

**Evaluar es necesario: Encuentros interinstitucionales sobre prácticas, tradiciones
y necesarias renovaciones en la evaluación académica
Primera parte**

Coloquio sobre propósito y desafíos de la evaluación

Relatoría¹

Montevideo, 7 de marzo 2024

Prorectorado de Investigación

Unidad Académica de CSIC y Núcleo Interdisciplinario CiTINDe

Universidad de la República

¹ Realizada por Natalia Gras, Profesora Adjunta de la UA de CSIC y Co-Coordinadora de CiTINDe, Universidad de la República

Esta relatoría no da cuenta lineal de la desgrabación y los apuntes tomados durante el Coloquio. Está organizada, como se desprende del índice, en torno a las grandes temáticas que fueron abordadas en el conjunto de las intervenciones.

INDICE

Introducción	3
1. La relevancia, complejidad y no neutralidad de la evaluación académica	4
2. Malestares que impulsan transformaciones en la evaluación académica	5
3. Los orígenes de esos malestares y sus consecuencias	6
3.1 Argumentos para la transformación de los sistemas de evaluación académica	7
3.2 Sobre las consecuencias en las prácticas académicas y calidad de vida de las personas	8
4. Propósitos de la evaluación	9
5. Sugerencias para la transformación de la evaluación.....	11
5.1 Sobre como evaluar la actividad académica y el papel de lo cuantitativo	11
6. Implementación de las transformaciones	12
7. Hacia una evaluación académica cualitativa, integral, plural y más humana: algunos desafíos.....	15
7.1 Desafíos para la evaluación cualitativa	15
7.1.1 El problema de la masa crítica y el volumen de la evaluación.....	15
7.1.2 El problema del “sobrecalentamiento” del sistema de evaluación	15
7.1.3 El problema de la discrecionalidad y la subjetividad de la evaluación	16
7.2 Desafíos para la evaluación integral.....	17
7.3 Desafíos para la evaluación plural	18
7.4 Desafíos para la evaluación más humana y con perspectiva de género	18

Introducción

Entre la comunidad académica local existen preocupaciones comunes en torno a diversos aspectos vinculados con la evaluación que llaman a la creación de un espacio de diálogo e intercambio interinstitucional que ponga en común prácticas y tradiciones y que contribuya a pensar alternativas para la evaluación académica en Uruguay. La necesidad de repensar los sistemas actuales de evaluación es compartida internacionalmente, tanto en el plano de la reflexión como en las prácticas.

En ese contexto, el Prorectorado de Investigación de la Universidad de la República (Udelar) junto a la Unidad Académica de la Comisión Sectorial de Investigación Científica (CSIC) y el Núcleo Ciencia, Tecnología e Innovación para un Nuevo Desarrollo (CiTINDe) del Espacio Interdisciplinario organiza el ciclo *“Evaluar es necesario: encuentros sobre prácticas, tradiciones y necesarias renovaciones en la evaluación académica”*.

El ciclo consta de tres encuentros interinstitucionales y el primero *“Coloquio sobre propósitos y desafíos de la evaluación”* fue el pasado 7 de marzo en el Campus Luisi Janicki: pioneras universitarias de la Udelar. En él participaron 70 personas pertenecientes a diversos ámbitos del Sistema de Ciencia, Tecnología e Innovación (SCTI)² del país, quienes intercambiaron y reflexionaron sobre para qué evaluar, qué mejoras son necesarias en la evaluación y los desafíos que eso supone.

El coloquio se desarrolló durante toda una mañana que comenzó con la apertura a cargo del Prof. Rodrigo Arim, Rector de la Udelar, que destacó el papel fundamental de la evaluación en la vida académica y en la creación de conocimiento, advirtiendo sobre sus implicancias y planteando la importancia de repensarla y transformarla, sin desconocer la complejidad y los desafíos que eso impone. Continuó con una bienvenida realizada por la Profa. Virginia Bertolotti, Prorectora de Investigación, que incluyó la presentación de los objetivos del ciclo, su fundamento y relevancia. Posteriormente, cerrando esta primera ronda de intervenciones, la Profa. Judith Sutz, integrante de CiTINDe, realizó una exposición sobre las transformaciones de la evaluación y su puesta práctica en distintas partes del mundo. Finalmente, y ocupando la mayor parte del coloquio, tuvo lugar un diálogo abierto entre todas las personas asistentes, que permitieron identificar focos de coincidencia acerca de los propósitos y mejoras necesarias en materia de evaluación académica en Uruguay; y los desafíos que eso impone.

Esta relatoría se realiza con base en las intervenciones de las personas asistentes, las notas tomadas por Melissa Ardanche³, Natalia Gras, Judith Sutz y Mariana Vaz⁴ y la desgrabación del audio de la actividad, revisada por Claudia Cohanoff⁵.

Se estructura en siete secciones que recogen los temas e intercambios que se desarrollaron tal como fueron discutidos en el coloquio.

² Agencia Nacional de Investigación e Innovación (ANII), Sistema Nacional de Investigadores (SNI), Instituto Pasteur de Montevideo (IPMon), Instituto de Investigaciones Biológicas Clemente Estable (IIBCE), Programa de Desarrollo de las Ciencias Básicas (PEDECIBA), Instituto Nacional de Investigación Agropecuaria (INIA), Universidad Tecnológica del Uruguay (UTEC), Asociación Uruguaya de Posgraduandos y Posgraduados (AUPP), Asociación de Docentes de la Universidad de la República (ADUR), Investiga Uy y la Udelar a través de sus Comisiones Centrales de Asuntos Docentes y de Dedicación Total y los Prorectorados de Enseñanza y Extensión y Relaciónamiento con el medio.

³ Docente de la Unidad Académica de CSIC, Udelar

⁴ Integrante de CiTINDe, Espacio Interdisciplinario, Udelar

⁵ Docente de la Unidad Académica de CSIC, Udelar

1. La relevancia, complejidad y no neutralidad de la evaluación académica

Diversas intervenciones hicieron notar la importancia de la evaluación académica y de reflexionar en torno a ella debido a su complejidad e implicancias en lo académico, lo institucional, lo colectivo y lo individual.

El rector de la Udelar planteaba, como primer punto de su intervención, la importancia de la evaluación en los ámbitos universitarios y académicos en general y en el contexto actual del Uruguay y el grado de desarrollo relativo de sus sistemas de creación de conocimiento, cuya funcionalidad social o compromiso es generar conocimiento e incrementar el acervo de lo que la humanidad sabe en un momento del tiempo. Resaltaba que desde mediados del SXIX la evaluación es un componente inseparable de la vida universitaria y de la creación de conocimiento, que no hay instituciones ni sistemas de creación de conocimiento sin algunos mecanismos de evaluación más o menos formalizados a lo largo de la historia, y que en las últimas décadas eso se ha rigidizado bastante.

Asimismo, varias de las intervenciones pusieron de relieve que la evaluación académica debía entenderse como un proceso social complejo que involucra a personas e instituciones con objetivos e intereses diversos, sus tensiones y contradicciones.

