

FACES

Facultad de Ciencias Económicas y Sociales

Año 11

Nº 22

enero-abril 2005

Facultad de Ciencias Económicas y Sociales
Universidad Nacional de Mar del Plata

CENTRO DE DOCUMENTACIÓN
Instituto de Investigaciones
Facultad de Ciencias Económicas y Sociales
Universidad Nacional de Mar del Plata
cendocu@mdp.edu.ar
<http://eco.mdp.edu.ar/cendocu/>

Evaluación en Ciencia y Tecnología en universidades latinoamericanas

Science and Technology Evaluation in Latinamerican Universities

Marta Arana¹

RESUMEN / SUMMARY

La evaluación científico - académica constituye actualmente un eje de debate en el campo de la educación superior, en el nivel internacional. Esta tendencia se asocia al surgimiento de una nueva ética en el campo de la educación superior, vinculada al desarrollo de la calidad, la eficiencia y el desempeño. Es posible observar una redefinición de las relaciones entre estado y gobierno, por un lado, y los sistemas de educación superior, por el otro: desde el sector público surge el estado evaluativo y en el sistema de educación superior aparece una ética competitiva como principal motor de desarrollo.

At present scientific-academic evaluation constitutes a point of debate in the field of higher education at international level. This trend is associated to the rise of a new ethics in the field of higher education linked to quality, efficiency and performance development. It is possible to observe a redefinition of the relationship between state and government on the one hand and the systems of higher education on the other. From the public sector the assessing state arises and from the higher education system a competitive ethics as a development tool is implemented.

¹ Investigadora y Profesora Titular de la Facultad de Ciencias Económicas y Sociales de la Universidad Nacional de Mar del Plata. aranamarta@gmail.com.ar

PALABRAS CLAVE / KEYWORDS

Educación superior - evaluación científico académica.

Higher education - scientific academic evaluation

ACERCA DE LA EVALUACIÓN ACADÉMICA: MECANISMOS Y PROCEDIMIENTOS

Aproximación desde una perspectiva internacional

El eje central de las Políticas Públicas implementadas en Educación Superior, en la década de los noventa, trata sobre la evaluación de todo el sistema de educación superior. Entre los Programas implementados en el marco del PRES (Programa de Reforma de la Educación Superior) el correspondiente a la implementación de Incentivos (Decreto 2420/93) a docentes investigadores, no escapa a esta realidad. Según señala Albornoz (2003: 14):

El problema de la asignación de valor a los conocimientos científicos ha sido una cuestión siempre presente, a partir de la primera institucionalización de las comunidades científicas en el siglo XVII. Muy tempranamente, la *Royal Society* debió desarrollar metodologías y criterios que le permitieran seleccionar los trabajos científicos que serían publicados, y establecer un orden de prelación entre ellos. Los primeros sociólogos de la ciencia, como Robert Merton, Joseph Ben David y Derek de Solla Price describieron los mecanismos por los cuales la comunidad científica se autorregularía mediante la asignación y reconocimiento de valor a la actividad de sus miembros y a los resultados de su trabajo. Esta dimensión del proceso de evaluación remite exclusivamente a la calidad y relevancia teórica de los resultados de la investigación. Sin embargo, a partir de la atención prestada por los gobiernos a la ciencia (proceso que en gran escala se desarrolló a partir de la segunda guerra mundial) otras finalidades sociales, propias de la lógica de los intereses públicos, han tornado más complejo el proceso de evaluación, introduciendo en él nuevos criterios y actores. Las decisiones relativas a las líneas de trabajo que habrán de ser financiadas y desarrolladas, así como los campos del

saber que serán explorados tienen ahora que ver con otras dimensiones, tales como la pertinencia o relevancia social, o la correspondencia con un sistema de prioridades políticas. En el campo de la tecnología, obviamente, los criterios de utilidad y eficiencia ocupan un lugar central en el proceso evaluador. Estos rasgos han convertido a la evaluación en un núcleo central de las políticas y en la gestión de la ciencia y tecnología.

La evaluación científico - académica constituye actualmente un eje de debate en el campo de la educación superior, en el nivel internacional. Guy Neave (1998) señala que esta tendencia se asocia al surgimiento de una nueva ética en el campo de la educación superior, vinculada al desarrollo de la calidad, la eficiencia y el desempeño. Sustrayéndonos de cualquier valoración previa, es posible observar una redefinición de las relaciones entre estado y gobierno, por un lado, y los sistemas de educación superior, por el otro: desde el sector público surge el estado evaluativo y en el sistema de educación superior aparece una ética competitiva como principal motor de desarrollo. En el marco descripto, las evaluaciones se ejercen *a posteriori*, ocupándose más de los resultados que de los procesos, y no tendrían como finalidad guiar el sistema en función de las prioridades nacionales. Neave (1998, 2001) realiza su trabajo basándose en la experiencia europeo - occidental, tomando los casos de Francia, los Países Bajos, Suecia y Gran Bretaña.

Joaquín Brunner (1990) propone una mirada diferente de la que denomina el "discurso latinoamericano" acerca de la evaluación científico - académica. Según dicho investigador, la aparición de este rol del estado como evaluador representaría una completa redistribución y racionalización de funciones entre el centro y las instituciones, al permitir que el centro mantenga el control estratégico del sistema con precisos instrumentos de políticas académicas, en tanto las instituciones aumentarían su autonomía, quedando libres para actuar frente a las presiones y demandas del mercado.

En el caso de la Argentina, en el marco del Programa de Reforma de la Educación Superior (PRES), se ha fortalecido la tendencia de contar con organismos intermedios y especializados entre el gobierno y las instituciones, que realicen las principales tareas de evaluación. Las evaluaciones en Francia

dependen del *Comité National d'Evaluation* que reporta directamente al Presidente de la República; en España, el Consejo de Universidades realiza la tarea de evaluación curricular, mientras que en la Argentina, la Comisión Nacional de Evaluación Académica Universitaria (CONEAU), dependiente del Ministerio de Educación, es responsable de esta tarea.

