabajo al completamiento de los formularios asociados a la presentación, tiempo que obviamente se restó al desarrollo de otras tareas en lo cotidiano.

En síntesis, la evaluación de la actividad de investigación de los académicos en Argentina, en la actualidad, se realiza a través de dos mecanismos importantes:

- la categorización y recategorización del Programa de docentes investigadores de universidades argentinas, dependiente del Ministerio de Educación de la Nación (Programa para la investigación Universitaria Argentina - PRINUAR-)
- la categorización y recategorización de docentes investigadores, perteneciente al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) existente desde 1958.

Con respecto al primero, su objetivo es institucionalizar la figura del/a investigador/a universitario/a, con igualdad de género, federalismo y autonomía; atendiendo las líneas propuestas por las universidades y las políticas nacionales de desarrollo del conocimiento. Este Programa pretende articular la investigación con la docencia, la extensión y las artes, aportando una perspectiva integral para el desarrollo científico y tecnológico. El principal eje es la categorización de las/os docentes investigadores/as de Universidades Nacionales Argentinas, a partir de la aplicación de un instrumento de evaluación de desempeño estandarizado donde el/la docente investigador/a debe incluir

sus antecedentes y expertise académica en el ámbito de la investigación, formación de posgrado, formación de recursos humanos, entre otros importantes aspectos. Una vez completado en el sistema el formulario de postulación/presentación, se sucede una instancia de evaluación de pares que se realiza a posteriori, cuyo resultado da lugar a la adjudicación de una categoría acorde a sus antecedentes. La obtención de alguna de las categorías vigentes de 1 a 5, siendo la categoría 1 la más alta y 5 la de ingreso al sistema, permite a los académicos/as no solo pertenecer a un ranking de investigadores/as a nivel nacional sino también acceder a ciertos espacios, actividades y prestigio vinculados al ethos que rodea a dicha categoría. En ambos sistemas, que entre sí en la práctica son complementarios, la evaluación es por comisión de pares.

Esta nueva apertura de postulaciones y repostulaciones al PRINUAR se realizó a consecuencia de una seria revisión de la Resolución del Ministerio de Educación Nro. 472/2023, de marzo de 2023; que retoma el Decreto N° 2.427/93, que estableció la evaluación de la actividad científica y tecnológica en el Sistema Universitario Nacional mediante un Proceso de Categorización para docentes investigadores de Universidades de Gestión Pública, y de acuerdo a los lineamientos contenidos en el Manual de Procedimientos aprobado por Resolución Ministerial N° 1543/2014. Esta revisión fue motivada por la pandemia sin lugar a dudas y también por una sentida necesidad de incorporar temas de perspectiva de género, así como la integración paulatina de las actividades de investigación de arte como objeto de estudio.

En 2019, el Ministerio de Educación, Cultura, Ciencia y Tecnología de la Nación crea el Sistema Nacional de Docentes Investigadores Universitarios (SiDIUN) con el objetivo de jerarquizar la investigación científica y el desarrollo tecnológico en el ámbito del Sistema Universitario Argentino (Resol. N° 1216/2019). Este mecanismo de evaluación y categorización, como todo que se precie de tal, pretende conducir la investigación en las universidades hacia un desarrollo más armónico en relación con las distintas instituciones científicas nacionales que integran el Sistema Nacional de Ciencia, Tecnología e Innovación.

Durante el 2022 y 2023 se realizó esta nueva convocatoria que pretende atender el retraso en las categorías de docentes investigadores/as y que ordena y coordina el sistema científico universitario.

Por su parte, la carrera de investigación del CONICET tiene rasgos particulares que lo diferencian de otros organismos análogos de otros países de América Latina. Este sistema ofrece un puesto permanente de investigación con una remuneración de base. Los/as académicos/as pueden desempeñarse en una universidad o centro de investigación dependiente o no de CONICET, siendo este organismo el empleador y evaluador (Vasen et al, 2021).

El CONICET no ofrece incentivos monetarios para premiar la publicación en revistas con mayor factor de impacto, esto parecería favorecer las condiciones para promover cambios pero hace tiempo se puede verificar una fuerte creencia en el prestigio de las revistas del cuartil número 1 de SCOPUS que son invariablemente consideradas condición sine quanon para el ingreso o promoción de un investigador en ciertas áreas del conocimiento.

El PROINCE y el CONICET representan dos culturas evaluativas -que aunque distintas- revisten gran impacto nacional con profundo arraigo en la comunidad académica argentina.

