

	<p style="text-align: center;">Relatoría del <i>Taller de trabajo y reflexión sobre la evaluación académica</i></p> <p style="text-align: center;">Unidad Académica de la CSIC</p> <p style="text-align: center;">Setiembre, 2018</p>	
---	--	---

1. Introducción (recordatorio de la actividad del Taller)

En la actualidad las modalidades prevaletientes de la evaluación académica están en cuestión. Más específicamente, lo que está siendo cuestionado no es la evaluación en sí sino las modalidades en que se está llevando a cabo en muchas partes del mundo. La justificación para expresar “problemas en las modalidades prevaletientes de la evaluación académica” radica en la existencia de un proceso de homogeneización a nivel internacional de las mismas. La experiencia de distintos países, con sistemas científicos y demandas a dichos sistemas muy diversas, muestra o señala problemas similares. Como consecuencia de los problemas identificados, están emergiendo propuestas alternativas a las formas prevaletientes de evaluación académica.

A la luz de las críticas a las modalidades prevaletientes de evaluación, así como de la aparición de propuestas alternativas, la CSIC entendió importante promover la discusión del tema a través de la organización del *Taller de trabajo y reflexión sobre la evaluación académica* que tuvo lugar el 8 de junio de 2018 en la Facultad de Ciencias Económicas y de Administración (FCEyA) de la Udelar.

El objetivo de dicho taller fue generar un espacio de diálogo y reflexión para conocer la percepción de los participantes sobre el tema, legitimar la pluralidad de perspectivas, alcanzar algunos consensos y/o clarificar las diferencias, y socializar algunas contribuciones teóricas y empíricas de la literatura especializada en el tema.

El taller se desarrolló durante todo un día de trabajo. La mañana se concentró en una plenaria que comenzó con la apertura a cargo del Prof. Roberto Markarián, Rector de la Udelar, una bienvenida realizada por el Prof. Rodrigo Arim, actual Ex - Decano de la FCEyA, y con la presentación de los objetivos y forma de trabajo por parte de la Profa. Cecilia Fernández, Pro-rectora de investigación. Posteriormente, los profesores Luis Acerenza, Marita Urquhart y Cecilia Fernández realizaron una exposición trayendo a la memoria la experiencia del PEDECIBA del año 2004, cuando a partir de su preocupación por los problemas que percibían en la evaluación de los investigadores, realizaron un conjunto de propuestas para superarlos. A continuación, la Profa. Judith Sutz, Coordinadora de la Unidad Académica de la CSIC, realizó una presentación sobre algunos aspectos destacados de la discusión internacional sobre el tema. La sesión plenaria de la mañana culminó con cerca de dos horas de intercambio entre todos los asistentes al taller.

En la tarde, la actividad continuó en tres talleres en simultáneo. En cada uno de esos talleres los participantes trabajaron, intercambiaron y reflexionaron, durante dos horas aproximadamente, en torno a tres ejes: los problemas de la evaluación, sus consecuencias y propuestas alternativas. La actividad del día culminó con una sesión plenaria de cierre, donde los voceros -elegidos en cada uno de los talleres- presentaron una pequeña síntesis de lo discutido, luego de lo cual se abrió un espacio para el intercambio entre todos los presentes y por último hubo un pequeño cierre.

Las sesiones plenarias de la mañana fueron transmitidas en vivo por el canal de YouTube de CSIC y están disponibles en [_https://www.youtube.com/watch?v=ah3KDDMr2y4&t=3403s](https://www.youtube.com/watch?v=ah3KDDMr2y4&t=3403s)

Al taller asistieron en el entorno de 280 personas, incluidos docentes-investigadores, evaluadores, tomadores de decisiones de política, gestores de

CTI y estudiantes, pertenecientes a los distintos servicios y dependencias de la Udelar, a las Comisiones Sectoriales, al Espacio Interdisciplinario, a la Comisión Coordinadora del Interior, a los Centros Universitarios Regionales, etc. y a otras instituciones como INIA, PEDECIBA, IIBCE, CEI, ANII, MIDES, MEC y LATU. Si bien la cantidad de asistentes fue disminuyendo hacia la tarde, cerca de 160 personas participaron de los tres talleres en simultáneo y de la sesión plenaria de cierre.

2. Relatoría del Taller

Esta relatoría se realiza con base en las notas tomadas por los docentes de la UA de la CSIC a lo largo de toda la actividad. Consiste en cuatro secciones: la primera es un listado sintético de temas que emergieron en las intervenciones y diálogos de los asistentes; las tres secciones restantes desarrollan dichos temas tal como fueron discutidos en el taller.