La organización de estos procesos de evaluación en el nivel nacional ha exigido implementar las condiciones procedimentales de la evaluación, que se sustentan en combinar mecanismos basados en el juicio de pares académicos conjuntamente con otros, desarrollados por los encargados de los organismos intermedios con funciones de evaluación. Asimismo, cada situación nacional ha condicionado estos procedimientos de evaluación externa con procedimientos internos de auto evaluación, obligándose en cada caso a definir los indicadores cuantitativos y cualitativos a utilizar.

El surgimiento del Estado Evaluativo implica el reconocimiento explícito de que la Educación Superior ya no puede seguir siendo administrada de manera independiente desde un único centro, sino que debe ser ajustada a los diferentes contextos nacionales y con mecanismos de coordinación que la vinculen, por ejemplo, con el mercado. Esto conlleva una redistribución de las relaciones de poder entre la autoridad del estado, las elites académicas y sus instituciones, y los correspondientes mercados y sus agentes. Este triángulo de coordinación (estado-mercado-elites académicas) no funciona de manera equilibrada: en los últimos tiempos se observa tanto una mayor presencia del mercado, como una forma más eficiente de intervención del Estado, así como un rol destacado de las elites académicas a las que pertenecen los actores que habitualmente ejercen las funciones de evaluación, mediante los procedimientos de la revisión por pares (*peer review process*).

Nuestra realidad latinoamericana no ha estado ajena a lo que Neave (1998) describe para Europa Occidental, pero son incipientes las cristalizaciones de prácticas relativas a la evaluación académica, así como de un discurso que promueva una nueva ética social. Entre las diferencias que presentan los sistemas de educación latinoamericanos y europeos es de destacar la masividad de los primeros, las severas restricciones presupuestarias que se constatan, así como la aparición de un nuevo y dinámico sector privado en educación superior que brega por ocupar un

espacio tan protagónico como el del sector público. En este nuevo escenario latinoamericano, se ha instalado un constante debate acerca de la evaluación de la producción científico académica, hacia adentro y hacia afuera de las instituciones en un contexto de crisis, que prácticamente es estructural.

Los principales actores que perciben la situación de crisis en la educación superior, que involucra el sistema de evaluación, y tratan de adoptar una actitud movilizadora, en aras de mejorar la calidad educativa son:

- a. los profesionales graduados, que constituyen la base del modelo clásico de universidad formadora de las "grandes profesiones" (médicos, abogados, ingenieros), quienes observan el deterioro generalizado de la enseñanza universitaria debido a la masificación, el bajo nivel de escolarización de los ingresantes y el descrédito de las prácticas docentes;
- b. los estudiantes y sus familias, al considerar que la educación superior ya no proporciona a los jóvenes las garantías de empleo ni el prestigio que anteriormente les brindaba;
- c. los científicos y académicos, que destacan la falta de condiciones laborales, de estímulos adecuados y la baja tasa de inversión por parte del Estado;
- d. los agentes externos, provenientes del mercado, que manifiestan sus quejas a la hora de contratar mano de obra y servicios de consultorías (las empresas, por ejemplo, consideran poco calificada y de bajo nivel la prestación de servicios a los usuarios);
- e. los gobiernos, que recién a partir de los años ochenta han comenzado a preocuparse por la eficiencia del "gasto" realizado en educación superior, sobre todo por la proliferación de instituciones privadas libres y la ausencia de marcos reguladores.

El panorama expuesto fundamenta por qué no puede hablarse de Estado Evaluativo en América Latina, como sucede en Europa Occidental. Pero es necesario destacar el surgimiento de una clara conciencia de la necesidad de la evaluación académica frente a la crisis de la educación superior. Dentro del sistema universitario latinoamericano, se observa una tensión entre quienes están a favor de la introducción de las prácticas de evaluación y quienes se resisten a ellas, generándose corporaciones a favor y en contra de las

evaluaciones. En el primer grupo se identifican, entre otros, núcleos de académicos activos; un número restringido de autoridades institucionales; circuitos tecno - burocráticos del gobierno que se autodefinen como "modernizadores"; algunos "especialistas del campo", que son estudiosos de los sistemas de educación; y consultores internacionales vinculados a la temática de la evaluación. La resistencia, por el contrario, está compuesta por un sector de actores e instituciones que reclaman una clara definición de los conceptos "calidad", "eficiencia" en la producción y que la evaluación se realice de manera contextualizada por las distintas realidades a considerar, tanto disciplinares como regionales. Se reclaman diferentes criterios que incluyan actividades realizadas en diversos campos, como la transferencia de tecnologías blandas y duras.

La Tabla 1, propuesta por Mario Albornoz (2003: 18) tomada del *Evaluating Federal Research Programs* (National Academy Press, Washington, 1999) se refiere a los métodos utilizados para evaluar I+D.

Dentro del sistema universitario, en el área de investigación se ha avanzado, sin duda, mucho más en la evaluación, que en otras áreas como gestión e incluso docencia. Existe una mayor tradición en la práctica de la evaluación por pares, que se utiliza tanto para la distribución de recursos como para determinar la competencia de proyectos e investigadores. Y también se percibe, en esta área, la necesidad de desarrollar sistemas nacionales de investigadores, que cuenta incluso con evaluaciones externas que contribuyen al desarrollo de una masa crítica.