La evaluación de la producción en el CONICET y su incidencia en los estilos de publicación

La carrera de investigación del CONICET (CIC) tiene rasgos singulares que lo diferencian de otros organismos análogos de otros países de América Latina. Por un lado, no es un sistema que ofrece un complemento al salario del investigador de parte de su universidad o institución, sino que ofrece un puesto permanente de investigación con una remuneración de base. En cuanto al proceso de evaluación, según el reglamento del organismo es el directorio quien determina los criterios generales que se dan a conocer en cada convocatoria para ingresos y promociones a la CIC (Res. 033/08, CONICET, 2008 -actualizado a 2021-). Posteriormente, cada comité asesor disciplinar revisa las pautas y los criterios ajustándolos a las "particularidades de las especialidades" (art. 31).

El análisis de las trayectorias completas de publicación muestra mayor diversidad de formatos, un porcentaje menor de publicaciones en inglés y mayor de publicaciones en Argentina (Beigel y Gallardo, 2021). Pero lo cierto es que hay una ponderación mayor de cualquier publicación en inglés respecto de otra realizada en español o en el país.

Hace algunos años, el CONICET realizó convocatorias a temas estratégicos (TE), sin embargo, Fischer, Goldberg y Jeppesen, (2022) concluyen que más de dos tercios de esos investigadores no han establecido vinculaciones con los actores de los sectores socio-productivos en los que supuestamente se insertan. También se detecta entre los investigadores en sus informes mayor preocupación por volcar los resultados de las investigaciones en producciones académicas. Esto supone la necesidad no solo de diseñar un acompañamiento en la gestión de la vinculación con los sectores estratégicos sino, además, de pensar un proceso evaluativo que pondere y valore las actividades de vinculación y transferencia en la carrera científica (Vasen, 2018).

Uno de los asuntos centrales a tener en cuenta en la revisión de la Resolución 2249/2014 (https://idh.unc.edu.ar/files/CCSH_-RD-20140625-2249.pdf) es el peso de la evaluación cuantitativa, a pesar de los esfuerzos realizados y de los documentos elaborados en el CONICET para fomentar un adecuado balance con el aspecto cualitativo. Por eso resulta relevante que una nueva Resolución contribuya a unificar criterios y objetivar la clasificación de las revistas pero que no se desentienda de una de las distorsiones principales que afecta las evaluaciones de nuestro organismo y en general de todo el sistema evaluativo a nivel internacional. La reintroducción de las 5 producciones relevantes como elemento importante en la evaluación de ingreso a CIC ha sido muy positiva. Pero aún falta adecuar los formularios y capacitar a los pares evaluadores para que comprendan la necesidad de leer los materiales y opinar sobre la originalidad de la contribución. Siguen completándolo en función de la indexación de las revistas.

Con la aplicación de la Res. 2249 se consolidó la autonomía de las ciencias sociales y humanas para jerarquizar las publicaciones de acuerdo a sus especificidades, y esto es aceptado por las otras áreas -a veces a regañadientes- en instancias superiores de evaluación, como la Junta de Calificación. Las recomendaciones de la Resolución 2249 implican otorgar el máximo puntaje a revistas con indexaciones en Scopus, WoS (Clarivate), Scielo pero, por usos y costumbres, se agregó Latindex Catálogo.

La heterogeneidad de las prácticas evaluativas en el PROINCE (hoy PRINUAR)

Existe una copiosa literatura que analiza la historia y la normativa que da vida al PROINCE. En su mayoría coinciden en puntualizar su impacto en la profesionalización académica, en primer lugar, un incremento en el número de investigadores y proyectos,

mayor en las universidades con menor tradición en la investigación y, dentro de cada universidad, mayor en las disciplinas con menor tradición en la investigación: las humanidades y las ciencias sociales, ciertas especialidades del campo tecnológico, y las áreas en las que tiene un fuerte peso el ejercicio profesional. Prati (2003) señala como uno de los efectos más relevantes del Programa un cambio en los criterios de status en el medio académico, a partir de la creciente importancia que adquirieron las actividades de investigación en la adquisición de prestigio social. La mayor parte de la literatura especializada coincide en afirmar que el Programa contribuyó a otorgar identidad institucional a la investigación en las universidades, a partir de la creación de una nueva figura, el “docente-investigador”, que realiza investigación en relación de dependencia respecto de la propia universidad,

Entre los estudiosos del campo universitario argentino es muy común separar las universidades nacionales del CONICET, tanto por sus modos de producción diferenciados como por sus culturas evaluativas ciertamente singulares. Sin embargo, en las universidades y en su sistema de evaluación nacional par excellence, el PROINCE, investigadores de CONICET y docentes-investigadores conviven: son objeto y sujetos de la categorización. Así, a partir de la creación del PROINCE convivieron conflictivamente dos culturas evaluativas que tienen carácter nacional, las que no obstante suelen solaparse, sobre todo en el ámbito de las ciencias naturales y aplicadas.