2.1. Los problemas que emergieron (a grandes rasgos)

Se listan los principales temas que emergieron de dicha discusión, sintetizados en 24 puntos que se organizan en dos grandes ejes: i) críticas y justificaciones a las formas vigentes de evaluación académica y ii) características para una evaluación académica alternativa.

Críticas y justificaciones a las formas vigentes de la evaluación académica

1. Falta de consideraciones de contexto al evaluar la investigación.
2. Énfasis en lo individual en detrimento de lo colectivo.
3. Evaluaciones declarativamente enfocadas en la mejora académica pero que son recibidas por los investigadores como juicios netos sin orientación a dicha mejora.
4. Evaluaciones superpuestas y sin armonización de criterios en los diversos sub-sistemas académicos a los que pertenecen los docentes (SNI, RDT -acceso y renovación, cargos docentes en los servicios).
5. Preocupación por el tipo de señales que están recibiendo los doctorandos en su formación como investigadores.
6. Evaluación excesivamente sesgada hacia la investigación en detrimento de la valoración de otras funciones docentes muy demandantes de tiempo -por ejemplo, construcción institucional.

7. Excesivo énfasis en cierto tipo de productos, lo que lleva a una homogeneización de criterios de evaluación que no toma en cuenta la diversidad de formatos en que se produce conocimientos en las diferentes áreas.
8. Inadecuación de los criterios bibliométricos por limitaciones en las bases de datos que éstos utilizan.
9. Tendencia a la “delegación”, por la cual se valora una contribución más por el medio en que es difundida que por sus méritos intrínsecos.
10. La extendida crítica a los indicadores bibliométricos es respondida con las razones que justifican su utilización: i) permiten utilizar criterios que homogeneizan la evaluación; ii) evitan los círculos particularistas de poder; iii) evitan los sesgos de la endogamia; iv) cuando hay cientos de evaluaciones que realizar no hay tiempo para leer trabajos; v) pérdida de legitimidad de la evaluación por pares.
11. El énfasis en obtener resultados publicables en revistas internacionales tuvo impacto positivo en ciertos ámbitos donde se creía que ello no era posible, impulsando la investigación.
12. Preocupación por los posibles sesgos introducidos en las agendas de investigación por los criterios de las revistas internacionales.
13. Preocupación por el impacto de la preeminencia dada en la evaluación a revistas internacionales por sobre las revistas nacionales y regionales -que no están en inglés-.
14. Excesivo énfasis en productos (productividad) en el marco de un incremento de las actividades docentes que somete a una presión indeseable sobre la calidad de las actividades académicas.
15. En el afán por dar cuenta de resultados (derivada del tipo de evaluación vigente) se agregan innúmeras actividades en los CVs, haciéndolos ininteligibles y poco manejables.
16. El énfasis cuantitativo conspira contra investigaciones de mayor riesgo, de largo plazo, más volcada a problemas nacionales.
17. Los formatos únicos de evaluación perjudican a los esfuerzos interdisciplinarios, a la investigación en temáticas emergentes y a las áreas que, por su naturaleza, requieren mucho tiempo para producir evidencias.

Características de las alternativas delineadas a la evaluación académica vigente

18. Incrementar la valoración cualitativa.
19. Asociar los criterios de evaluación a los objetivos de la investigación.

20. Evaluar productos pero también procesos (de investigación).
21. Debe ser integral, valorando el conjunto de actividades docentes.
22. Las diversas modalidades del contexto de la investigación tienen que tomarse en cuenta (área cognitiva, nivel de desarrollo, tipo de investigación).
23. La evaluación, más allá de tomar en cuenta diversas especificidades debe tener en común un alto nivel de exigencia, lo que incluye reconocimiento internacional.
24. Atención a las devoluciones de la evaluación, que deben reflejar el propósito de generar aprendizajes a partir de las mismas.

2.2. Problemas actuales de la evaluación académica en Uruguay

2.2.1 Falta de claridad en los objetivos de la evaluación

La evaluación académica puede tener diversos objetivos, a saber: i) acceder a cargos y a la carrera académica; ii) fomentar y estimular trayectorias académicas individuales, grupales y/o institucionales; iii) asignar recursos físicos y económicos; iv) proveer incentivos materiales y simbólicos -prestigio-; v) rendir cuentas a la sociedad. La evaluación académica puede ser a la vez considerada como un instrumento para la mejora y el fortalecimiento individual, grupal e institucional. Sin embargo, y desde la percepción de los participantes en el taller, un problema que parece recurrente es la falta de claridad respecto a los propósitos de la evaluación académica en Uruguay.