Un integrante del Grupo de Trabajo de ADUR sobre el Título II del Estatuto del Personal Docente (EPD), referido fundamentalmente al Régimen de Dedicación Total, subrayaba que la normativa para evaluar -y su eventual transformación- como herramienta para evaluar mejor es una parte de lo que en realidad es la evaluación, pero hay otra dimensión que no se puede desconocer cuando se piensa en transformar y es que la realizan personas que forman parte de una misma comunidad -evalúan y son evaluadas- con sus sesgos, prejuicios, valores, creencias, etc.. Esto puede poner en cuestión la aplicación de los criterios de evaluación incluso estando éstos explícitamente normados. Resaltaba así que no alcanza con que las instituciones “decreten” nuevos procesos de evaluación, siendo preciso transformar las prácticas asociadas a ella. En cierta sintonía, una persona adscripta al INIA señalaba que, aunque existe mucha literatura sobre la evaluación, sus problemas y sus posibles mejoras, es bastante más complejo implementar los cambios para mejorar la evaluación y superar sus principales problemas. Esto último se debe a que hay que tener presente tres planos que delimitan un espacio de conflictos: lo que se dice, lo que se hace, y lo que se desea. Así, aunque es importante reconocer “el amor” como componente que motiva y moviliza a las personas en su compromiso con la sociedad, no puede desconocerse que también se trata de una disputa por dinero, poder y reconocimiento.

Sumando a dicha complejidad, el rector introdujo como aspecto central de la discusión, la no neutralidad de la evaluación. Señalaba que los sistemas de evaluación –de forma explícita o implícita- generan una estructura de incentivos –premios y castigos- que indica a las instituciones, colectivos y personas qué es deseable que hagan. En definitiva, se trata de una priorización sobre qué es lo deseable, y, eso a su vez, tiene derivas institucionales. En particular, resaltaba la incidencia directa en los procesos de construcción institucional que se puedan estar generando y lo explicaba así: “(...) *en la medida en que hay un sistema que estructura premios y castigos, también estructura jerarquías. Entonces, al interior de los sistemas de ciencia, de tecnología, de educación superior, de creación cultural, y en la medida en que la evaluación señala qué es lo deseable o qué es lo más deseable que suceda en términos de generación de conocimiento, también va a señalar y ordenar a instituciones, grupos y personas, en función de esos criterios.*”. En ese sentido, destacaba también un riesgo importante, en particular en el contexto reciente de expansión y generalización de la educación superior, en cuanto a que lo deseable en términos de diversificación de los sistemas de educación superior, se transforme en algo no deseable: la estratificación de dichos sistemas.

2. Malestares que impulsan transformaciones en la evaluación académica

La presentación realizada por Judith Sutz, con base en lo que está sucediendo en el mundo respecto a la discusión sobre la evaluación académica, comenzaba señalando que hay malestar con la evaluación y a efectos de revertir esa situación, diversas universidades, agencias de financiamiento, academias de ciencias y agrupaciones de países de distintas partes del mundo están repensando sus sistemas de evaluación e implementando cambios. La necesidad de repensar los sistemas actuales de evaluación es compartida internacionalmente, tanto en el plano de la reflexión como en el plano de las prácticas⁶.

Asimismo, planteaba que las principales preocupaciones de las instituciones internacionales que impulsan transformaciones en la evaluación académica son: i) aumento de malas prácticas académicas; ii) insatisfacción con la calidad de vida en el trabajo; iii) disminución de la calidad, integridad, reproducibilidad, relevancia conceptual, pertinencia social, originalidad y visión de largo plazo de los esfuerzos académicos; iv) el descuido de actividades fundamentales como la formación, el relacionamiento con actores no académicos y la construcción institucional; y, para el caso exclusivo de los países desarrollados, vi) la desafección de jóvenes de la academia y migración hacia otros espacios laborales que ofrecen posibilidades de seguir siendo investigadores fuera de la academia.

Adicionalmente, señalaba que los resultados obtenidos con la encuesta *Consulta Investigan* (2022) realizada en el marco de CiTINDe, muestran un claro alineamiento de lo que piensan investigadores e investigadoras en el Uruguay con las preocupaciones que se observan a nivel internacional: i) el 81,3% coincide en que el Sistema Nacional de Investigadores (SNI) influye en la planificación de su trabajo; ii) el 72,9% considera que la evaluación actual genera malas prácticas y el 86,1% que hay que cambiar el sistema para evitarlo; iii) el 80,3% percibe que dificulta el relacionamiento con actores no académicos y atender sus problemas y el 82,9 que hay que cambiar para revertirlo; iv) el 61,3% indica que la evaluación le genera estrés y ansiedad; y, además, v) el 67,3% está de acuerdo con que si la evaluación fuera más cualitativa se propondría ser más innovador en el conjunto de las otras actividades académicas que realizan (por ejemplo: docencia, relacionamiento con el medio, etc.).

Diversas intervenciones del público presente aportaron sus visiones sobre estas cuestiones, complementando y/o reforzando las preocupaciones y/o malestares con las formas actuales de evaluación ya señaladas y argumentando la necesidad de cambio en la evaluación.

De acuerdo con la visión de una persona integrante de una Comisión Técnica Asesora (CTA) del SNI, otra problemática de los sistemas de evaluación tiene que ver con que en su mayoría se han centrado en las personas individuales, desconociendo que su desempeño también es resultado de un proyecto de construcción colectiva. En ese sentido, destaca la importancia de pensar los cambios de la evaluación actual para fortalecer también ese proyecto de construcción colectiva y, tal vez, transitar a la evaluación de colectivos. Vinculado a esto, una persona adscripta al IIBCE, ponía de relieve la tensión que generan los sistemas de evaluación entre el trabajo colaborativo y la exigencia de independencia en referencia a las evaluaciones de PEDECIBA y del SNI (la discusión sobre múltiples autores y su ordenamiento). Entiende que, entre las transformaciones a los sistemas de evaluación, debe haber cambios que contribuyan a eliminar la tensión entre lo individual y lo colaborativo.

⁶ Para una revisión de los debates sobre evaluación académica, propuestas de cambios y transformaciones implementadas ver GYA-IAP-ISC (2023) The Future of Research Evaluation: A Synthesis of Current Debates and Developments. DOI: 10.24948/2023.06 Disponible en <https://www.interacademies.org/sites/default/files/2023-05/2023-05-11%2BEvaluation%2B-%2BWEB.pdf>

Por otra parte, esta persona integrante de una CTA del SNI, cuestionaba el poco o nulo espacio que dejan los sistemas de evaluación para la formación propia de investigadores e investigadoras institucionalizadas y la actualización permanente, ya sea para iniciar o avanzar con nuevas líneas o por asumir riesgos con temas difíciles o de largo aliento. A partir de allí considera central para la calidad académica romper con la idea productivista de los sistemas de evaluación y jerarquizar la formación y actualización permanente de los y las investigadoras.

Desde el IIBCE se enfatizaba la importancia de que el cambio en la evaluación considere y valore la contribución a la sociedad, comprendiendo sus aportes y la diversidad de productos que genera que no son exclusivamente ni mayormente “papers”. Reforzando el argumento, señalaba la dificultad de colegas del IIBCE que, teniendo aportes significativos para la sociedad, quedan excluidos del SNI por no alcanzar los estándares de publicación de ese sistema y por lo tanto también del acceso a fuentes de financiamiento de la investigación de la ANII⁷.

La intervención que se hacía desde el INIA afirmaba los argumentos anteriores y los complementaba indicando que en la evaluación es preciso jerarquizar: i) la enseñanza y la formación; ii) la vinculación con las políticas públicas y los actores sociales y sus problemáticas; iii) las contribuciones a la innovación empresarial y social; iv) la comunicación y divulgación de los conocimientos, debido a la importancia de la comprensión pública de la ciudadanía de diversos temas de la realidad nacional; y, v) la construcción institucional. Todas estas son actividades importantes que insumen tiempo y que no se reflejan en publicaciones o patentes, pero que son fundamentales para fortalecer el Sistema de CTI y sus instituciones. Señalaba el “eterno problema del equipo o el individuo” y coincidía en que quizás había que avanzar en la evaluación de colectivos y, también, en contextualizar la evaluación.