El PROINCE categoriza nacionalmente a investigadores/as en base a un conjunto de criterios entre los cuales las publicaciones son un porcentaje menor y solo se exige que las mismas sean “preferentemente indexadas” para la categoría superior. Es un sistema de evaluación que admite perfiles muy heterogéneos. Uno de los aspectos principales que singularizan el PROINCE en relación a las tendencias evaluativas globalizadas es el hecho de que nunca se incorporó un criterio de evaluación de la producción basada en el impacto o el índice de citas, cosa que sí hicieron en otros países y que es moneda corriente en el CONICET (Beigel y Baranger, 2019)

La evaluación en PROINCE tiende a ser permeable a la diversidad de estilos de producción y circulación: más allá de las discusiones, el físico que publica solo en revistas mainstream y la cientista social que publica libros solo en español podrán terminar alcanzando la categoría solicitada si tienen producción suficiente. Y esto tiene ventajas y desventajas. Entre las primeras, se cuenta el favorecer tanto las agendas locales como las “globales”, admitir la publi-diversidad y la especificidad geo-espacial

de las trayectorias. Entre las segundas, la sensación que queda en evaluadores y evaluados de que prima una “valuación” burocrática o cuantitativa, o que la calidad no hace la diferencia, o que es un sistema poco exigente.

En relación a las prácticas evaluativas concretas desarrolladas en el PROINCE, el estudio mostró que no existe en los comités de pares un criterio uniforme sobre lo que implica la calidad de la investigación ni mucho menos una forma única de ponderar el desempeño de sus pares docentes. Pudo observarse que la evaluación en este programa asume características *sui generis* que no se ajustan estrictamente a lo que la literatura disponible define como evaluación por pares (peer review). Si bien, algunos elementos están presentes, en las evaluaciones observadas aparecen varias diferencias: los pares que evalúan no son expertos en la temática o línea disciplinar, sino que pertenecen a áreas disciplinarias.

Modelos de evaluación con referencia a herramientas usadas a nivel internacional

La evaluación de una trayectoria académica, en referencia a las actividades de investigación, implica la ponderación de ciertos aspectos de la misma en función de criterios que se acuerdan y establecen como prioritarios en la comunidad académica (representada por quienes offician de evaluadores). Tras la implementación del PRINUAR, ya presentado y analizado en este documento, existe cierto consenso acerca de la necesidad de revisar esos criterios, con el objetivo de realizar ajustes que permitan trasladar la centralidad actual en la producción de papers (con criterios referidos al impacto social) hacia otros aspectos más relevantes para nuestro contexto (relevancia social). Esos criterios responden a objetivos sociales, económicos, políticos y académicos globales que van más allá de las instituciones universitarias y sus investigadores/as, aunque los afecta de manera directa (Rafols, I entrevista 2023)

La clave estaría entonces en pensar cuáles son los intersticios, los espacios no cubiertos por estas evaluaciones que universalizan criterios y allí pensar el surgimiento de una evaluación más situada y significativa para el contexto y su comunidad. Desde nuestro posicionamiento ello tiene relación directa con imaginar respuestas a las preguntas: para quién se investiga, para qué se investiga y con quién se investiga; con un sentido claramente orientado a provocar contribuciones sociales en territorio de significativa relevancia.

La Declaración de Leiden y la Declaración sobre Evaluación de la Investigación

En un informe de síntesis publicado por un grupo del Center for Science Futures, junto con socios del Global Young Academy e InterAcademy Partnership², se destaca que existen al menos 15 esfuerzos distintos en el mundo que instan a los/as formuladores/as de políticas, financiadores/as y/o directores/as de instituciones de Educación Superior a minimizar el daño potencial de los sistemas de evaluación actuales. Todas estas iniciativas han llegado a una amplia audiencia y son progresistas en su enfoque en métricas responsables como requisito previo para mejorar la cultura de investigación y traer igualdad, diversidad, inclusión y pertenencia a la comunidad de investigación.

Dos Declaraciones, apoyadas por miles de investigadores de primera línea de todo el mundo, resultan sumamente importantes. La primera es La Declaración sobre Evaluación de la Investigación (DORA) rubricada en San Francisco en el año 2012. La segunda es El Manifiesto de Leiden sobre indicadores de investigación firmado en el 2014.