2.2.2 Evaluaciones descontextualizadas

La evaluación académica no considera suficientemente el contexto en el que desempeñan sus actividades los individuos, grupos e instituciones y tampoco sus condiciones materiales de base (acceso a recursos, infraestructura, materiales, equipamiento), ni las dificultades diferenciales que ello implica en términos del desempeño académico posible. A la vez, la evaluación se centra en el desempeño académico individual y no reconoce que el mismo depende del trabajo colectivo y del grado de desarrollo y/o avance de las trayectorias académicas individuales y grupales. A pesar de lo anterior, los participantes del taller reconocen que algunos pasos se han dado hacia la implementación de evaluaciones más contextualizadas pero insisten en que se debe avanzar más en esa dirección. También señalan que si bien en la retórica se resalta que la

finalidad de la evaluación es mejorar la calidad académica, en la práctica la comunidad académica nacional no lo percibe. En ese sentido, señalan necesario mejorar los mecanismos de evaluación y la comunicación.

2.2.3 Sobre-evaluación y sus sistemas de señales orientadoras en conflicto

Los asistentes al taller perciben -como un problema relativamente generalizado- que están sometidos a evaluaciones superpuestas que proveen *sistemas de señales* diferenciales y con frecuencia en conflicto. En particular destacan dos tipos de tensiones. Uno asociado entre el sistema de señales que se deriva de las evaluaciones realizadas en el Sistema Nacional de Investigadores y las que se derivan de las evaluaciones implementadas en la Udelar. El segundo, al interior de la propia Udelar, entre la evaluación que se realiza en el RDT y la que se realiza para el acceso y la renovación de los cargos docentes.

Algunos participantes del taller resaltaron la necesidad de incidir desde la Udelar en el sistema de señales que genera el SNI, propendiendo a una mayor alineación/coordinación entre los sistemas de señales que generan ambas instituciones con sus evaluaciones. Finalmente, algunos participantes también manifestaron cierta preocupación por el tipo de señales que reciben los doctorandos en sus procesos de formación como investigadores.

2.2.4 Las limitaciones de la evaluación estandarizada para apreciar la diversidad académica

De acuerdo con la percepción de asistentes al taller, un problema recurrente es la ausencia de herramientas para evaluar la diversidad de actividades académicas que trascienden a la publicación de artículos científicos -*papers*-. Según ellos, hay una gran tensión entre el discurso político respecto a la integralidad de las funciones -enseñanza, investigación, extensión, gestión y construcción institucional que deben desempeñar los docentes de la Udelar- y las formas de evaluación de la institución que se concentran fundamentalmente en el reconocimiento de algunos productos derivados de la investigación, y por lo tanto sub-valoran el desempeño en las restantes funciones. Consideran que hay una sub-valoración de las funciones docentes y

que debería evaluarse la producción docente o en enseñanza y la construcción institucional. Esta última, además, requiere mucho tiempo, esfuerzo y representa la base para el desarrollo de las actividades de investigación, sin embargo muchos de sus aspectos no son atendidos en la evaluación.

Asimismo, los participantes resaltan que en la evaluación académica no se validan otros productos de la investigación distintos de los artículos científicos. Plantean un conjunto de críticas respecto al uso de indicadores bibliométricos como herramienta para la evaluación, ya que se calculan sobre bases de datos bibliográficas acotadas y, por lo tanto, sesgadas en términos de su cobertura disciplinar, idiomática y de los productos de investigación que contienen. Esto representa una reducción de la actividad académica en la medida que se califica únicamente la publicación de artículos científicos en revistas arbitradas indexadas a ciertas bases de datos bibliográficas. En ese sentido, señalan que la cantidad de artículos publicados es un mal indicador de la producción científica.