¿Qué y cómo se pretende evaluar? Procedimientos e instrumentos

La evaluación pretende garantizar determinadas normas y valores académicos, con el propósito de asegurar que se cumplan niveles básicos de rendimiento, con resultados satisfactorios respecto de las inversiones realizadas. Se trataría de racionalizar la aplicación de los escasos recursos, tomando conciencia de la responsabilidad pública de las instituciones respecto del cumplimiento de sus objetivos. Se buscaría mejorar los programas y las instituciones, facilitando la asignación de recursos. Un régimen eficaz de evaluación debería combinar la capacidad institucional para lograr las metas propuestas, el funcionamiento efectivo de la

TABLA 1: MÉTODOS UTILIZADOS PARA EVALUAR I+D

Métodos	Pros	Contras
Análisis Bibliométrico	Cuantitativo. Útil sobre bases agregadas para evaluar calidad de ciertos programas y campos.	Las mediciones son sólo cuantitativas. No es útil en todos los programas y campos. Las comparaciones entre campos y países son difíciles. Pueden ser artificialmente influenciados.
Tasa de retorno económico	Cuantitativo. Muestra los beneficios económicos de la I+D	Mide sólo beneficios financieros, no beneficios sociales (tales como el mejoramiento de la calidad de la salud). El tiempo transcurrido entre la I+D y el beneficio económicos es frecuentemente largo. No es útil en todos los programas y campos.
Revisión por pares	Método de fácil comprensión. Permite evaluar la calidad de la investigación y a veces otros factores. Una parte importante de los programas financiados con fondos federales evalúan la calidad con este método.	Se focaliza principalmente sobre la calidad de la investigación. Otros elementos son secundarios. Generalmente sirve para evaluar proyectos, no programas. Gran diversidad entre las diferentes agencias. Reparos contra el uso de redes de “viejos”. Los resultados dependen de que se involucre gente de un alto nivel en el proceso evaluador.
Estudio de casos	Permite comprender los efectos de los factores institucionales, organizacionales y técnicos que influyen en el proceso de investigación. Ilustra acerca de todos los tipos de beneficio del proceso de investigación.	Casos aislados, no comparables entre programas. Focalizar la evaluación sobre casos puede acarrear dificultades para evaluar los beneficios de una programación.
Análisis retrospectivo	Útil para identificar vínculos entre programas de I+D e innovaciones en períodos largos de inversión en I+D.	No es útil para evaluación de corto plazo, debido a que el intervalo entre la I+D y sus resultados prácticos es largo.
Punto de referencia (<i>benchmarking</i>)	Provee una herramienta de comparación entre programas y países.	Puede ser focalizado sobre campos, no sobre programaciones.

organización y constatar la medida en que se satisfacen los estándares de la institución y/o del gobierno. La evaluación debería orientarse de manera que posibilitara verificar la efectividad del funcionamiento de una institución o programa, buscando también el reaseguro de ciertos estándares definidos por la propia comunidad científico académica.

Por su parte, Brunner (1990) plantea diferentes tipos de evaluación: a) la **autoevaluación** por la unidad institucional; b) **juicio de pares externos** que puede realizarse o no mediante visitas, guiándose por normas provistas por el sistema, focalizándose en el logro de metas y/o la efectividad del sistema; c) **aplicación de indicadores de performance**, que pueden ser definidos por el sistema o el gobierno. Este procedimiento descansa en una información cuantitativa referida a procesos o productos. Este mecanismo no implica autoevaluación ni tampoco la utilización de pares externos. Por otro lado, discrimina la **evaluación como proceso estructurado**, que utiliza algunos de los tipos expuestos en el párrafo previo, pero sus resultados se aplican para la asignación de recursos, incentivos, sanciones; mientras que la **producción de decisiones categoriales** se suele utilizar para la elaboración de *rankings*. Por ejemplo, en el Programa de Incentivos a docentes investigadores, implementado en nuestro país, se da un **proceso estructurado** de evaluación, combinado con el mecanismo de **juicio de pares**.

De lo expuesto se deriva que existen diferentes regímenes de evaluación que contemplan, a su vez, diferentes dimensiones a considerar. Hablar de régimen significa considerar también que existen diversos tipos de instrumentos a aplicar, pudiéndose combinar los indicadores con el juicio de pares o la autoevaluación con la evaluación externa. El tema de la evaluación en educación superior, en su conjunto y con propósitos públicos, exige pensar en el diseño de un régimen de evaluación específico. Esto no sólo depende de decisiones técnicas, ya que hacen falta condiciones políticas favorables tanto en el sistema, como en el estado y en la opinión pública en general. Parece necesario negociar las decisiones del diseño de la evaluación para hacer posible su posterior aplicación.

¿Cómo evaluar comunidades científico - académicas?

A partir de la década de los setenta se han desarrollado criterios aplicables en un nivel internacional para evaluar la productividad en el trabajo

científico; estos parámetros surgen vinculados al caso de la Física y se generalizan a todas las disciplinas de investigación. La crítica fundamental a estos criterios se basa en que no reconocen las diferentes tradiciones disciplinares, ni las diversas formas de tratamiento de los objetos de estudio, así como de los mecanismos de difusión de los conocimientos y la vinculación con el medio socio-económico de las disciplinas humanísticas, sociales, naturales o de transferencia tecnológica. Se cuestiona la legitimidad de los espacios de validación de los diferentes conocimientos que se generan en las comunidades científicas, como las publicaciones en inglés o la validez de los referatos, situación que inevitablemente nos remite al evento descrito por Alan Sokal (1998) en su libro "Imposturas Intelectuales". Estos criterios y procedimientos de evaluación tienden a una homogeneización, desde las políticas públicas, de realidades muy diversas, tanto disciplinares como socio políticas. A esta altura de los acontecimientos, no se puede eludir el tema del poder que se debate entre las diferentes comunidades académicas, actitudes muchas veces corporativas que tratan de cubrir la lucha por el acceso a recursos y financiamientos siempre escasos.

La evaluación de comunidades científicas debería entenderse como algo diferente de la simple medición de las publicaciones que la suma de sus integrantes produce. La "cientometría" puede considerarse como un elemento auxiliar en esta evaluación. La literatura específica habla de evaluación de investigación, realizada por la aplicación de indicadores de producción, productividad e impacto de las publicaciones, teniendo en cuenta la participación en revistas internacionales de prestigio con referato.

Es necesario, sin embargo, contextualizar estas prácticas sociales. Tal como manifiesta Albornoz (2000: 6):

Desde todos los ángulos, el establecimiento del "mapa" del genoma humano constituye un logro muy representativo de nuestra sociedad contemporánea. Ciencia, política y economía se entrelazan íntimamente. Sin embargo, los panegiristas que entonan el himno al conocimiento y a las fronteras casi infinitas que éste ofrece fa la humanidad parecen olvidar que la realidad social es más compleja y encierra intereses contrapuestos. Los beneficios de la ciencia no son iguales para todos. La ciencia reproduce la estructura social y se

convierte en un instrumento que hace más ricos a los ricos, más fuertes a los fuertes y más pobres a los pobres. Finalmente, en la realidad prevalece el "efecto Mateo: Dios le da más al que más tiene".