3. Los orígenes de esos malestares y sus consecuencias

En su presentación, Sutz postulaba que en el origen de los problemas se encuentra el productivismo, en particular, “más artículos es mejor”, así como la aproximación a la calidad por métricas.

En sintonía, el rector apuntaba que, desde la década de 1980, las métricas y las evaluaciones se han convertido en un fin en sí mismo, más que una herramienta para transformar prácticas o mejorar lo que hacen las personas, los colectivos y las instituciones. Ejemplificaba con el problema de una métrica en particular: los rankings de universidades, la percepción de prestigio –“espurio”- que producen y las distorsiones en el funcionamiento institucional que generan. Esto lo vincula con la emergencia de las métricas y la preponderancia de mecanismos de evaluación cuantitativos y rígidos que restringen las autonomías institucionales y limitan su avance en ciertas direcciones, con implicancias que pueden ser muy negativas para el desarrollo institucional. Por ejemplo, el proceso de expansión y profundización de la Udelar en todo el territorio nacional es fenomenal con impactos muy importantes y deseables para el desarrollo del país, de su sistema de educación superior y su sistema de CTI. Sin embargo, las métricas utilizadas para evaluar el desempeño valoran mal ese esfuerzo de la Universidad, haciéndole perder posiciones relativas en el ranking y su consecuente reducción de “prestigio”. Eso podría haber inducido a re-direccionar la política en un giro conservador, centralizador y concentrador y con ello, truncan un proceso de desarrollo necesario y deseable, ya no para la Universidad, sino para el país. Subraya, además, que eso es válido para todos los sistemas de evaluación, incluyendo la evaluación de individuos, y plantea la importancia de que dichos sistemas se entiendan en su contexto, porque ciertos mecanismos de evaluación pueden ser apropiados en un determinado momento, pero dejar de

⁷ Aunque pertenecer al SNI no es un requisito explícito -formalmente solicitado- para acceder a recursos en fondos concursables de ANII, es importante señalar que las modalidades dirigidas a investigadores consolidados, de los Fondos Clemente Estable y María Viñas, se indica que la definición de “investigadores consolidados” utilizada se equipara con la establecida en los requisitos del SNI en los niveles I, II y III (Gras y Barletta, en revisión).

serlo en otro.

3.1 Argumentos para la transformación de los sistemas de evaluación académica

La exposición de Sutz, sintetizaba algunos argumentos esgrimidos por las instituciones internacionales (International Research Council y Welcome Trust) que impulsan transformaciones. Por un lado, lo ya adelantado por el rector: “la mayoría de las instituciones pone una atención incesante en lo individual, presionando a todos para que marquen casillas y se ajusten”. Por otro, planteaba que “el uso de indicadores bibliométricos como *proxy* de desempeño es conveniente, pero profundamente defectuoso”. En relación a esto, resultan particularmente elocuentes las siguientes intervenciones del público asistente:

- a) “(...) nuestro trabajo como evaluadores es de peor calidad que nuestro trabajo como investigadores. Aceptamos reglas del juego que no aceptaríamos en nuestro trabajo de investigación, como confundir datos brutos con datos analizados. Si vas a contar el número de artículos, el número de citas o el coeficiente *H* o cualquier otro indicador, se toman esos datos, se procesan y se analizan. No usaríamos esos datos en nuestras investigaciones sin ese tratamiento como sí los usamos en los procesos de evaluación. ¿Por qué? Porque el volumen de trabajo de evaluador es enorme y diría hasta desproporcionado.” y,
- b) “(...) está esta idea de que estamos evaluando mal, y sí, caemos en cierto cuantitativismo, en cierto burocratismo, en ciertas prácticas formales y no estamos haciendo el trabajo que realmente sería aspirable hacer. (...) es raro, porque hay veces que uno siente como que tiene un divorcio mental, digamos, usa palabras que no saben muy bien qué son: “excelencia académica” y realmente no tenemos idea de qué es. Son conceptos vacíos que incluso en términos colectivos son muy difíciles de consensuar. (...) Capaz que en un momento fue un lenguaje que nos ayudó a crecer como comunidad académica, pero hoy es preciso repensarlo”⁸.

Asimismo, fueron citados argumentos del Welcome Trust: “Queremos ayudar a construir una mejor cultura de investigación, una que sea creativa, inclusiva y honesta, las prácticas actuales priorizan los resultados a casi cualquier costo. Esto está dañando el bienestar de las personas y socavando la calidad de la investigación. Todos podemos ayudar a re-imaginar cómo se lleva a cabo la investigación”. Señalaba además que lo que preocupa particularmente a las instituciones internacionales que promueven cambios se vincula con la situación de: i) los jóvenes; ii) las mujeres; y, iii) la interdisciplina.

Sobre estos aspectos distintas intervenciones, realizaron múltiples comentarios. Uno de ellos enfatizaba que no solamente los jóvenes, las mujeres y el trabajo interdisciplinario quedan peor valorados o invisibilizados con los estándares actuales de los sistemas de evaluación, sino que también les ocurre, en sentido amplio, a las “las periferias”. Las hay al interior del país, de los campos de conocimiento, de las instituciones. Esas diferencias en el desarrollo relativo, se indica, deben ser comprendidas en su contexto, dando lugar a la promoción y no a la exclusión mediante la imposición de estándares de evaluación homogeneizantes de los campos o áreas de mayor desarrollo relativo del país. En esta línea argumentativa, diversas intervenciones plantearon la necesidad de reconocer la heterogeneidad existente en el país y promover la diversidad; así, el cambio de la evaluación debería ir en una dirección que contemple mayor reflexividad, contextualización y pluralismo.

La intervención de la AUPP cuestionaba el excesivo cuantitativismo de los sistemas de evaluación, su carácter jerárquico –“evaluación de arriba-abajo o como mucho entre pares, pero

⁸ Sobre la necesidad de repensar el concepto de “excelencia” ver Kraemer-Mbula, Erika, Tijssen, Robert, Wallace, Matthew & McLean, Robert (2021), Transformando la excelencia en la investigación. Nuevas ideas del Sur Global (Primera edición en español). Bogotá: Universidad del Rosario, African Minds, Disponible en <https://www.africanminds.co.za/transforming-research-excellence-new-ideas-from-the-global-south/>

sin o muy poco espacio de abajo-arriba”-, y la nula o escasa valoración de la diversidad de actividades académicas que se deben desempeñar. Hacía notar que como colectivo se encuentran en la parte del sector académico más precarizado, que está bajo las presiones más fuertes de los sistemas de evaluación y, también, de la incapacidad del sistema de CTI de absorberlos, resaltando así que entre ellas y ellos estarán una buena parte de quienes queden excluidos de dicho sistema. Asimismo, agregaba que es importante que los sistemas de evaluación consideren todos esos aspectos contextuales y también los ciclos vitales de las personas –“no olvidar que por jóvenes nos estamos refiriendo a personas que superan los 30 – 35 años y hasta 40 años en no pocos casos”-, lo que se agrava aún más en el caso de las mujeres; se plantea también la necesidad de cuidar la salud mental de las personas.