El DORA es un espacio catalizador del cambio en tanto pretende aportar a la reforma de los modos actuales de evaluación de la investigación en los/as académicos/as en el mundo. Desde la amplitud de disciplinas, busca contribuir en el desarrollo y promoción de mejores prácticas en la evaluación de la investigación académica, con vocación de construir diseños de investigación que aborden directamente las desigualdades estructurales en el mundo académico. Mediante sus aportes pretende ofrecer la reflexión sobre la necesidad de desarrollar nuevas políticas y prácticas para las decisiones de contratación, promoción, financiación y la prueba de nuevas herramientas y procesos en la evaluación de la investigación y el uso responsable de métricas que se articulen con los valores académicos fundamentales y promuevan la coherencia y la transparencia en la toma de decisiones. Durante la última década, ha habido una serie de manifiestos y principios de alto perfil sobre evaluación de la investigación para abordar estos desafíos, incluido el Manifiesto de Leiden (2014).

² El futuro de la evaluación de la investigación: una síntesis de los debates y desarrollos actuales. Disponible en: <https://council.science/es/publications/the-future-of-research-evaluation-a-synthesis-of-current-debates-and-developments/> Fecha de visita 30/01/24

En síntesis, los indicadores cuantitativos ofrecen un buen apoyo para la evaluación, pero las métricas no pueden sustituir las decisiones informadas. La valoración de la calidad de una investigación o de un artículo debe basarse en la investigación misma y por métricas a su nivel y no solamente, por ejemplo, por el factor de impacto de la revista donde se haya publicado. No hacerlo así significa que los/as evaluadores/as abandonan su responsabilidad.

A partir de lo anterior, algunas conclusiones muy concretas que se pueden obtener a partir de un análisis sobre DORA y Leiden (Codina, Ll. 2019):

- Las métricas, mal entendidas, pueden acabar utilizándose de forma amenazante para los verdaderos objetivos de la ciencia. En lugar de apoyar su progreso, pueden ser un verdadero obstáculo.
- Las dos declaraciones examinadas resultan necesarias y valiosas y hay que promover su conocimiento. Ambas tienen el respaldo de colectivos científicos que han demostrado de forma sobrada su excelencia en la ciencia, y están apoyados por sociedades científicas de proyección internacional.
- Ambas descartan formas taxativas de evaluación basadas en indicadores aislados que sustituyan decisiones informadas, entre otras cosas, porque supone que los evaluadores hacen abandono de sus responsabilidades.
- En su lugar, recomiendan aproximaciones basadas en el portafolio completo del investigador. Si este portafolio está acompañado, además de buenas métricas, mucho mejor. Pero nunca en sustitución de una decisión informada e integral.
- Las declaraciones son especialmente beligerantes con el uso del factor de impacto, que es una medida de la revista, no del artículo, cuando se utiliza como sustituto de la evaluación de la calidad de artículos e investigaciones individuales.
- Por último, muchos expertos, además de desaconsejar su uso en evaluaciones de carreras académicas y de la calidad de aportaciones científicas, alertan sobre los importantes problemas intrínsecos que presenta el Journal Impact Factor, en sí mismo, como estadístico, independientemente de su aplicación.

Es necesario destacar que ninguna de las dos Declaraciones pone en duda la necesidad de disponer de indicadores y métricas, dado que estos instrumentos brindan

transparencia y ayudan a reconocer los esfuerzos de académicos e investigadores. Pero a su vez, los mismos son insuficientes para valorar una actividad tan compleja, razón por la que conviene que una evaluación tenga un componente de autoevaluación, de evaluación externa e independiente es imprescindible sobre todo en aquellas universidades donde la actividad investigativa se financia con fondos públicos.

DORA SPACE

DORA es quizás la iniciativa global más activa. Ha catalogado los problemas causados por el uso de indicadores basados en revistas para evaluar el desempeño de investigadores/as individuales y proporciona 18 recomendaciones para mejorar dicha evaluación; desaconsejando categóricamente el uso de métricas basadas en revistas para evaluar la contribución de un/a investigador/a o cuando se busca contratar, promover o financiar.

SPACE es una herramienta para la evaluación que se puede adaptar a diferentes contextos institucionales, geografías y etapas de preparación para la reforma, lo que permite a las universidades hacer un balance de las limitaciones y capacidades internas que probablemente afecten su capacidad para reformar la forma en que evalúan la investigación y los investigadores.

Marco SCOPE

En 2021, INORMS REG publicó una versión actualizada del marco SCOPE que incluyó un proceso de cinco pasos que las instituciones deben seguir durante las evaluaciones:

- comenzar por explicitar lo que se desea evaluar,
- considerar el contexto,
- explorar opciones para medir,
- sondear en profundidad,
- valorar la evaluación.