Citando a Derek de Solla Price, Brunner (1990) menciona el desarrollo de sofisticados paradigmas de "medición de ciencia", que intentan comparaciones de disciplinas, instituciones y países, mediante la construcción de modelos de análisis sobre el comportamiento y la dinámica de diversas disciplinas y áreas de investigación, su estructura interna y las redes de "co - citación"; recordemos que actualmente en algunas disciplinas muchos investigadores ya no cuentan sus publicaciones internacionales con referato, sino las veces que son citados por colegas altamente calificados en instituciones de primer nivel internacional, esto hace que se generen mecanismos distorsionantes por los cuales se construyen las mencionadas redes de "co-citación" como mecanismos de reciprocidad, por llamarlos de alguna manera ética. Preguntarse si es posible evaluar las comunidades científico académicas, implica bucear en la sociología de la ciencia y tratar de definir qué se entiende por comunidad científico académica. En este trabajo se hace referencia a la entramada red de grupos específicos de practicantes que se hallan investigando en una misma disciplina o especialidad, algunos de los cuales también se dedican a la enseñanza superior, y que se encuentran vinculados entre sí por relaciones personales, profesionales e institucionales. Se trata de evaluar disciplinas o especialidades, unidades de investigación de diverso nivel (Grupos, Proyectos, Programas), con diferente grado de desarrollo (en formación, formados). Lo que se busca es poder evaluar el sistema científico académico global de un país.

En este punto podemos señalar *in extenso* lo expresado por Mario Albornoz (1997: 97):

Como doctrina económica, el "pensamiento único" reposa sobre tres pilares macroeconómicos ortodoxos: rigor monetario, rigor presupuestario y flexibilidad salarial. En ciencia y tecnología, el pensamiento único se basa en la hegemonía casi absoluta de la óptica de la innovación por sobre cualquier otra dimensión en base a la cual pudiera ser orientada la actividad científica. No es casual que esto

ocurra, ya que esta perspectiva implica la reducción del conocimiento científico y tecnológico en un hecho fundamentalmente económico; no solamente esto, sino que además se le adjudica el carácter de instrumento fundamental para el logro de un valor cargado de intereses e ideología: la competitividad.

Continúa en el mismo texto diciendo que:

La reflexión acerca de los instrumentos conduce, simultánea pero distintamente, al plano de los fines y al de la eficiencia. Contiene implícito el análisis, propio de los estudios sociales de la ciencia y de la tecnología, de los procesos sociales del conocimiento y de los actores intervinientes; actores dotados de intereses reales y concretos cuya articulación determina el sentido de las políticas. Por una parte, la discusión sobre los fines queda expedita cuando se le quita el chaleco de la compulsión hacia el modelo exclusivo. Cuando ello ocurre, se iluminan los distintos rumbos que las decisiones políticas pueden adoptar, en el escenario concreto de los actores y sus circunstancias, entendidas como la intersección de capacidades, oportunidades y restricciones. Por otra parte, la discusión sobre los instrumentos instala el tema en el plano de las políticas públicas, en el que se rescata su dimensión técnica, sin perder de vista su politicidad esencial.

En un intento de sistematización de la temática referida a la evaluación, Brunner (1997) plantea que los principales rasgos a considerar son:

- a. **Los alcances de la evaluación:** tipos de investigación involucradas, niveles en que operará la investigación (país, sistema universitario, disciplina, universidad, grupos de investigación) y tiempo a que se refiere la evaluación (*ex - ante, interim, ex - post*).
- b. **El propósito de la evaluación:** se debe tener claro cuál es el objetivo de la misma, se busca medir producción, se quiere conocer la posición relativa del país respecto a una región o al mundo, se quiere calcular la rentabilidad del gasto público y/o privado invertido, se busca consolidar el desarrollo.

- c. **Los criterios a emplearse:** éstos son variados y deben tener en cuenta la pertinencia para el desarrollo de país, la incidencia en áreas prioritarias, la vinculación con el medio productivo. En general, cuando se evalúan proyectos de investigación competitivamente se considera la excelencia del mismo (definida por pares) y/o la trayectoria del o de los investigadores, muchas veces se olvidan los contextos y se produce el mencionado efecto Mateo, debiendo las instituciones responsables prever áreas prioritarias de desarrollo.
- d. **La organización del proceso de evaluación:** Se debe definir claramente quiénes serán los evaluadores (en países con comunidades científico académicas pequeñas se puede recurrir a pares extranjeros), pero de acuerdo con la temática a investigar se puede recurrir a funcionarios públicos y usuarios, sobre todo en temas de gestión o de transferencia al medio productivo (la manera de evaluar el FONTAR en Argentina es un buen ejemplo de esta modalidad). Es importante definir previamente tanto los criterios, los circuitos administrativos, los cronogramas y garantizar, de ser necesario, la periodicidad de las evaluaciones.

Desde su perspectiva, Albornoz (2003: 22) expone detalladamente que:

[...] otros modelos organizativos de la evaluación distintos a los tradicionales, que se integran exclusivamente por miembros reconocidos de la comunidad científica: los pares. Los nuevos modelos organizativos difieren fundamentalmente en que la constitución de los comités es heterogénea por la inclusión de otros actores. Todo esto lleva al concepto de complejidad (multiplicidad de perspectivas legítimas, no linealidad, emergencia de los problemas, organización, multiplicidad de escalas, incertidumbre).