Por su parte, la CH del SNI indicaba que, en lo posible y hasta cierto punto, en las evaluaciones del sistema siempre se trata de tener en cuenta las circunstancias vitales de las personas. Describía los mecanismos regulados e instrumentados en el SNI, tales como las prórrogas por maternidad y enfermedad en el período objeto de evaluación. Señalaba que además de esos casos, que están contemplados en su Reglamento, también se tiene en cuenta las circunstancias en sentido amplio: i) los medios con que cuentan los y las investigadoras; ii) la institución o lugar donde están insertos para llevar adelante su actividad; iii) la disciplina en la que trabajan –“porque hay disciplinas que requieren de más medios que otras”-. Resaltaba que en lo posible y hasta cierto punto se contemplaban todas estas cuestiones que, en definitiva, hacen parte de evaluaciones más cualitativas.

3.2 Sobre las consecuencias en las prácticas académicas y calidad de vida de las personas

La presentación realizada por Sutz, mostraba la gran importancia que le asigna la literatura internacional al estudio de las consecuencias sobre las prácticas académicas y calidad de vida de las personas. Ejemplificaba con base en la realización de diversas encuestas de opinión que exploran estos asuntos en Flandes, Noruega, Polonia, Inglaterra, Austria entre muchas otras. Recuperaba los principales resultados de la conducida por el Wellcome Trust, debido a que es una de las más comprensivas, ya que abarcó a más de 4000 investigadores e investigadoras de diversas partes del mundo y cuenta, además, con cerca de 100 entrevistas en profundidad en el Reino Unido y discusión en talleres sobre alternativas para la evaluación. En el resumen ejecutivo del informe cualitativo se dice: *“Las largas jornadas de trabajo, la presión por encontrar resultados positivos, la falta de seguridad laboral y los intensos niveles de competencia crean las condiciones propicias para el estrés, la ansiedad, la mala salud mental, el acoso y comportamiento agresivo. Muchos hablaron sobre el impacto de esto no sólo en su propio bienestar sino también en el de sus familias”*⁹.

En sintonía, la Prorectora de Investigación de la Udelar en su intervención ponía de relieve este aspecto y señalaba: *“es una preocupación que el Prorectorado de Investigación ha tenido siempre y que a mí me preocupa particularmente, entre otras vinculadas a la evaluación: los sistemas de evaluación repercuten fuertemente en la calidad de vida y en la efectividad de los investigadores (...) los investigadores en Uruguay estamos comenzando a perder calidad de vida”*.

La intervención de la AUPP profundiza en la magnitud del problema derivado del escaso cuidado a la salud mental de las personas y el papel que en ello juegan los sistemas de evaluación: *“como posgraduandas y posgraduandos nos llegan múltiples situaciones, pero en general que ya están prendidas a fuego y ante lo cual es muy difícil responder apropiadamente. Entre otras cosas porque no tenemos espacios institucionalizados a dónde recurrir, más allá de que en todas las*

⁹ Wellcome Trust: Research Culture - Qualitative Report Disponible en <https://cms.wellcome.org/sites/default/files/what-researchers-think-about-the-culture-they-work-in-qualitative-research.pdf>

Sobre los resultados cuantitativos ver Wellcome Trust: Research Culture – Quantitative phase Disponible en <https://cms.wellcome.org/sites/default/files/what-researchers-think-about-the-culture-they-work-in-quantitative-research.pdf>

organizaciones e instituciones a las que hemos acudido y empezamos a negociar hay muy buena disposición, pero termina siendo siempre una disposición individual y no se avanza en una respuesta institucional con la preparación y el desarrollo de las capacidades necesarias para afrontar este tipo de problemas, que ya sabemos que son muy graves y complejos. Además, las evaluaciones son parte de este problema porque nos están generando muchísimo estrés, muchísima incertidumbre, en un contexto de mucha precarización, que nos sentimos afortunadas si tenemos una beca, que también supone mucho esfuerzo y trabajo obtener; y si bien hemos logrado que en algunos casos se reconozcan derechos básicos laborales, en muchos otros y para la mayoría de nosotros no los tenemos garantizados”.

Reforzando lo ya mencionado sobre la tensión entre “lo individual y lo colectivo” y su relación con el requisito de trabajo “independiente” versus el “colaborativo”, el intercambio en el coloquio aporta sobre la relación entre esas cuestiones y la competencia que generan los sistemas actuales de evaluación y sus efectos en el estrés y la ansiedad de las personas y en sus vínculos interpersonales.

Así lo planteaba una persona integrante del GT Título II del EPD de ADUR: “(...) hay un aspecto del estrés que sí podemos abordar y que quizás todos conocemos: (hay) climas que se establecen al interior de grupos de trabajo donde se trabaja colectivamente, donde digamos todos ayudamos a todos y hay otros climas donde nos pisamos la cabeza para salir adelante. (...) las señales [que emite la evaluación de las personas] pueden disminuir ese estrés o no. (...) aquí hay un aspecto colectivo y creo que este es un asunto central de la cuestión que sí podemos abordar. (...) Cuando te evalúan para el cargo, la renovación, el SNI o la DT, en ningún caso hay una evaluación del colectivo, del grupo, por más que exista (...) existen los grupos de investigación [financiados por CSIC], pero en realidad son más que nada - y lo voy a decir de una manera muy dura- mecanismos para que las personas progresen (...) esa evaluación no se considera en ninguna de esas instancias. Creo que ese es el punto quizás donde podemos ver cómo hacerlo”.

4. Propósitos de la evaluación

En relación a la pregunta sobre ¿para qué evaluar?, la presentación de Sutz apuntaba tres hitos que pueden identificarse como puntos de inflexión:

- i) hasta la década de 1980 para asegurar la calidad principalmente de la investigación a través fundamentalmente de la evaluación por pares;
- ii) desde ese momento y con la preeminencia del *New Public Management*, para rendir cuentas y auditar a quienes investigan y a las instituciones en que trabajan; y,
- iii) más recientemente -últimos cinco años- para promover:
 - a. la calidad de la actividad académica en sí;
 - b. la relevancia de los resultados de la actividad académica para la sociedad; y,
 - c. la calidad de las diversas actividades académicas, evitando sesgos que, por asegurar la calidad de algunas, descuidan la calidad de otras

Diversas intervenciones del público presente ponen en evidencia multiplicidad de objetivos que son perseguidos con la evaluación académica de personas, que dependen en algunos casos de los contextos institucionales, en otros de las motivaciones individuales a la hora de actuar como evaluadores y evaluadoras y en otros refieren a consideraciones sobre lo que se cree o entiende para qué se debería evaluar.

En relación a la distancia entre para qué se evalúa y para qué se debería evaluar, la intervención de una persona integrante del GT sobre Título II del EPD de ADUR, planteaba que se evalúa principalmente para decidir sobre el acceso y/o permanencia a cargos docentes, al RDT o al SNI y para decidir la asignación de recursos para la investigación a través de proyectos, es decir que en su perspectiva se evalúa para asignar reconocimiento, prestigio, poder y dinero. Sin embargo, sostiene que debería evaluarse para promover y estimular ciertas prácticas, tales como la construcción colectiva con actores académicos y, también, con actores no académicos.

En el plano más asociado a las motivaciones individuales sobre para qué evaluar, una persona adscripta al IIBCE indicaba, por un lado, que evaluaba en SNI o PEDECIBA para “mejorar la calidad de la ciencia” y para promover las carreras académicas; en los concursos de acceso a cargos para “seleccionar la o el mejor candidato”; y, en cuanto a los proyectos para identificar “a los mejores” a financiar. Por otro lado, expresaba que esos propósitos no siempre se alcanzan.