CoARA

CoARA es un Acuerdo sobre la Reforma de la Evaluación de la Investigación generada a partir de una Coalición global, de organizaciones e instituciones que participan de la

investigación, para participar activamente en los procesos de toma de decisiones, publicada en 2022, tras una acción conjunta entre la Asociación Universitaria Europea (EUA), Science Europe y la Comisión Europea.

CoARA presentó 4 compromisos centrales:

- 1) reconocer diversas contribuciones durante las evaluaciones,
- 2) utilizar evaluaciones cualitativas de pares y métricas basadas en revisiones sobre indicadores cuantitativos para la evaluación de la investigación,
- 3) evitar el uso de métricas basadas en revistas y publicaciones,
- 4) abolir el uso de clasificaciones universitarias durante las evaluaciones.

TARA

El proyecto TARA surge en 2021 como una iniciativa de DORA con el objeto de generar herramientas para avanzar en la evaluación de la investigación en el mundo académico, apoyando la defensa comunitaria de políticas y prácticas de evaluación de investigación responsable. En 2022, se presentaron las dos primeras herramientas en una convocatoria comunitaria que incluyó investigadores, editores y profesores de todo el mundo con la finalidad de obtener comentarios de la comunidad sobre el desarrollo futuro de herramientas para 2023. Esta herramienta incluye ideas para defender la transparencia, la toma de decisiones basada en políticas, fomentar una diversidad de opiniones que inviten a todos los puntos de vista y ampliar las posibilidades más allá de las normas históricas.

FOLEC

Latinoamérica cuenta con el Foro Latinoamericano de Evaluación de la Investigación, conocido como FOLEC (<https://www.clacso.org/folec/>), que es una iniciativa del Consejo Latinoamericano de Ciencias Sociales (CLACSO). El objetivo de FOLEC es apoyar el intercambio de conocimientos sobre las reformas de evaluación de la

investigación en instituciones de Nivel Superior de América Latina. FOLEC ha convocado reuniones nacionales e internacionales para facilitar la transferencia de conocimiento orientado a la reforma de la evaluación científica. La producción de documentación encaminada a pensar la reforma de la evaluación de la investigación en la región ha sido prominente y de gran relevancia. En 2020 FOLEC-CLACSO desarrolló una etapa de diagnósticos, propuestas y orientaciones en relación con los procesos y sentidos de reforma de la evaluación académica en la región, plasmados en distintos documentos de trabajo, encuentros y actividades. Desde 2022, integra la Junta Ejecutiva de la Declaración de San Francisco sobre Evaluación de la Investigación (DORA).

FOLEC Se ha manifestado con múltiples formas de resistencia a la mercantilización de la ciencia a nivel internacional, realizando fuertes críticas a las formas dominantes de evaluación de la producción científica e instando a los académicos de la región a producir un giro hacia una Ciencia Social Relevante, de carácter abierto; Rescatando el importante lugar de las Ciencias Sociales y las Humanidades no solo por sus diferencias sino también por sus aportes. En numerosos documentos y eventos académicos ha tomado la iniciativa que orienta el debate hacia la necesidad de evaluar la evaluación de la producción científica en América Latina y el Caribe, poniendo sobre el tapete la necesidad de una transformación que dé cuenta de las características, particularidades, identidades e historia propias de la región, en franca interacción con el mundo (CLACSO, 2020).

Conclusiones

El sistema de investigación está bajo gran presión en todos los países del mundo debido a las crecientes expectativas de múltiples actores (incluidos los financiadores, los gobiernos y la industria editorial), las tensiones entre las dinámicas de competencia y cooperación, los recursos limitados, un sistema de comunicación académica en evolución y una industria editorial y de análisis de datos cada vez más agresiva (De Rijcke et al. 2023). La política científica debe manejar estas demandas y tensiones, mientras mantiene la calidad académica, defiende la integridad de la investigación, es inclusiva y diversa, y salvaguarda tanto la investigación básica como la aplicada. Durante la última década, estas tendencias han ido acompañadas de reflexiones críticas sobre los sistemas de evaluación de la investigación y medición del desempeño que

aparecen en el centro de las presiones y, sobre todo, en el núcleo de las posibilidades de una transformación. Las metodologías apropiadas y sensibles al contexto para evaluar la investigación son importantes, pero se han intensificado los debates sobre la calidad de la evidencia que informa la formulación de políticas, las prioridades en investigación y las decisiones sobre el financiamiento, que inciden directamente en las trayectorias profesionales individuales y el crecimiento de las instituciones.