La inclusión de nuevos actores en los procesos de investigación científica implica una auténtica democratización del conocimiento. En muchos casos ahora se hace necesario incluir a quienes pueden opinar sobre aspectos de su interés (problemas tales como enfermedades, polución y preservación ambiental, pero también opresión, discriminación y explotación). Adicionalmente, cierta parte del conocimiento se presenta en nuevos contextos (comunidades

indígenas o métodos terapéuticos tradicionales, por ejemplo). En este sentido, se puede afirmar que la ciencia moderna está renovando su localización social en un contexto enriquecido. La tendencia actual es hacia el diseño de procesos de evaluación que combinen diversos métodos:

Evaluación por pares científicos- basada en criterios de excelencia, importancia o calidad.

Evaluación por distintos actores- basada en criterios de oportunidad, permeabilidad, explotabilidad y aplicabilidad de los conocimientos.

Es importante señalar que la multiplicidad de criterios implica necesariamente la multiplicidad de actores ya que, del mismo modo que los no científicos están incapacitados para opinar acerca de la calidad científica de una propuesta, los pares científicos no son, de por sí, los adecuados para opinar acerca de la utilidad de ellos. Ahora bien, el método mixto tiene el problema de que requiere una definición acerca del papel de los actores y del peso relativo de los criterios de evaluación. Por otra parte, la inclusión de los no-pares genera desconfianza en la comunidad científica.

La evaluación en un caso concreto: el Programa de Incentivos a Docentes Investigadores (Decreto 2420/93)

Uno de los objetivos primordiales del Programa de Incentivos a Docentes - Investigadores es el de imponer una cultura de la evaluación en el sistema universitario argentino, sistema diagnosticado como trasgresor y poco avenido a ser evaluado. Como se mencionara, en términos de Brunner, para este Programa se ha seleccionado un modelo de evaluación de proceso estructurado combinado con la evaluación por pares. Si bien se denomina a Docentes - Investigadores, queda claro que lo que se evalúa es el área de investigación; la evaluación de la docencia, según el Ministerio, es responsabilidad de cada universidad, estando aún pendiente de resolución este punto, ya que se controlan la cantidad de horas de clases, pero no se evalúa la calidad de la actividad en docencia.

A pesar de la implementación del Programa desde 1994 y el cumplimiento de las evaluaciones anuales periódicas de los Informes de Avance y Finales,

así como de la Auditoría realizada por la Coordinación del Programa en 1999, aún no se pueden evaluar cambios en la calidad académica de los graduados. Pero la actividad de investigación está impregnada de prácticas de evaluación consensuadas y realizadas explícita o implícitamente. Se dan, por ejemplo, cada vez que se acepta un *paper* para ser publicado en una revista con referato o cuando un autor es citado como referente por un colega; o bien cuando se expone un trabajo en un congreso o se adjudican premios y subsidios, y en cada oportunidad que se logra el reconocimiento de la comunidad de pares.

En el marco del Programa de Incentivos se efectúa una evaluación "ex - ante" de los Proyectos que participarán o no del Programa y evaluaciones "ex - post" anuales y periódicas a través de los informes de avance y finales de los Proyectos, siendo evaluados como "satisfactorios" o "no satisfactorios". Por otro lado se realizan las "categorizaciones", donde se evalúa al docente - investigador que se postula a la categoría que considera es la que le corresponde (de acuerdo con la normativa originaria de Programa las categorías eran cuatro: A, B, C y D, siendo la A las de más alta jerarquía correspondiendo a Investigador Principal del CONICET; actualmente se prevén cinco categorías que son I, II, III, IV y V, jerarquizadas también de manera descendente). Para poder participar del Programa, y acceder al cobro del incentivo, es necesario estar previamente categorizado. En casos de no coincidir la categoría otorgada con la solicitada por el docente - investigador, están previstos Tribunales de Alzada para resolver recusaciones con plazos para impugnaciones.

La evaluación por pares, que se utiliza, es un sistema mediante el cual la excelencia intelectual de un trabajo es juzgada por expertos de la misma área de conocimiento. Entre los criterios considerados figuran la calidad del proyecto, su coherencia, la factibilidad de llevarlo adelante, la dirección, el impacto de la investigación, así como la originalidad del tema. A fin de evitar arbitrariedades se utilizan evaluadores de categoría de investigación igual o mayor a la del director del proyecto. Existe una base de evaluadores que provee la Secretaría de Políticas Universitarias (SPU) y los investigadores pueden recusar a algún miembro del listado de evaluadores, con motivos fundados, así como pueden indicar la comisión disciplinaria (organizadas por campo de conocimiento) en la que quieren ser evaluados. Entre los investigadores existe un consenso muy alto con respecto a este sistema de

evaluación. Del mismo modo, los organismos que financian investigación lo consideran el único sistema posible de ser utilizado para juzgar la excelencia científica. Sin embargo, a pesar de esta aceptación, se han detectado algunas desventajas como por ejemplo:

- prejuicios por parte de los evaluadores;
- en comunidades científicas pequeñas, los jueces también son parte (los evaluados también son evaluadores) y esta situación puede prestarse a arbitrariedades, ya sea por amiguismo o por enemistad;
- de acuerdo con los evaluadores que participen pueden privilegiarse los proyectos tradicionales sobre los innovadores.

Si bien se acepta que consta de algunas imperfecciones, la evaluación por pares es el mecanismo con mayor consenso para evaluar en el campo de la investigación. El poder que pueden ejercer quienes elijan a los evaluadores desde la gestión, en las respectivas Secretarías de Ciencia y Técnica de cada universidad, se ve limitado por la posibilidad del evaluado para recusar algún miembro de los potenciales jueces que están publicados en la página *web* de la Secretaría de Políticas Universitarias.