Desde el INIA se resaltaba que la evaluación debería tratar de promover y asegurar la calidad de la investigación, la docencia, la vinculación con el entorno. Una persona adscripta a la Udelar subrayaba que la evaluación debía proveer señales hacia lo colaborativo y cooperativo (en coincidencia con otras intervenciones), la exploración de objetivos académicos ambiciosos, “de frontera” y pertinentes o útiles para la sociedad. Desde la Secretaría General de ADUR se destacaba que la evaluación debe contribuir a lograr los objetivos institucionales trazados y cuestionaba, por un lado, hasta qué punto las formas de evaluar existentes en la institución contribuyen al cumplimiento de esos objetivos y por otro, cuáles son realmente esos objetivos institucionales.

La intervención realizada en representación de la UTEC compartía que con la evaluación buscan: i) identificar prácticas, actividades y resultados, y reconocer a las personas y a los grupos que las realizan; ii) recoger insumos e información valiosa para la toma de decisiones sobre la generación de nuevas herramientas e instrumentos de política para la promoción de prácticas y actividades; y, iii) promover prácticas y líneas de trabajo. En 2023 realizaron la primera convocatoria a sus docentes con el fin de generar un registro de investigadores de la UTEC. Eso implicó la evaluación de personas y permitió generar información para establecer una línea de base para la evaluación, monitoreo y ajuste de las políticas que implementan. Señala, además, que están en una fase inicial, definiendo qué van a evaluar y cómo. Se plantean dos áreas de acción: i) la evaluación de docentes y sus actividades, junto con la generación de un sistema de incentivos; y, ii) la evaluación de las estructuras de investigación creadas.

La CH del SNI respondía a la pregunta sobre los propósitos de la evaluación de ese sistema recuperando el mandato y misión que establece su ley de creación, que dice que es para fortalecer la comunidad de investigación en Uruguay y que su misión es identificar, categorizar y evaluar periódicamente a los investigadores en el Uruguay o uruguayos en el exterior o extranjeros que colaboran con la comunidad de investigación en Uruguay y, también, para proveer incentivos económicos a la investigación. Adicionalmente, describía la organización del proceso de evaluación y señalaba que: *“hay un primer análisis que realizan las CTA integradas por especialistas en las disciplinas, organizadas por áreas temáticas, que proveen la mirada más específica o más conocedora del área y sus formas de trabajo (cómo se comunica el conocimiento y cómo se mide, digamos, la calidad, si se quiere). Luego de ese análisis pasa al Comité de Selección (CS) que está integrado por dos personas de destacada trayectoria investigativa de cada área de conocimiento, lo que le da otro de los aspectos interesantes que tiene el sistema (...) que es el cruzamiento entre distintas áreas del conocimiento, para poder contemplar las distintas cosas y que las distintas áreas, de alguna manera, se complementan en ir mejorando mutuamente, en sus visiones y entendimiento. Finalmente, nuestra comisión toma la última decisión, y analizamos particularmente los casos en que las comisiones anteriores no estuvieron de acuerdo y también unos cuantos cientos más, por ejemplo, por promociones, o casos de no renovación – que no soy muchas-, o de apelaciones –que tampoco son muchas-, entre otros (...). Así es como funciona el SNI, y creo que trata de dar buenas garantías. Además, hay otra instancia, que es la del Comité de Revisión, que es un comité ad hoc, integrado por personas que participaron del proceso de evaluación de ese llamado y personas que no participaron, buscando el cruzamiento entre las distintas áreas del conocimiento, cuyo objetivo es analizar las apelaciones que se puedan haber recibido.”*

5. Sugerencias para la transformación de la evaluación

Entre las instituciones internacionales que impulsan cambios en la evaluación, Sutz planteaba que las sugerencias para la transformación que hacen son: i) valorar la calidad principalmente de forma cualitativa; y, ii) valorar adecuadamente todas las actividades académicas.

En sintonía, el rector planteaba que *“debemos comenzar a transitar rápidamente un camino en donde vayamos desde una evaluación muy centrada en la cantidad a una evaluación mucho más centrada en la calidad. (...) tenemos que migrar del esquema publicar o morir a un esquema en que importe lo que se genera sustantivamente”*.

Además de migrar a una evaluación centrada en la calidad, sostiene que también hay que avanzar en el reconocimiento y valoración de “prácticas riesgosas”, hay que generar incentivos para que la academia asuma riesgos y tener cuidado en la estructura de incentivos que se genere, en particular para los y las jóvenes, atendiendo también la dimensión de la salud mental que para nada es una trivialidad.

Asimismo, y en convergencia con otras intervenciones del público, plantea que es importante también el reconocimiento de los objetivos de la construcción institucional en la evaluación. En su intervención el rector sostenía: *“el reconocimiento de lo colectivo y de la construcción institucional está muy subvaluado en todos los procesos de evaluación que tenemos [en el país], su peso relativo es muy bajo”*.

5.1 Sobre como evaluar la actividad académica y el papel de lo cuantitativo

En su presentación, Sutz recogía la experiencia de Noruega para avanzar con propuestas alternativas para la evaluación.

Así citaba a la Academia de Ciencias y Unión de Universidades de Noruega: *“Desarrollar un sistema de evaluación más equilibrado y abierto no significa reducir los requisitos de calidad. Más bien, se trata de reducir el uso de números como “proxy” de la calidad. (...) El objetivo es poner más énfasis en la calidad, el contenido, la integridad académica, la creatividad y las contribuciones a la investigación y/o la sociedad.”*¹⁰

Al tiempo que enfatizaba que nadie, de quienes están impulsando transformaciones a nivel internacional, propone eliminar el uso de indicadores cuantitativos en la evaluación; sino que lo que recomiendan es que dichos indicadores deben ser: i) complementarios a la apreciación cualitativa que es la parte central y medular de la evaluación; y, ii) utilizados con precaución de acuerdo a las guías de uso responsable de indicadores bibliométricos, siguiendo -por ejemplo- los diez principios señalados en el Manifiesto de Leiden de 2015¹¹.

La intervención del rector ponía atención en que la estructura de incentivos que genere la evaluación tiene que estar alineada con sus objetivos. En su visión la evaluación debe ser un instrumento para mejorar prácticas, para asegurar que las personas asuman programas de investigación valiosos para la acumulación del acervo de conocimiento útil de la humanidad, “útil” en su sentido más profundo, no instrumental. Es decir, que sea relevante, que le *“sirva a la humanidad desde el punto de vista cultural, desde el punto de vista social, desde el punto de vista productivo, en todas las dimensiones que hacen a nuestra existencia”*. Entiende que es muy importante ser cuidados y equilibrados con cómo se construye la nueva estructura de incentivos de la evaluación para, por un lado, no reproducir incentivos perversos y, por otro, alcanzar los objetivos que se persiguen: i) que haya ciencia de calidad en Uruguay; ii) que esté conectada internacionalmente; y, iii) que incorporen agendas de investigación innovadoras que contribuyan al bienestar de nuestras sociedades.

¹⁰ Universities Norway (2021) NOR-CAM - A toolbox for recognition and rewards in academic careers. Mayo 2021 Disponible en <https://www.uhr.no/en/f/p3/i86e9ec84-3b3d-48ce-8167-bbae0f507ce8/nor-cam-a-tool-box-for-assessment-and-rewards.pdf>

¹¹ Manifiesto de Leiden 2015 Disponible en http://www.leidenmanifesto.org/uploads/4/1/6/0/41603901/manifiesto_cast.pdf

Desde el punto de vista de la estructura de incentivos que genera la evaluación, la intervención de la UTEC indicaba que allí se plantean, por una parte, que es muy relevante que la evaluación tenga un buen balance entre los resultados o productos (publicaciones, proyectos, dirección de tesis, etc.) y las prácticas o procesos que conducen a esos productos; porque reconocen que la investigación es un proceso de muy largo aliento, ya que generar los ambientes de trabajo conducentes para la investigación y la producción de conocimiento, así como la conformación de grupos lleva su tiempo y eso es muy importante reconocerlo en los procesos de evaluación. Por otra parte, que también sea una evaluación integral, de todas las funciones docentes desarrolladas por las personas con el propósito de promover y estimular el desempeño de los docentes de la UTEC en su integralidad.