Lo hasta aquí reseñado, independientemente de enfoques y matices respecto de la evaluación, pone en juego la concepción de excelencia. Aquí conviene detenerse, especialmente por su interés para esta ponencia, en el papel de las agencias financiadoras en la institucionalización de los discursos de excelencia a través del desarrollo y ejecución de mecanismos competitivos y regímenes de evaluación (Aksnes, Langfeldt y Wouters, 2019). Como sitios de intervención en los que se unen la política científica, el financiamiento y la promoción de la investigación, estas organizaciones pueden producir cambios importantes en el resto de las instituciones del sistema. La revisión de literatura ofrecida por Jong, Franssen, y Pinfield, (2021) devela que éstas han sido poco exploradas, mientras juegan un papel importante en la proliferación de los discursos de excelencia existentes.

Dar por sentado que hay una definición global de calidad, universalmente aceptada, implica asumir como propios criterios descontextualizados que tienen una larga historia y que están siendo profundamente cuestionados en los últimos años. Aun con las mejores intenciones, estas ambigüedades conceptuales producen desigualdades y arbitrariedades que se reproducen cotidianamente en los comités de pares. Pero, además, conspiran contra las orientaciones generales que una agencia de financiamiento de la investigación, legítimamente, pretende imprimir a sus decisiones.

Sin lugar a dudas el Manifiesto de Leiden y el Espacio DORA han significado un gran avance hacia la reflexión y la práctica concreta de la evaluación de la investigación en el mundo académico. Los esfuerzos realizados en la última década se orientan fundamentalmente a la construcción y prueba de herramientas evaluativas que escapen a los modelos hegemónicos actuales (DORA SPACE, CoARA, TARA).

En América Latina se están desarrollando experiencias de sistemas regionales de información para incluir la producción científica publicada en la región, y promover su valorización en los sistemas de evaluación. Un ejemplo de ello son los esfuerzos del

Foro Latinoamericano sobre Evaluación Científica (FOLEC) perteneciente al Consejo Latinoamericano de Ciencias Sociales (CLACSO) y las herramientas que ofrece para comprender por ejemplo las características actuales de las revistas nacionales; orientando esta discusión hacia la idea de una ciencia con relevancia social y una ciencia abierta.

El FOLEC suscribe y participa de DORA, proponiéndose adaptar los principios de este Espacio a las particularidades de la Región Latinoamérica. Estas importantes iniciativas aportan una nueva visión de la evaluación sobre la investigación orientada fundamentalmente a reconocer las características del contexto, las misiones y necesidades institucionales, la diversidad de paradigmas, historia y cultura de las disciplinas y la compleja trayectoria de los/as académicos/as que se traduce en una multiplicidad de prácticas relacionadas a la investigación; las que no pueden ser medidas con instrumentos únicos. El gran desafío a asumir en la actualidad debe tomar los principios que se destacan y defienden a nivel global pero adaptándolos a las necesidades locales y regionales. Ello no implica en absoluto abandonar la idea de medición sino por el contrario pensar en otras formas que la complementen.

Llevando estas ideas y principios a la evaluación que se realiza actualmente en Argentina, resulta importante promover la reflexión en torno a la necesidad de pensar una nueva cultura de evaluación que incentive a los pares evaluadores a opinar sobre la calidad de las contribuciones académicas y sobre la relevancia social de los resultados obtenidos en un sentido mucho más amplio. Sin lugar a dudas, esto implicaría una reforma o revisión del sistema de evaluación en su conjunto.

Bibliografía

-Aksnes D. W. Langfeldt L., y Wouters P. (2019) Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories SAGE Open January-March 2019: 1 –17

-Araujo S. (2003) Universidad, investigación e incentivos. La cara oscura. La Plata. Ediciones Al Margen

-Arocena, R., y Sutz, J. (2020). “The need for new theoretical conceptualizations on National Systems of Innovation, based on the experience of Latin America”, Economics of Innovation and New Technology, 29, (7), pp. 814-829.

-Avanço, K. et al. 2021. Future of Scholarly Communication . Forging an inclusive and innovative research infrastructure for scholarly communication in Social Sciences and Humanities. Research Report. OPERAS. (2021, June 29).

-Aguado López E. y Becerril García A. (2021). Performatividad en la ciencia mexicana: el dispositivo de evaluación del SNI. Revista mexicana de ciencias políticas y sociales. Vol. LXVI N° 243. 19-53.