Señalan que esa forma de evaluar la producción científica desconoce las especificidades de los diversos campos disciplinares en términos del tipo de productos que generan y los canales más apropiados para su difusión. En ese sentido ven con preocupación la tendencia a valorar los productos de la investigación en función del canal de difusión en que se hacen visibles y no por los méritos intrínsecos de ese producto. Así, los productos de investigación orientados a informar la toma de decisiones en el ámbito de las políticas públicas y el Estado quedan invisibilizados y son escasamente valorados. A la vez, señalan que esta forma de evaluar no contribuye ni a la construcción ni al fortalecimiento de la comunidad académica nacional, porque los procesos de evaluación así implementados delegan y no asumen la responsabilidad de orientar el trabajo de investigación hacia la superación de sus principales debilidades. Es decir, esa evaluación no contribuye al aprendizaje y el fortalecimiento de las capacidades de investigación y, además, refuerza las relaciones jerárquicas de producción de conocimiento existentes a nivel mundial. Importa resaltar que lo planteado no significa que no sea importante publicar en revistas arbitradas de circulación internacional, todo lo contrario, el

problema está en que suele reducirse la diversidad de la actividad académica y sus productos sólo a eso.

De ahí es que sugieren la necesidad de buscar y encontrar mecanismos propios de evaluación que contribuyan a la construcción colectiva de la comunidad académica nacional y su aprendizaje. En esa dirección, resaltan la necesidad de utilizar todas las herramientas disponibles para la evaluación, tanto cuantitativas como cualitativas, para dar garantías, evitar prácticas endogámicas y discrecionales y para apreciar sustantivamente el contenido de la producción de conocimiento de los distintos campos del conocimiento en los diversos formatos en que se expresa y difunde.

2.2.5 Las métricas como criterios superadores de la discrecionalidad en la evaluación

De acuerdo con la percepción de algunos asistentes, los indicadores bibliométricos como criterios de evaluación se adoptaron también como una forma de evitar la discrecionalidad y la endogamia en la evaluación y para establecer criterios comunes para todos. Según ellos, la prevalencia de grupos de poder en la institución y la endogamia existente en los procesos de evaluación, han hecho mucho daño al desarrollo y solidez de la comunidad académica nacional. Así el uso de estos indicadores busca contrapesar *“la arrogancia y la mediocridad”*.

2.2.6 Las métricas como criterios para afrontar el volumen creciente de evaluaciones

Entre los asistentes al taller una pregunta recurrente fue cómo evaluar cuando hay cientos de propuestas. Algunos de ellos destacaron que en las evaluaciones como las realizadas en el SNI no se pueden leer todos los trabajos, además de la pérdida de legitimidad de la evaluación por pares; por estas razones es que predominan los criterios de evaluación adoptados por el área básica, donde algunos índices y bases de datos tienen más valor que otros.

En ese sentido se cuestionan sobre qué parámetros -cualitativos y cuantitativos- de evaluación hay que considerar y cómo definirlos respetando

la especificidad de los campos científicos, sus estilos de producción, formas de expresión y difusión de sus resultados. Además, algunos consideran que ese proceso tiene que ser liderado por los integrantes de las comisiones evaluadoras por áreas; mientras que para otros ese proceso debe ser resultado de una amplia discusión que involucre a todos, evaluadores y evaluados. Entre los asistentes se reconoce, por un lado, el descontento con los mecanismos actuales de evaluación y, por otro lado, que tanto la evaluación puramente cuantitativa como la puramente cualitativa tienen problemas y que lo ideal sería llegar a un modelo equilibrado, con criterios de evaluación cuantitativos y cualitativos complementarios entre sí.

2.2.7 Las exigencias de las evaluaciones

Para algunos asistentes, la evaluación tiene que estar más humanizada y menos “industrializada” la producción de conocimiento. Hay una deshumanización del sistema y una auto-explotación de los académicos que quizás no se percibe con total claridad, pero cada vez se exigen más actividades, de publicaciones -de hecho se mide *productividad*-, de clases, de alumnos, de tutorías de tesis, de proyectos, etc. es algo que se vuelve inalcanzable y se pierde gente valiosa por tener que trabajar bajo presión, en condiciones de estrés o por razones que van más allá de la producción.

A la vez o como contracara del mismo fenómeno, es decir, en la medida en que avanza la tendencia a contar, se agregan más actividades e incluso una misma actividad en diferentes ítems, y el resultado son CVs inmanejables e hipertrofiados que no reflejan necesariamente el desempeño real.

2.3. Consecuencias de la evaluación académica

2.3.1 Efectos sistémicos no deseados de la evaluación sobre la actividad académica

Entre los asistentes al taller se llamó la atención sobre la necesidad de reflexionar sobre la lógica de la evaluación y sus efectos sobre la actividad académica. Señalan que la evaluación se concentra en el desempeño de la investigación realizada por los individuos, pero pierde de vista que esos individuos realizan varias funciones o diversas actividades académicas.