En la evaluación se tienen en cuenta aspectos cuantitativos, debidamente ponderados, y aspectos cualitativos. Entre los primeros cabe mencionar infraestructura científica, ayudas a la investigación, contratos, publicaciones científicas, patentes, participaciones en congresos y reuniones científicas, dirección de tesis de postgrado y estancias académicas en centros extranjeros. En cuanto a los aspectos cualitativos, como los factores de impacto para las ciencias experimentales, matemática y ciencias de la salud, existe una herramienta importante que es el *Science Citation Index (SCI)*. Los análisis se efectúan sobre los resultados de investigaciones que aparecen en las publicaciones científicas periódicas que recoge el SCI. Si entre los científicos de estas ciencias existe un consenso en cuanto a su utilización, no es así para las humanidades y sociales en las cuales existe mayor dificultad para consensuar criterios para la evaluación. El mismo objeto de estudio condiciona las posibilidades de publicación en inglés por ejemplo, así como sus espacios de legitimación. No obstante, en términos generales, hay coincidencia en evaluar:

- a) la coherencia del programa,

- b) la distribución de los resultados entre los miembros del grupo,
- c) la originalidad del proyecto,
- d) la calidad de las publicaciones,
- e) la productividad, definida por el número total de estudios y de tesis doctorales y, en algunas disciplinas, los informes externos.
- f) la viabilidad del tema,
- g) la relevancia de la investigación.

Algunas desventajas del proceso de evaluación vigente

El sistema de evaluación vigente posee algunas desventajas que son importantes considerar:

Costos de la evaluación anual: Como se indicara oportunamente, la evaluación en el Programa de Incentivos es anual, involucrando los nuevos Proyectos presentados, Informes de Avance e Informes Finales. Esto implica gastos de viáticos y horas / hombres para el desplazamiento de docentes - investigadores que han sido convocados como evaluadores a fin de mantener en las Comisiones el 50 % de la externalidad; el alquiler de lugares para reunión de Comisiones Evaluadoras, gastos administrativos. En muchas oportunidades, cuando no se encuentran especialistas o ellos no pueden desplazarse por motivos personales, se les envía el material a evaluar por correo, generándose en la universidad otro gasto extra. Por otro lado, es necesario considerar los costos de comunicaciones telefónicas, faxes y las horas/hombre del personal asignado para realizar tareas de apoyo mientras las Comisiones se encuentran reunidas evaluando. Algunas universidades han optado por pagar a los evaluadores por proyecto o informe; otras instituciones lo consideran una obligación académica.

Desde la perspectiva del investigador, se observa que para la solicitud de categorización, la formulación de los Proyectos y la elaboración de los Informes de Avance y Finales, debe poseer un equipamiento informático adecuado para el *software* provisto por la SPU. Es necesario contar con papel para las impresiones requeridas y cartuchos para imprimir; se realizan gastos en fotocopias y se invierten varias horas hombre dedicadas a la tarea de organizar todas las certificaciones acumuladas de las actividades realizadas. Pareciera un dato menor,

pero si se tiene en cuenta los montos del incentivo asignados, resulta un costo de importancia.

Dificultades para reunir las Comisiones Interdisciplinarias de pares. Esto se debe a que el cronograma de evaluación es el mismo para todas las universidades nacionales, situación que genera demoras ya que la comunidad científico - académica argentina es pequeña y los especialistas que cumplen con los requisitos necesarios para ser evaluadores son pocos, demorándose el trabajo y la devolución de los resultados, lo que dificulta las tareas de llenado de las bases exigidas por el Programa.

Concentración de trabajo en determinadas áreas disciplinarias que cuentan con pocos investigadores I y II.

Cronogramas de evaluación que a veces coinciden con compromisos de los evaluadores, muchos de los cuales también deben ser evaluados.

Falta de respeto de los convenios de confidencialidad.

Dificultades para compatibilizar criterios de evaluación, ya sea entre las diferentes disciplinas como entre las universidades.

Se intenta aplicar criterios homogéneos para todas las disciplinas lo que genera serios conflictos, tanto por sus diferentes intereses y desarrollo como por el grado de acceso a los recursos disponibles.

CONSIDERACIONES FINALES

Después de un barrido de diferentes perspectivas de evaluación en los contextos europeos y latinoamericanos, se puntualizan problemas, debates, modalidades, necesidades, respecto de esta cuestión en la que operan intereses de orden y magnitud diferentes. En relación con la importancia de esta temática, parece productiva, en lo que hace a una mirada conclusiva, una cita de Mario Albornoz (2003) que sesga algunas variables complejas:

La evaluación requiere consenso acerca del campo, los fines y los criterios. La evaluación es una tarea colectiva. De allí, la importancia de la selección del equipo que tendrá a su cargo la evaluación (imparcialidad no cuestionada). La legitimidad social de los árbitros es una cuestión fundamental de la que depende. En temas de cierta complejidad o en comunidades científicas pequeñas la idoneidad y

legitimidad de los evaluadores plantea dificultades. Frecuentemente se apela a evaluadores extranjeros cuando no existe un par reconocido en la comunidad local, o cuando se requiere una opinión muy especializada y no comprometida con intereses en juego. Sin embargo, este recurso debe ser empleado con precauciones porque los pares extranjeros algunas veces opinan a con desconocimiento del contexto, entendido como las limitaciones o prioridades locales. En el caso de evaluaciones mixtas como, por ejemplo, en la integración de paneles con científicos e industriales se presentan a veces dificultades de diálogo que atañen al lenguaje, la lógica y el papel que se espera que desempeñen unos y otros.

En cuanto a los docentes - investigadores y los gestores participantes, están tomando conciencia de la necesidad de mejorar los datos y transparentar la información recuperada, a efectos de alimentar unas bases confiables que posibiliten un mayor encuadre y un diálogo más eficaz entre organizaciones nacionales y del exterior. Los niveles de resistencia, muy visibles a comienzos de la implementación del Programa, parecen estar cediendo, lo que no significa una aceptación total del mismo.

La mentada “cultura de la evaluación” es una realidad impuesta, de la que los actores involucrados están tomando conciencia. La evaluación por pares pone en evidencia que, en la práctica, la masa crítica existente en cada disciplina es la que está determinando las líneas prioritarias de investigación ya que es la responsable de aprobar o no proyectos e informes presentados. Más allá de identificar las debilidades y fortalezas de nuestras universidades, un objetivo a cumplir sería el de mejorar la calidad en la generación y transferencia de conocimientos. Es necesario destacar lo puesto en evidencia: más allá de tratarse de un Programa que vincula la productividad con los salarios percibidos, se han puesto en juego bienes simbólicos altamente valorados en las comunidades académicas como es, por ejemplo, el prestigio.