-Beigel, F. et al (2023) “Scientific production and gender inequalities in two academic elites: Brazil and Argentina”, Revue d'histoire des sciences humaines, N° 42, p.255-280. Indexada en DOAJ, ERihPlus. <https://doi.org/10.4000/rhsh.8374>

-Beigel F. y Baranger D. (2019). Culturas evaluativas: Impactos y dilemas del Programa de Incentivos a Docentes-Investigadores en Argentina (1993-2018), 2019, pp. 271-284 (14 pages) <https://doi.org/10.2307/j.ctvt6rksm.14>

-Beigel, F., y Bekerman, F. (2019). Culturas evaluativas Impactos y dilemas del Programa de Incentivos a Docentes-Investigadores en Argentina (1993-2018). CLACSO; Buenos Aires: CECIC, 2019; Buenos Aires: IEC-CONADU, 2019.

-Beigel, F., y Gallardo, O. (2021). Productividad, bibliodiversidad y bilingüismo en un corpus completo de producciones científicas. Revista Iberoamericana de Ciencia, Tecnología y Sociedad-CTS, 16(46), 41-71.

-Biagioli, M., y Lippman, A. (Eds.). (2020). Gaming the metrics: Misconduct and manipulation in academic research. Mit Press.

-CLACSO – Consejo Latinoamericano de Ciencias Sociales (2020). Diagnóstico y propuestas para una iniciativa regional. Serie Para Una Transformación de la Evaluación de la Ciencia en América Latina y el Caribe del Foro Latinoamericano sobre Evaluación Científica (FOLEC). 2da. Edición. Buenos Aires: CLACSO. www.clacso.org/folec/clacso-ante-la-evaluacion

-CoARA: una dirección común para reformar la evaluación de la investigación. Status quo de las evaluaciones de la investigación Disponible en: <https://sfdora.org/2023/07/18/empowering-fair-evaluation-and-collective-impact-inorms-and-coaras-efforts-to-drive-change/> 18 DE JULIO DE 2023

-Corda, M. C., y Ruiz del Ferrier, M. C. (2018). Constitución y desarrollo de LatinREV, una red latinoamericana de revistas académicas en el campo de las ciencias sociales y las humanidades. Memoria de la IV Jornada de LatinREV

-De Rijcke et al. (2023) The future of research evaluation: a synthesis of current debates and developments discussion . The Gloperbal Young Academy (GYA), InterAcademy Partnership (IAP) and the International Science Council (ISC)

-DORA. Un marco de evaluación de la calidad para el diseño, la planificación y la evaluación de la investigación: actualizaciones del Programa de eficacia de la investigación sobre sostenibilidad en Canadá. disponible en:

<https://sfdora.org/2023/09/06/a-quality-assessment-framework-for-research-design-planning-and-evaluation-updates-from-the-sustainability-research-effectiveness-program-in-canada/>

-Echeverry-Mejía, J. A. (2021). Esquema analítico de la gobernanza de la ciencia, la tecnología y la innovación en el ámbito universitario. En L. Córdoba, L. Rovelli, y P. Vommaro (Eds.), Política, gestión y evaluación de la investigación y la vinculación en América Latina y el Caribe (pp. 223-260). CLACSO. <https://doi.org/10.5281/zenodo.6359684>

-Espacio DORA Repensar la evaluación de la investigación: pilares del impacto. Disponible en: <https://sfdora.org/resource/rethinking-research-assessment-building-blocks-for-impact/>

-Faber, S. (2021). “Las trampas de la excelencia universitaria. La obsesión mundial con la “calidad” académica, ¿está destruyendo la calidad intelectual, humana y laboral de las universidades?” CTXT Contexto y Acción. N° 273. <https://ctxt.es/es/20210601/Politica/36260/universidad-excelencia-trampas-calidad-academica-Sebastian-Faber.htm>

-Fischer, M., Goldberg, M. y Jeppesen, C. (2023). Trayectorias en investigación orientada a la resolución de problemas. El caso de los temas estratégicos del CONICET, Argentina. Revista Iberoamericana de Ciencia, Tecnología y Sociedad.

-Foro Latinoamericano sobre Evaluación Científica Ciudad de México (2019). Relatoría El 26 y 27 de noviembre de 2019, en el Instituto Dr. José María Luis Mora de la Ciudad de México (FOLEC).

-Giménez-Toledo, E. (2015). La evaluación de la producción científica: breve análisis crítico. *RELIEVE-Revista Electrónica de Investigación y Evaluación Educativa*, 21(1).

-Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. Mit Press.