Diversas intervenciones de las personas presentes, casi todas, coincidieron en la importancia de transitar hacia una evaluación mucho más centrada y fundamentalmente basada en apreciaciones cualitativas con aspectos de carácter cuantitativo complementarios; justamente para alcanzar los objetivos ya señalados sobre calidad, originalidad-riesgo, relevancia, contribución a la sociedad, construcción institucional, formación, integralidad, etc.

Algunas intervenciones plantearon además que, si la evaluación tiene por objetivo promover la contribución de la actividad académica a la sociedad y su vinculación con actores no académicos, entonces quizás sea importante reflexionar sobre la pertinencia de incluir en la evaluación las miradas de esos actores –abrir la evaluación-. Se reconocía a la vez, que no es fácil y es todo un desafío, pero se sugería revisar la experiencia de la Universidad Intercultural Indígena Ecuatoriana, que está incorporando en sus sistemas de evaluación a las comunidades. Sobre este tema, desde el INIA se aportaba: “Nosotros en el desarrollo de las tecnologías, estamos incorporando usuarios en el proceso de evaluación de las tecnologías. Se llama SERTEC el proceso. Decisores de políticas, actores de la comunidad, sector productivo, nos han aportado bastante en los procesos de evaluación con puntos de vista diferentes que enriquecen y pluralizan respecto de una evaluación puramente interna”.

En su intervención, una persona integrante de la CH del SNI planteaba que *“Uruguay tiene una sana tradición de no ir hacia lo cuantitativo. En el siguiente sentido, para empezar, mi visión es que, en el mundo, cuando se está criticando el cuantitativismo, es esa ilusión de objetividad de que puede haber un algoritmo, una fórmula, [donde] pongo cantidad de papers, de citas, de qué revistas, y no sé qué cantidad de cosas, en la licuadora y sale algo que objetivamente es la calidad del investigador o de la producción o de lo que sea. Cosas casi que de ese estilo se han hecho en distintos lugares del mundo y se está dando marcha atrás. El año pasado la ANII fue, digamos, anfitriona de una reunión de agencias equivalentes a la ANII de todas las Américas, y nos invitaron a participar de una de las sesiones donde se trataba este tema y donde participaban gente de Holanda por videoconferencia porque eran también unos impulsores del asunto y quedó clarísimo que hay muchos sistemas que están dando marcha atrás en esto y se habló del currículum narrativo, y yo para mis adentros decía qué suerte, qué suerte que acá nunca, creo que nunca consideramos, nadie consideró ni propuso ir por ese camino. Entonces, no quiero decir cualquier cosa, ni con esto que sea perfecto lo que tenemos, sin duda es perfectible por muchos lados y todas las sugerencias son bienvenidas, pero creo que de ese punto de vista no estamos tan mal”*.

6. Implementación de las transformaciones

En relación a la implementación de las transformaciones que están siendo impulsadas internacionalmente, Sutz aportaba desarrollos recientes vinculados a CVs narrativos (Holanda y Reino Unido) que informan cualitativamente lo hecho junto con la presentación de algunos productos de investigación seleccionados para su análisis.

Recuperando la experiencia de Holanda, citaba: *“La Academia de Ciencias de Holanda va a introducir un formato de currículum vitae narrativo uniforme en la próxima ronda de una de las mayores subvenciones científicas del país. Esta medida se adecúa a un nuevo enfoque para reconocer y recompensar a los científicos y es una continuación del Position Paper “Room for*

*Everyone Talent*¹² que la Academia publicó recientemente con sus socios. Además, la Academia ha tenido buenas experiencias en los últimos años con el CV narrativo en el piloto previo (...) Un CV uniforme tiene varias ventajas: ya no hay diferencias entre los dominios y el CV es compatible con los principios de la declaración DORA que la Academia firmó recientemente. Como resultado, los CV se evaluarán mucho más por su calidad e impacto (sostenible) que por su cantidad y prestigio”.

Dicho CV narrativo consta de dos partes: i) descripción narrativa del perfil del candidato; permitiendo a los candidatos decidir qué es o no importante mencionar en su CV; y, ii) una lista de no más de 10 resultados clave con una descripción de por qué los candidatos los consideran importantes; de este modo ya no se les solicitará sus listas exhaustivas de publicaciones. Como resultado, se espera que las personas con una trayectoria profesional dinámica tendrán igualdad de oportunidades.

El CV narrativo que se solicita en el Reino Unido¹³, da cuenta de cuatro aspectos, además de permitir incorporar la composición del equipo de trabajo y el rol que tiene cada integrante: i) las contribuciones a la generación de nuevas ideas, herramientas, metodologías y/o al conocimiento; ii) las contribuciones al desarrollo de colegas y al mantenimiento de relaciones de trabajo efectivas y eficientes; iii) las contribuciones a la comunidad ampliada de investigación e innovación; y, iv) las contribuciones a los usuarios y audiencias más amplias de la investigación o la innovación y al beneficio social general.

Adicionalmente, Sutz mostraba antecedentes locales en torno a propuestas sobre evaluación académica en el país. Señalaba en primer lugar que en el taller sobre evaluación académica que hizo CSIC en 2018¹⁴, se discutió mucho y allí se plantearon muchos de los temas tratados en este Coloquio y también se hicieron propuestas que hoy en día están siendo implementadas en distintas partes del mundo. En segundo lugar, recordó una resolución del CDC del año 2012¹⁵ donde se aprobó un documento sobre el cambio en las prácticas de evaluación. Importa remarcar que cuando la Comisión Central de Asuntos Docentes tuvo que responder a un pedido –reciente- del CDC para recomendar una metodología que implementara la evaluación cualitativa que está requerida en el nuevo Estatuto del Personal Docente, la Dirección General Jurídica recordó que esa resolución del 2012 está vigente y constituye un documento oficial de la Udelar. En tercer lugar, destacó lo pionero de la Comisión Directiva del PEDECIBA cuando aprobó los “Criterios, herramientas y procedimientos generales para la evaluación de la actividad académica de los investigadores” en el año 2004¹⁶ y resaltó la capacidad que tuvo Uruguay para reflexionar sobre estos problemas bastante antes que muchos otros a nivel internacional –por ejemplo DORA 2012¹⁷–.