Reconocen que publicar es importante pero resaltan que poner el énfasis sólo en eso no representa ni estimula la diversidad y conduce a que los individuos se alejen de la gestión, la construcción institucional, la enseñanza, la extensión y, además, la investigación se reduce a publicar cierta cantidad de artículos en ciertos períodos de tiempo, empobreciendo, con ello, la calidad de la producción de conocimiento nacional.

De no considerar equilibradamente todas esas funciones y/o actividades, se desestimula:

- i) la dedicación a la enseñanza y además se socava el impulso a enseñar creativamente, reforzando el atraso acumulado de cuatro o cinco décadas en prácticas de enseñanza y pedagogía. Peor aún, en la formación de los estudiantes *“se promueve el avanza y no pienses”*, se señala una participación de estudiantes cada vez menor en la actividad de enseñar y que las nuevas generaciones pierden de vista la carrera docente;
- ii) la dedicación a la construcción institucional. Este planteo emerge con más fuerza en el interior del país, donde se están creando instituciones e infraestructura (física y académica); y,
- iii) las tareas de gestión académica y de apoyo a la investigación.

Asimismo, los participantes del taller señalaron que la evaluación que prioriza la cantidad tiende a estimular el trabajo de investigación más automático y que esa señal no es deseable y menos para orientar el trabajo de los más jóvenes. El énfasis en la cantidad de artículos científicos publicados promueve la publicación de trabajos inconclusos, incorrectos, fragmentados y “cooperativas” -co-autorías- mal entendidas. También desestimula el desarrollo de nuevas líneas de investigación porque la probabilidad de obtener resultados rápidos y “seguros” es menor. Tampoco estimulan el abordaje de problemas de investigación que pueden ser útiles, pertinentes o relevantes para el país. Por el contrario, promueve prácticas de investigación orientadas por lo que es publicable. Lo que no es publicable rápidamente, queda de lado con independencia del desafío intelectual que implique. Así la producción de conocimiento tiende a concentrarse en procesos de investigación de corto plazo, y ya no se orienta por el avance de la ciencia, su relevancia y/o por su

pertinencia para la resolución de problemas -sociales, productivos, ambientales, energéticos, etc.-.

En ese sentido, plantean la necesidad de reflexionar sobre estos temas para cambiar el foco único de la evaluación y, también, tener en cuenta los períodos temporales de la investigación para la obtención de resultados originales, relevantes y pertinentes. También llaman la atención respecto a que las revistas internacionales tienen sus propios intereses en materia de producción de conocimiento y que están condicionando las agendas de investigación locales, sus temas, enfoques teóricos y metodológicos. Ven con preocupación que esos criterios en la evaluación “matan” las revistas científicas nacionales, regionales y las que no están en inglés; desestimula o perjudica a la *ciencia emergente*, a la interdisciplina y a todo lo que no se ajusta a los parámetros de desempeño del modelo hegemónico de evaluación.

Otros participantes del taller coincidieron en que las publicaciones en revistas internacionales no pueden ser la única forma de evaluación, pero señalan que sí han ejercido un efecto positivo en algunos contextos específicos. En particular en aquellos donde se consideraba que no era posible publicar y donde la máxima aspiración ascendía a las presentaciones en congresos. En la actualidad eso cambió. Se implementaron mecanismos que hicieron posible una mejora, un crecimiento de las publicaciones como forma de difundir el conocimiento que allí se produce. A la vez y como resultado de eso y de la mirada externa, también observan cambios positivos en las valoraciones sobre sus investigadores.

2.4. Alternativas para la evaluación académica

2.4.1 Tender hacia un modelo de evaluación académica integral, flexible y adaptable

En el taller los participantes llegaron a ciertos acuerdos, a saber: i) que no están satisfechos con la forma en que se está procesando la evaluación; ii) que hay que cambiarla; y iii) que el cambio debe ir en una dirección que contemple, en líneas generales, los siguientes atributos: integralidad, flexibilidad y adaptabilidad. Consideran que la evaluación debe estar orientada por objetivos

y que es una actividad que encierra apreciaciones esencialmente cualitativas. Sostienen que debe ser integral, valorando equilibradamente el conjunto de funciones desempeñadas en la Udelar, los productos concretos y los procesos.

Existe también acuerdo en que la integralidad de funciones requiere de mecanismos de evaluación integrales, y se cuestionan respecto a cómo evaluar todo lo que no es investigación y qué elementos deben cuantificarse y cuáles deben cualificarse. Como punto de partida, plantean que es probable que esos elementos varíen según el campo de conocimiento y las disciplinas, y en ese sentido marcan como importante determinar cuáles son las evidencias de la producción de conocimiento en cada campo y cómo se reflejan (expresan y difunden).