-Larivière, Vincent e Yves Gingras. (2010). On the relationship between interdisciplinarity and scientific impact. En *Journal of the American Society for Information Science and Technology*. Vol. 61, no. 1, 126-131.
< <https://doi.org/10.1002/asi.21226> >

-Invernizzi, N. (2022). Los sistemas de evaluación como conformadores de agendas científicas. *Ciencia, tecnología Y política*, 5(9), 080.
<https://doi.org/10.24215/26183188e080>

-Jong, L., Franssen, T. and Pinfield, S. (2021) *'Excellence' in the research ecosystem : a literature review*. Working Paper. RoRI Working Papers, 5 (5). Research on Research Institute

-Molas-Gallart, J.; Boni Aristizábal, A.; Giachi, S.; Schot, J. (2021). A formative approach to the evaluation of Transformative Innovation Policies. *Research Evaluation*. 30(4):431-442. <https://doi.org/10.1093/reseval/rvab016>

-Moore, F., Rovelli, L. I., y Corengia, A. (2021). Tendencias recientes en la Carrera Docente/Académica en el sector privado argentino. Un estudio de caso en dos universidades de investigación. *RAES*, 13(23), 238-253.

Naidorf, J. y Perrotta, D. (2016). “La cultura académica argentina frente al cambio de ciclo”, *Revista del IICE* /39: 45-66

-Piovani, J. I. (2019). Sobre la utilidad de las ciencias sociales en tiempos de neoliberalismo y posverdad. *La política científica en disputa: diagnósticos y propuestas frente a su reorientación regresiva*, 115-133.

-Piro, F.N., Sivertsen, G. (2016). How can differences in international university rankings be explained?. *Scientometrics* 109, 2263–2278 (2016).
<https://doi.org/10.1007/s11192-016-2056-5>

- Pölönen, J., y Sivertsen, G. (2021). Grey-zone between legitimate and predatory scholarly publishing. The 25th International Conference on Science, Technology and Innovation Indicators (STI). <https://doi.org/10.5281/zenodo.5513289>
- Pontille y Thorny in Eve, M. y Grey, J. (2020). “Reassembling scholarly communications.”
- Prati, M. (2003). El impacto del Programa Incentivos a partir de las percepciones de los académicos. Trabajo final Licenciatura en Sociología: UNLP.
- Ràfols, I. (2019) SyT Indicators in the wild. *Research Evaluation*, 28(1), 2019, 7–22, doi: 10.1093/reseval/rvy030
- Rafols, I. (2023). Entrevista. Lugar: Universidad Nacional de Cuyo, Mendoza. Fecha: 30 y 31 de octubre-1 de noviembre 2023. Marco: Workshop "Fuentes nacionales, regionales y globales para una nueva perspectiva de circulación del conocimiento". Entrevistadora: Fernanda Beigel.
- Resolución Ministerial N° 1216/2019 -SiDIUN
- Resol 472/2023. Anexo. REGLAMENTO PROGRAMA PARA LA INVESTIGACIÓN UNIVERSITARIA ARGENTINA (PRINUAR)
- Robinson Garcia, N., Costas, R., van Leeuwen, T. N., y Nane, T. (2019). Towards a multidimensional valuation model of scientists. In G. Catalano, C. Daraio, M. Gregori, H. F. Moed, & G. Ruocco (Eds.), 17th International Conference on Scientometrics and Informetrics, ISSI 2019 - Proceedings (Vol. 2, pp. 2565-2566). (17th International Conference on Scientometrics and Informetrics, ISSI 2019 - Proceedings; Vol. 2). International Society for Scientometrics and Informetrics.
- SCOPE: Estructura Un proceso de cinco etapas para evaluar e investigar responsablemente. Disponible en: <https://inorms.net/wp-content/uploads/2022/03/21655-scope-guide-v10.pdf>.
- Schimanski, L. A., y Alperin, J. P. (2018). The evaluation of scholarship in academic promotion and tenure processes: Past, present, and future. *F1000Research*, 7.

- TARA. Introducción al kit de herramientas del Proyecto. Dos nuevas herramientas para apoyar la evaluación responsable de la investigación: composición del comité sin sesgos y elementos básicos para el impacto. 2 DE MARZO DE 2023. Disponible en: <https://sfdora.org/2023/03/02/introducing-two-new-tools-to-support-responsible-research-assessment-debiasing-committee-composition-and-building-blocks-for-impact/>
- Vasen F. (2018). La ‘torre de marfil’ como apuesta segura: Políticas científicas y evaluación académica en México (Arizona State University, 2018-07)
- Vasen, F.; Sarthou, N.; Romano, S.; Gutiérrez, B.; Ortiz, M. E. & Pintos, M. (2021). “Sistemas Nacionales de Categorización de Investigadores en Iberoamérica: la configuración de un modelo regional”, Documento de trabajo N° 1, Proyecto PICT2018-2794, Julio 2021
- Vessuri, H., Guédon, J.-C., y Cetto, A. M. (2014). Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology*, 62(5), 647-665. <https://doi.org/10.1177/0011392113512839>