En relación a los CVs narrativos algunas intervenciones ponían de relieve la necesidad de asegurar un espacio en ellos en el que las personas pudieran contar diversas situaciones vinculadas a sus

¹² Ver VSNU, NFU, KNAW, NWO and ZonMw (2019). Position paper 'Room for everyone's talent': towards a new balance in the recognition and rewards of academics. Noviembre 2019. Disponible en https://www.nwo.nl/sites/nwo/files/media-files/2019-Recognition-Rewards-Position-Paper_EN.pdf

¹³ Disponible en <https://www.ukri.org/publications/resume-for-research-and-innovation-r4ri-template/>

¹⁴ Ver Unidad Académica de CSIC (2018) Relatoría del Taller de Trabajo y Reflexión sobre la Evaluación Académica. Disponible en <https://citinde.ei.udelar.edu.uy/uploads/bibliografia/c95dfb08b2a9c30f91693be920b48cf8b40ded0b.pdf>

¹⁵ Ver Resolución Nro. 4 CDC en sesión ordinaria del 31 de Julio de 2012 Disponible en <https://www.expe.edu.uy/expe/resoluci.nsf/e3365ff03c2a3d6103256dcc003b9031/4caf4bf2d2feffb003257a45005a1677?OpenDocument>

¹⁶ Ver PEDECIBA (2004) Criterios, herramientas y procedimientos generales para la evaluación de la actividad académica de los investigadores. Documento aprobado por la Comisión Directiva del PEDECIBA en su sesión del 17 de junio de 2004. Disponible en <https://www.pedeciba.edu.uy/uploads/reglamento/6b936851737683dda980694825d4fd4690d1d424.pdf>

¹⁷ Ver San Francisco Declaration on Research Assessment (2012) Disponible en <https://sfdora.org/read/>

circunstancias vitales y personales y que, en términos más generales, ese aspecto sea tenido en cuenta en la evaluación. Así lo planteaba una persona adscripta al IIBCE: “quizás ahí en los currículums narrativos uno pueda narrar *yo trato de hacer esto, pero en realidad tengo que cuidar cuatro hijos, etc.* Eso es algo que me parece que hay que tener en cuenta”.

Otra intervención aportaba la experiencia de evaluación del área de biología del PEDECIBA y señalaba que para las renovaciones vienen implementado un formulario desde 2016 que integra aspectos cuantitativos y cualitativos, con espacios para narrar diversas situaciones (enfermedad, maternidad, imprevistos, mudanza del laboratorio, cambios de línea de trabajo). Además, indicaba que las comisiones evaluadoras cuentan con unas pautas generales con los requisitos mínimos que se deben cumplir, y resalta que las colaboraciones se valoran positivamente y que en lo que se hace más énfasis es la formación de recursos humanos –“jóvenes investigadores”.

Desde la CCAD se compartía la experiencia en relación al proceso de generación de la Propuesta de Ordenanza de Evaluación Docente. El primer punto a considerar fue que el nuevo EPD establece que la evaluación de los docentes de la Udelar tiene que ser: i) cualitativa; e, ii) integral de todas sus funciones: “[los docentes] tenemos que realizar todas las tareas universitarias y nos tienen que evaluar en consecuencia. Y, a medida que la investigación se toque con la enseñanza y con la extensión, obviamente vamos a ser mejores docentes. Eso, en algún lugar tiene que aparecer y se tiene que valorar” También se destaca que la propuesta de ordenanza de evaluación docente es el resultado de un proceso de negociación y participación colectiva, y que provee lineamientos generales que todos los servicios debieran cumplir y que, justamente, se trata de lineamientos generales para respetar también sus especificidades y que éstos puedan enriquecerla con sus particularidades. Adicionalmente, se resalta que en el Anexo de la propuesta de ordenanza se provee lineamientos para la elaboración de informe y plan de actividades a efectos de transparentar y que a todas las personas se las evalúe con base a esos contenidos. Otro aspecto que hace a las garantías de la evaluación docente y que se resalta de la propuesta, es que busca incorporar la mayor cantidad de voces sobre el desempeño docente y esto es un enorme desafío porque “¿quiénes son quienes deben evaluar, el superior jerárquico, el grupo, la unidad académica, el instituto, el departamento, el servicio, los pares, que están cerca, los que están lejos, los estudiantes, etc... Pero bueno, estratifiquemos una visión general, el departamento quiere ir para ese lado, es relevante esa investigación para el servicio. Vayamos incorporando capas para que cada uno de los actores diga qué es lo que le parece sobre ese desempeño docente”. Además, también se subraya otro aspecto importante de la propuesta, vinculado con lo anterior, que tiene que ver con que es de carácter formativo en tanto busca que sea una herramienta de aprendizaje y mejora continua tanto para la persona como para el colectivo en que está inserto. “Es una propuesta para ir construyendo y que, a nivel institucional, el sistema empiece a agregar esas capas de evaluación que creo que necesitamos todos y que debe ser integral.”

Por su parte, la CH del SNI compartía la experiencia con el CVUy y si bien reconocía que está lejos de los CV narrativos actuales, consideraba que incluye diversos espacios que permiten narrar diversos aspectos. Así indicaban que: i) incluye una sección donde las y los investigadores deben establecer cuál es la significación de su trabajo; ii) para las renovaciones del SNI, está la sección resumen del período donde se le pide especialmente a los y las postulantes que cuenten cuál ha sido el impacto/significación de su trabajo en el período; iii) hay otro apartado para dar cuenta de los aspectos de construcción institucional; y más recientemente, iv) se incorporó un nuevo campo para documentar mejor “las formas de producción de conocimiento no tradicionales” para que las comisiones lo puedan evaluar. Se destaca que es un desafío, y aunque tiene una pauta sobre cómo entregar esa información, no es fácil porque “si bien de nuestro lado, es importante lo que el investigador narre, las comisiones -por supuesto- tienen que verificarlo y para eso tiene que haber alguna manera verificable. Porque si no se vuelve todo muy poco sólido, a mi juicio. Se trata de tener elementos para evaluar la contribución a la sociedad en un sentido amplio, formas

no tradicionales, aspectos que, por supuesto, tienen que estar en el CVUy, siempre tienen que tener una raíz en la investigación, en la creación de conocimiento original.” Adicionalmente, otra persona integrante de la CH del SNI enfatizaba la importancia del CVUy y de sus espacios para narrar en las instancias de evaluación, ya que la información que esté contenida allí es la que tienen las y los evaluadores para evaluar lo realizado. Además, agrega: “cuando apareció por primera vez el CVUy, teníamos la ilusión de que teníamos que llenar un CV único, y lo íbamos a presentar donde fuera, pero todavía no, no llegamos a eso. Por eso quería marcar la importancia que existe en tener un único CVUy que contemple todos esos espacios”.

Asimismo, desde la ANII se resalta que el CVUy es permeable a las necesidades informacionales para la evaluación del SNI, de la Udelar, por lo que están abiertos y se ponen a disposición para atender los cambios y mejoras al CVUy que la Udelar proponga.

7. Hacia una evaluación académica cualitativa, integral, plural y más humana: algunos desafíos

7.1 Desafíos para la evaluación cualitativa

En el cierre de su intervención, el rector planteaba un conjunto de desafíos para la evaluación y en particular para la evaluación cualitativa, vinculados a: i) comunidad pequeña; ii) masa crítica y necesario cuidado en las estructuras de incentivos y las señales que generan; iii) desiguales relaciones de poder y endogamia; y, iv) el tiempo y recursos que requiere. Lo expresaba así: “estoy convencido que las evaluaciones cualitativas son mucho más relevantes que las cuantitativas (...), que hay que premiar la calidad, las construcciones colectivas, que las personas asuman riesgos en sus programas de investigación, pero también creo que tenemos que ser cuidadosos con la estructura de incentivos que generamos -tanto al interior de nuestras instituciones como también con cómo dialogan con lo internacional-, y ese diálogo, es el problema básico que tenemos cuando hacemos evaluación cualitativa y que tiene que ver con la exigencia que supone para ciertos cuerpos docentes o académicos relativamente pequeños en relación a lo que hay que evaluar. Es mucho más difícil hacer una buena evaluación cualitativa que hacer una evaluación cuantitativa. Y en simultáneo, las lógicas que puede haber en cuanto a ciertas reproducciones, no necesariamente de buena calidad, en los poderes relativos en las instituciones”.