Adicionalmente y como líneas generales plantean que la evaluación tiene que considerar: i) las especificidades disciplinares, los objetivos de investigación y sus plazos; ii) las especificidades contextuales y los objetivos institucionales; y, iii) los grados de avance diferenciales en las trayectorias académicas.

La evaluación tiene que estar ajustada al contexto -incluyendo las características del mercado laboral y la capacidad de retener y ampliar las capacidades de investigación- y reconocer que en ciertos lugares no hay masa crítica suficiente y que los individuos no trabajan aislados de su entorno. A la vez, se resalta que las personas, los grupos y los servicios parten de situaciones o condiciones materiales diferentes y, en ese sentido, se requieren mecanismos de evaluación propios capaces de atender a esas especificidades. Para algunos, la CSIC debe liderar el proceso y desarrollar instrumentos de medida propios, ya que las soluciones de "talla única" no sirven. A la vez, los objetivos institucionales tienen que ser considerados en la evaluación, y la construcción institucional tiene que ser valorada, en particular a instancias del fortalecimiento y profundización de los procesos de descentralización de la Udelar.

También deben considerarse en la evaluación las especificidades disciplinares, como por ejemplo, en el campo de la salud, las tareas clínicas y de atención a pacientes, que insumen la mayor parte del tiempo. Simultáneamente, destacan que no todas las evaluaciones tienen que pasar por el mismo filtro y que los

puntos de evaluación no pueden ser rígidos, sino acompañar las variaciones existentes entre áreas de conocimiento: *“no se puede evaluar de igual forma a las ciencias sociales y a las ciencias básicas”*. Adicionalmente, la evaluación de los individuos tiene que considerar su punto de partida y analizar, en términos relativos a su trayectoria y condicionantes, a dónde logró llegar. Es decir, la evaluación debe valorar las oportunidades diferenciales que ha tenido el investigador.

En la medida que la investigación es una tarea cada vez más colectiva, la evaluación debe incorporar también la dinámica de los colectivos. En este sentido, sería deseable contar con una narrativa de los individuos sobre el funcionamiento de su colectivo de referencia, que explique su contexto, oportunidades y condicionantes para su desarrollo.

Asimismo, la evaluación tiene que considerar el impacto de las actividades académicas en ámbitos no académicos, tiene que dar tiempo suficiente para la producción científica de calidad, ambiciosa, relevante, original y pertinente para la atención de los principales problemas que afectan el desarrollo nacional, y no penalizar los procesos de investigación que no alcanzan resultados inmediatos y publicables. Es decir, debe tomar en cuenta las especificidades de los procesos de investigación, sus objetivos y atender a las eventuales disparidades territoriales, institucionales y grupales; además de considerar las formas en que se produce, acumula y difunde el conocimiento en cada campo. Se trata de encontrar mecanismos para que todos tengan incentivos para desarrollar sus actividades académicas.

Los participantes también destacaron que si bien hay que tener en cuenta todas esas especificidades, no se pueden aceptar soluciones simplistas porque eso tampoco contribuye al crecimiento y fortalecimiento de la comunidad, sus actividades y a las áreas de conocimiento. Es decir, no ven como algo deseable ni aceptable, bajar los estándares de excelencia –*“sí, hay que tener una vara alta y no ser autocomplacientes”*. Plantean que más allá de las especificidades hay elementos que son transversales y comunes a todos, cuyos estándares de excelencia deben mantenerse para todos, y que la referencia internacional y la mirada externa son, también, aspectos muy importantes.

2.4.2 Evaluación académica: definir sus objetivos, hacer de ella una actividad más transparente y generar mecanismos para evaluarla continuamente

Entre los participantes un tema recurrente fue la necesidad de esclarecer y definir cuáles son los objetivos de la evaluación, preguntándose para qué evaluar a los académicos y qué sentido tiene tanto para la Udelar como para el país. Plantearon que la evaluación debe ser considerada como un instrumento o un medio para generar diversidad y a la vez aumentar y fortalecer las capacidades académicas locales, mejorar la calidad de los procesos de producción de conocimiento, de enseñanza, de relacionamiento con el medio, estimular el aprendizaje y mejora continua de individuos, grupos e instituciones. Destacan que dependiendo de los objetivos de la evaluación deberán definirse sus criterios, y que no es lo mismo la evaluación de un proyecto, grupo, individuo, o si se trata de acceso o ascenso en la carrera docente. Además, resaltan que si el objetivo fundamental es contribuir al aprendizaje y a mejorar la calidad y el desempeño, entonces las evaluaciones tienen que generar devoluciones para los evaluados con recomendaciones y orientaciones que contribuyan a ese objetivo, y no solamente con comentarios del estilo “*publique más*”. Sugieren que una buena práctica a seguir es la que implementa la CSIC a través de sus devoluciones -derivadas de la evaluación de propuestas de investigación- a los investigadores.