Para el desarrollo de la Ciencia y Tecnología del país, es necesario incrementar el aporte del PBI a esta área: el país que invierte más del 1% de su PBI en I+D es un país que empieza a desarrollarse de manera más acentuada y a salir de una economía débil. Para llegar a una economía fuerte hay que

invertir más del 2,5 del PBI. Japón está en el 3%, Suiza en el 3,8, Estados Unidos en el 2,6, pero como su PBI es mucho mayor que el del resto de los países, es incomparable. La media latinoamericana es el 0,5 o 0,6 y la Argentina invierte sólo el 0,3. Brasil es el único país de la región que supera el 1%: en el 2003 invirtió el 1,5 de su PBI en investigación tecnológica. El actual Ministro de Educación, Daniel Filmus, promete pasar del 0,3 al 1% en cuatro años. Por otro lado, es necesario comprometer al sector privado a invertir en Ciencia y Tecnología.

En este artículo, como en todo trabajo de investigación, se ha pretendido responder algunas preguntas, pero se han abierto muchas otras. En relación con los propósitos con que fue ideado e implementado el Programa de Incentivos a Docentes e Investigadores, los explícitos están claramente formulados en el Decreto 2427/93 y las posteriores normativas emanadas de la SPU. ¿Pero cuáles serían los implícitos? En el Prólogo del informe de la Auditoría Técnica de los Informes Finales o de Avance de los Proyectos de Investigación (1998) del Programa de Incentivos a Docentes Investigadores”, Juan Carlos Del Bello, manifiesta que, como es sabido, el Programa ha sido cuestionado por unos pocos grupos minoritarios de la comunidad académica, con hipótesis nunca confirmadas (como suponer que los estímulos a la investigación son un mero aumento salarial, que está sobreestimado el número de docentes-investigadores activos, y que la universidad no es un buen "hogar" para las actividades de investigación). Sin embargo, a partir de los favorables resultados de la auditoría técnica, se concluye que el Programa de Incentivos ha alcanzado sus objetivos de estimular la investigación en la universidad estatal, integrar la investigación con la docencia y también mejorar los ingresos monetarios de los docentes universitarios con vocación de mejorar sus prácticas académicas.

La referencia a que “la universidad no es un buen hogar para las actividades de investigación” parece remitirnos a un viejo debate respecto de dónde debe realizarse la investigación: en las universidades, en las empresas o en los organismos de gobierno como el CONICET. No podemos dejar de señalar que en este período se crea la Agencia Nacional de Investigación y se da un profundo debate acerca del modelo de organización de investigación a seguir: ¿el francés del CENERES del cual se origina el CONICET o el de la *National Research Foundation* de USA?

Respecto de las externalidades generadas por el Programa, se presentan conflictos dentro de las comunidades científico académicas, sobre todo a partir de la vigencia del Manual de Procedimientos en el que la evaluación de la gestión tiene un alto puntaje. Esto condujo a Celso Aldao, investigador I, de la UNMDP a escribir lo que él mismo denomina su paper más conocido: al aplicar la grilla de evaluación, toma los casos de un Rector XX categorizado I y de Albert Einstein, quien al final de su vida logra la categoría II por haber recibido el premio Nobel (60 puntos). Cabe aclarar que, debido a tales inconsistencias propiciadas por la grilla en cuestión, Aldao y otros investigadores renuncian a participar como evaluadores.

Las consecuencias no deseadas de la implementación del Programa son propias de la aplicación de una política homogeneizadora a una realidad socio cultural muy diferente del ámbito donde es pensada en los orígenes (Programas *merit pay* en USA). En el momento de cerrar estas líneas, se está realizando una nueva convocatoria con ajustes, tanto de la grilla de evaluación como de condiciones formales de participación. Según sus gestores, las modificaciones realizadas aportarían a una contextualización necesaria, propiciatoria de una continuidad del Programa en un sentido más armónico. O no...

Es posible afirmar a estas alturas, que gran parte de los efectos no deseados que se adjudican a la implementación del Programa de Incentivos estaban ya radicados en la comunidad científicoacadémica nacional. Aquellos investigadores de vocación continúan trabajando, a pesar de todas las adversidades del medio, quizás otros se han sumado con el mero objetivo de cobrar incentivos; pero esta realidad no es exclusiva de nuestro país. Pierre Bourdieu (2003:50) en “El Oficio de Científico”, libro póstumo, recupera:

[...] un *Dictionary of useful research phrases* que circula de laboratorio en laboratorio y recuerda los discursos irónicos y paradójicos a propósito del discurso científico que producen los propios científicos: *Post-prandial Proceeding of the Cavendish Physical Society, Journal of Jocular Physics, Journal of Irreproducible Results, Revue of Nuclear Physics.*

Según el modelo de las listas de “debe decirse/no debe decirse” de los manuales de idiomas, los autores establecen un cuadro comparativo que confronta dos versiones de la acción: la producida para la presentación formal y la descripción informal de lo que ha sucedido realmente. A un lado “lo que escribió” (*what he wrote*); al otro “lo que pensaba” (*what he meant*) (Gilbert y Mulkay, 1984:176):

1. Sabemos desde hace tiempo...//No me he tomado la molestia de buscar la referencia.
2. Aunque todavía no sea posible ofrecer unas respuestas definitivas a esas preguntas...//El experimento no ha funcionado, pero he pensado que, por lo menos, podría aprovecharlo para una publicación.
3. Han sido elegidas tres de las muestras para un estudio detallado...//Los resultados de las otras carecían de todo sentido y han sido ignorados.
4. Dañado accidentalmente durante el montaje... // Se cayó al suelo.
5. De gran importancia teórica y práctica...//Interesante para mí.
6. Sugerimos que... Sabemos que... Parece...//Creo.
7. Se cree generalmente que...//También lo piensan otros tíos.