Adicionalmente, algunos resaltaron la necesidad de transparentar los procesos de evaluación. Es decir, mejorar la comunicación sobre los criterios para determinar quiénes son los evaluadores, cómo se seleccionan y qué criterios de evaluación se utilizan; tanto para evaluadores como para los evaluados. De acuerdo con ellos, las señales o reglas del juego tienen que ser claras para todos porque la mala comunicación de los incentivos y heterogeneidad es un problema que favorece la discrecionalidad y la endogamia. Complementariamente, consideran como una práctica deseable la realización de declaraciones de conflicto de interés por parte de los potenciales evaluadores y la generación de instancias/mecanismos para que los evaluados puedan hacer los descargos que entiendan pertinentes ante evaluaciones que consideran no satisfactorias.

Además de lo anterior, reconocen que la evaluación de calidad requiere de tiempo, trabajo calificado y los mejores insumos posibles. Respecto al trabajo calificado, consideran que hay que ampliar el conjunto de evaluadores, involucrando también a los evaluados e incluso a los estudiantes - como veedores en los concursos y de garantía ética-, se trata de distribuir las tareas para disminuir la prevalencia de los grupos de poder que dominan la evaluación. A la vez entienden necesario instancias para formar a los evaluadores y hacer énfasis -en esa formación- con enfoque interdisciplinario, para que puedan desarrollar mejores evaluaciones. Atendiendo a los insumos para la evaluación, llaman la atención sobre los currículums vitae, destacando que son demasiado extensos -porque repiten información- y dificultan la tarea de evaluar. Consideran que hay que dar instrucciones claras y precisas que contribuyan al desarrollo de CVs acotados y relevantes, con formatos bien pensados.

Complementariamente, entienden necesario desarrollar mecanismos para evaluar los procesos de evaluación, como actividad sistemática y regular. Aquí también, sostienen que la CSIC debe asumir la tarea de sistematizar el estado del arte en el tema, y agregan que la revisión de las formas de evaluación debería ser parte de una “discusión perpetua”, de la mano con un cambio progresivo, atado a objetivos institucionales. Cada servicio debería hacer esto a la interna. Finalmente proponen líneas presupuestales dedicadas a la evaluación.

2.4.3 Simplificar y unificar sistemas de evaluaciones

Los asistentes al taller llaman a reflexionar sobre la mejora del sistema de evaluaciones, consideran que el sistema actual refleja y alimenta las relaciones de poder dominantes. Destacan la necesidad de buscar mecanismos que trasciendan la publicación en revistas arbitradas indexadas a ciertas bases de datos internacionales. Resaltan la importancia de ser cuidadosos en el diseño de políticas e instrumentos de evaluación por sus efectos que pueden conducir, de no cambiarse, a empobrecer la actividad académica, su integralidad y calidad; e insisten en la necesidad de una mayor alineación/coordiación entre los objetivos de la Udelar y los incentivos del Sistema Nacional de Investigadores.

Plantean procesos de evaluación continuos y transversales como alternativas a las múltiples evaluaciones puntuales. También proponen restringir las instancias de evaluación con períodos de tiempo más largos entre evaluaciones. Consideran importante unificar las evaluaciones del RDT con las del cargo docente, y llaman a retomar e implementar los lineamientos de evaluación plasmados en el documento de CSIC sobre la evaluación del RDT (Acceda al texto en el siguiente [link](#)). Enfatizan la necesidad de que la evaluación académica ponga el foco en la integralidad de las funciones, hay que evaluar profesores no solo investigadores -que es una condición necesaria pero no suficiente- y eso no se puede perder de vista. Además destacan que la discusión sobre la evaluación académica tiene que estar conectada con el nuevo Estatuto de Personal Docente y su paulatina implementación.

2.4.4 Propuestas sobre los métodos de evaluación

Los asistentes al taller coincidieron en que se debe alcanzar un conjunto equilibrado y complementario de criterios cualitativos y cuantitativos para la evaluación académica.