Este divertido listado permite descubrir la hipocresía de la literatura formal. Pero la doble verdad de la experiencia que los agentes pueden tener de su propia práctica tiene algo de universal. Conocemos la verdad de lo que se hace (por ejemplo, el carácter más o menos arbitrario o, en cualquier caso, contingente de las razones o de las causas que determinan una decisión judicial), pero para estar en regla con la idea oficial de lo que se hace, o con la idea obvia y evidente, es preciso que esa decisión parezca que ha sido motivada por unas razones, unas razones lo más elevadas (y jurídicas) posibles. El discurso formal es hipócrita, pero la tentación del “radicalismo chic” lleva a olvidar que las dos verdades coexisten, con mayor o menor dificultad,

en los propios agentes (es una verdad que me costó mucho trabajo aprender y que aprendí paradójicamente, gracias a los cabileños, tal vez porque es más fácil descubrir la hipocresía colectiva de los extraños que la propia).

Los trabajos de Sokal en “Imposturas Intelectuales” y David Lodge en “El mundo es un pañuelo”, que representa la circulación de académicos en congresos, refuerzan esta línea de pensamiento.

Sin embargo, algo no es discutible: la continuidad del Programa de Incentivos dependerá de un consenso de los actores sociales involucrados como el CIN, la CONADU, la SPU y las comunidades científico académicas. Sería de esperar que el acuerdo básico se centrara en la evaluación a aplicar que, más allá de las clasificaciones y formas desarrolladas por especialistas, apelara a un consenso que contemple una postura ética al respecto.

BIBLIOGRAFÍA

- Albornoz, M. (2003). Evaluación en ciencia y tecnología. Perspectivas Metodológicas. Ediciones de la Universidad Nacional de Lanús, Buenos Aires, Noviembre, Año III, N° 3: 17-34.
- Albornoz, M. (2000). Editorial El Genoma es ajeno (las penas son de nosotros). Revista de Estudios Sociales de la Ciencia. Universidad Nacional de Quilmes. Buenos Aires, Agosto 2000. REDES 15, 232 pp.
- Albornoz, M. (1997). La política científica y tecnológica en América Latina frente al desafío del pensamiento único. Revista de Estudios Sociales de la Ciencia. Universidad Nacional de Quilmes. Buenos Aires, Octubre 1997, REDES 10: 279 pp.
- Albornoz, M. (1996). De la anomalía Argentina a una visión articulada del desarrollo Ciencia y Tecnología. Revista de Estudios Sociales de la Ciencia. Universidad Nacional de Quilmes. Buenos Aires, Septiembre 1996, REDES 7: 26-35.
- Albornoz, M. y Fernández Polcuch, E. (1996). Indicadores de CyT. Revista de Estudios Sociales de la Ciencia. Universidad Nacional de Quilmes.

- Buenos Aires, Septiembre 1996, *REDES* 7: 8-45.
- Araujo, S. (2002). Evaluación, incentivos a la actividad investigadora y trabajo académico. Algunas conclusiones en el estudio de un caso en Argentina. En *La universidad cautiva*, Ediciones Al Margen, La Plata, Buenos Aires, 49-62.
- Araujo, S. (2001). Los universitarios en la lupa: evaluación de la calidad, incentivos a la actividad investigadora y sus efectos en la profesión académica. En *Pensamiento Universitario*, UBA, Buenos Aires, 58-69.
- Bourdieu, Pierre. (2003). *El oficio de científico. Ciencia de la ciencia y reflexividad*. Barcelona, Ed. Anagrama, 212 pp.
- Brunner, J. J. (1997). *Educación Superior en América Latina. Cambios y Desafíos*. Fondo de Cultura Económica, Chile, 248 pp.
- Brunner, J. J. (1990). *La evaluación de la investigación científica*, Centro de Investigación del Desarrollo Internacional del Canadá, Oficina Regional para América Latina, Santiago de Chile, Chile. 125 pp.
- Carullo, J. C. y Vaccarezza, L. (1997). El incentivo a la investigación universitaria como instrumento de promoción y gestión de la I+D. *Revista de Estudios Sociales de la Ciencia*. Universidad Nacional de Quilmes. Buenos Aires, Octubre 1997, *REDES* 10: 45-53.
- Clark, B. (1983). *El sistema de Educación Superior, Nueva Imagen*, México. 264 pp.
- Gibbons, M. (1998). Conferencia de UNESCO. París. 21 p.
- Gibbons, M. (1997). *La nueva producción del conocimiento*, Ediciones Pomares, Corredor S. A., Barcelona, 169 pp.
- Gilbert, G. Nigel & Mulkay, Michael. (1984). *Opening Pandora's box : a sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press. 216 pp.
- Gurmendi, M. L. (2000). *Gerenciamiento del sistema de información, Maestría Gestión Universitaria*, Fac. Cs. Económicas y Sociales, UNMDP, Mar del Plata. 128 pp.
- Hammersley, M. y Atkinson, M. 1994. *Etnografía. Métodos de investigación*. Paidós, Barcelona. 238 pp.
- Neave, Guy, (2001). *Educación Superior: Historia y Política. Estudios comparativos de la universidad contemporánea*, Gedisa, Barcelona, 365 pp.

- Neave, Guy. (1998). La reforma de la educación superior francesa, o la fábula del buey y el sapo. En *Prometeo encadenado: estado y educación superior en Europa*. Neave Guy y Van Vught, F. (comp.). Barcelona: Gedisa Editorial.
- Krotsch, P. (2002). *La universidad cautiva*, Ediciones Al Margen, La Plata, Buenos Aires, 225 pp.
- Krotsch, P. (2003). *Las miradas de la Universidad*, Ediciones Al Margen, La Plata, Buenos Aires. 172 pp.
- Prati, M. (2002). El Programa de Incentivos a los Docentes Investigadores: formulación, implementación y visiones sobre su impacto. En *La universidad cautiva*. Ediciones Al Margen, La Plata, Buenos Aires, 101-125.
- Salomón, J. J. (1994). Tecnología, diseño de políticas de desarrollo. *Revista de Estudios Sociales de la Ciencia*. Universidad Nacional de Quilmes. Buenos Aires, Septiembre 1994, *REDES 1*: 35-48.
- Sokal, Alan. (1998). *Imposturas intelectuales*. Paidós, Madrid, 146 - 165.