Respecto a los criterios cuantitativos, no desestiman el uso de indicadores bibliométricos como medidas de comparación e incidencia internacional. Proponen analizar las ventajas de las métricas y alertan que la calidad tiene que estar por encima de la cantidad y que el uso de esos indicadores tiene que estar ponderado en función de los objetivos institucionales y de investigación, de las prácticas y especificidades de producción de conocimiento en los distintos campos (tipos de productos y formas de difusión). La contabilización de artículos científicos por sí sola no da cuenta de la calidad ni de la diversidad de la producción de conocimiento, y proponen la lectura de los artículos -al menos de los tres o cinco más relevantes indicados por el evaluado-. La cantidad de citas tampoco es una buena aproximación de la calidad, hay que descontar las auto-citas y apreciar cualitativamente la calidad de la cita (el contexto de la cita). Asimismo, algunos señalan que existen realidades muy heterogéneas y que, en los lugares con menor desarrollo relativo, promover las publicaciones en revistas arbitradas puede tener impactos muy positivos en términos de aprendizaje; resaltan la importancia de estimular el avance y

desarrollo de los lugares que cuentan con procesos de investigación más incipientes.

También se resaltó que en el ámbito institucional se debe trabajar para que las revistas científicas nacionales se categoricen y priorizar la evaluación de las revistas que publica la Udelar y que la revisión de los artículos siga ciertos criterios, es decir, diseñar una política específica orientada a mejorar la calidad de las revistas científicas de la Udelar. Sostienen que tiene que haber más canales de publicación y dicen: *“nosotros queremos investigadores que publiquen fuera, pero también queremos investigadores que publiquen dentro. Debemos priorizar la evaluación de nuestras revistas”*.

Respecto a los criterios cualitativos, si bien preocupa su falta de legitimidad y sus costos en términos de tiempo y recursos, consideran que la apreciación cualitativa de los méritos intrínsecos de la actividad académica es la mejor manera para contribuir al fortalecimiento y crecimiento de la comunidad académica nacional. Proponen superar su falta de legitimidad ampliando el conjunto de evaluadores, generando y difundiendo ampliamente las pautas y criterios de evaluación, exigiendo declaraciones de conflicto de interés y generando mecanismos para que los evaluados apelen los resultados de las evaluaciones. Algunos consideran que deben definirse ciertos estándares y proponen cuatro parámetros: formación personal, participación en proyectos, formación de otros, publicación de resultados. La generación de acuerdos de cooperación y convenios con organizaciones extra-universitarias públicas, privadas o de la sociedad civil tiene que ser considerada y valorada. Consideran que es importante la evaluación externa para evitar problemas de endogamia y prácticas perversas, y reafirman la revisión por pares directa. A la vez, destacan la necesidad de desarrollar un sistema de registro de la actividad académica de la Udelar.

Con base en la experiencia de algunas Facultades proponen generar mecanismos de autoevaluación que contengan ciertas narrativas como insumos adicionales para la toma de decisiones de los evaluadores. Concretamente, proponen realizar una prueba experimental con los docentes en el RDT y que todo DT deba hacer al menos una evaluación en el año. Pasos que podría tener una nueva forma de evaluar el desempeño académico: i)

autoevaluación y narrativa; ii) evaluación de un par sobre el desempeño; iii) evaluación de otro par sobre la consistencia de la evaluación realizada por el evaluador. Considerar la narrativa del investigador sobre su trayectoria (significación de su trabajo). Para ello pedirle al evaluado que relate la relación entre las tesis que dirigió y los productos que obtuvo; entre los proyectos y sus productos, los cursos en los que participó y el material docente que elaboró, etc. La evaluación se realiza en distintos momentos de la vida de la persona y los instrumentos para evaluar deben ser diferentes. En la primera etapa hay que tener en cuenta factores de tipo personal que debe tener la persona, y su capacidad de relacionarse y de trabajar en grupo. Incorporar en la evaluación individual una narrativa sobre cómo funciona la investigación en su colectivo de referencia. Al evaluar sus publicaciones, solicitar las cinco mejores y pedir al investigador fundamentos. Considerar la formación de recursos humanos: “preocupación por la descendencia” y por otras actividades de enseñanza, y promover activamente la formación de investigadores autónomos. La evaluación tiene que considerar la temporalidad de la producción como parte del proyecto académico del docente, valorar el conjunto de actividades desempeñadas desde la evaluación anterior - *“qué hizo en 5 años”*-