7.1.1 El problema de la masa crítica y el volumen de la evaluación

Distintas intervenciones del público enfatizaban que la evaluación cualitativa es “necesaria”, “que se debe”, “que se puede”, “que hay que hacerla”, al tiempo que reflexionaban sobre cómo lograrlo debido a los desafíos ya señalados tanto en términos de los requerimientos de masa crítica como por el tiempo que insume. A modo de ejemplo se decía:

- i) “requiere que quien lo haga tenga una capacidad muy alta de poder entender en un mundo en donde la generalidad no aplica a las particularidades de los temas (...) pero, insisto, se debe avanzar en esa dirección”;
- ii) “siendo una comunidad académica tan pequeña, (...) no habría habido suficientes personas en la comisión para poder hacer una opinión pertinente y ecuaníme sobre los contenidos. No estoy diciendo que defiende ese sistema de evaluación [cuantitativo] que me parece aberrante. Estoy señalando que, si no resolvemos el problema práctico de cómo evaluar ese volumen de información, no vamos a poder cambiar la situación. Y lo planteo porque quiero cambiarla y no alcanza con tener la intención de cambiarla”.

7.1.2 El problema del “sobrecalentamiento” del sistema de evaluación

Otra preocupación vinculada con el punto anterior es la planteada por la Prorectora de Investigación respecto al ya existente “sobrecalentamiento” del sistema de evaluación actual y

decía: *“cada vez son más instancias las de evaluación, más cosas para evaluar, más sofisticación en los mecanismos de evaluación y en el caso de Uruguay, que somos una comunidad obviamente pequeña, la percepción es que estamos al límite digamos, de la decencia evaluativa por decirlo de alguna manera”*.

Distintas intervenciones coincidieron en este aspecto, y con cierto humor en referencia al título del ciclo de encuentros, una persona insistía “evaluar [menos] es necesario”, la enorme cantidad de evaluaciones compromete su calidad y nos conduce a cuantificar porque es más ‘fácil’ y rápido; la cantidad creciente de evaluaciones limita la posibilidad de tender a una evaluación más cualitativa que es lo deseable”; y sostenía “hay que evaluar menos si queremos mejorar la calidad de la evaluación”.

En ese sentido, docentes de la Udelar comentaron los esfuerzos que se están realizando orientados a unificar la evaluación en el RDT con la de renovación de los cargos, *“no totalmente, pero por lo menos sincronizarla y que el informe de actividades y el plan de trabajo sean los mismos. Son elementos muy importantes de avance. Después, probablemente existan dos evaluaciones distintas, porque se evalúan cosas ligeramente distintas.”*

Una persona adscripta al IIBCE, reflexionando en relación a su experiencia en esa institución, sugería que tal vez no eran necesarias tantas instancias de evaluación de las personas, en referencia a que una misma persona puede estar siendo evaluada por hasta cuatro instancias (cargo, RDT, SNI, PEDECIBA), además de las evaluaciones en fondos concursables, y decía: *“nuestro instituto es un instituto público que depende del Ministerio de Educación y Cultura, los cargos son puestos fijos, somos funcionarios públicos, no se nos evalúa para mantener el puesto, primera cosa a resaltar. Sin embargo, eso no quiere decir que no hagamos nada. No, no es así. Somos evaluados por PEDECIBA y SNI (...), a lo que voy es que capaz que esa necesidad de evaluar todo el tiempo no es necesaria y capaz alcanza con evaluar una sola vez (...)”*

7.1.3 El problema de la discrecionalidad y la subjetividad de la evaluación

A la vez, en la discusión y como aspecto que también hay que atender y resolver, se planteaba que las evaluaciones en general y las cualitativas en particular tienen el riesgo de incrementar el grado de discrecionalidad en los procesos de evaluación y la subjetividad y que, por lo tanto, es preciso cuidar y profundizar la transparencia de dichos procesos para asegurar las garantías y la legitimidad de dichos procesos.

En ese sentido, una intervención desde la Udelar planteaba que “la gente [y en particular la más joven] equivocadamente puede sentir que las evaluaciones cuantitativas son una salvaguarda de las garantías” en tanto permite objetivar los requisitos de evaluación a través de su aproximación mediante ciertas cantidades mínimas requeridas de cursos, de publicaciones, de factores de impacto, etc. y consideraba que este es un tema que hay que atender y resolver de cara a la transición a procesos de evaluación más cualitativos.

Asimismo, y desde la AUPP en coincidencias con otras intervenciones, señalaban que es preciso transparentar los procesos de evaluación actuales, sus criterios y procedimiento de selección, así lo expresaban: *“una cosa que nos parece importante tener en cuenta en la evaluación académica es la transparencia de los sistemas. En concreto voy a dar el ejemplo de las becas, porque es lo que más cercano se nos hace. Muchas veces no sabemos qué cantidad de becas, qué se va a evaluar más, cuáles son las áreas que van a tener más cantidad de becas, cuáles son los criterios que se van a aplicar. Eso muchas veces viene ex post pero no ex ante, entonces las reglas de juego pueden ser bastante cambiantes. La transparencia de los sistemas a veces se da en ciertas partes, por ejemplo, a veces se presentan muy claramente los criterios, pero no se presenta muy claramente cómo es el proceso de evaluación. Todo esto genera mucha incertidumbre, sobre todo cuando tenemos varios sistemas de evaluaciones que están solapados”*.

Asociado a lo anterior, pero vinculado al problema de la subjetividad por el temor a las injusticias se subraya la necesidad de que debemos confiar más. Una intervención del IIBCE resaltaba este asunto así: *“creo que le tenemos miedo a la subjetividad porque le tenemos miedo a no ser justos o justas, (...) tenemos que cambiar un poco eso y confiemos en los tribunales confiemos en la gente que está ahí (...)”*.

Asimismo, la intervención de la UTEC coincide y enfatiza que la evaluación debe ser transparente y establece que siempre tendrá cierta subjetividad: *“[respecto al] tema de la subjetividad y objetividad que tanto se discute, creo que no tenemos que tenerle tanto miedo a que un sistema de evaluación no sea justo si no es 100% objetivo, no creo que la evaluación pueda ser nunca 100% objetiva y más si estoy pensando en evaluar personas, siempre habrá y tiene que haber un grado de subjetividad. Lo que creo que sí es importante es que el sistema sea transparente, que las reglas estén claras y acordadas entre todos; pero no podemos pedir que sea aséptico (...)”*.

Adicionalmente, en relación a la transparencia, las relaciones desiguales de poder y los conflictos de interés en los procesos de evaluación, desde la Secretaría General de ADUR se plantea la necesidad de diversificar y rotar a quienes evalúan: *“¿Quiénes son los evaluadores a los cuales un docente se somete? ¿O un proyecto se somete? Si, por ejemplo, se sabe que hay conflictos de intereses y son siempre evaluados por las mismas personas que tienen conflictos de intereses, bueno, ahí podemos tener un problema en lo que tiene que ver con la objetividad de la evaluación”*.

7.2 Desafíos para la evaluación integral

Un desafío sobre el que se intercambiaba bastante se vincula con la cuestión sobre quiénes deben evaluar, en cuáles espacios y qué aspectos deben considerarse si el objetivo es lograr que la evaluación sea más integral.

La intervención de una persona integrante del GT Tít.II EPD ADUR, señalaba: *“si uno quiere que la evaluación sea más integral, y más colectiva, como se señala, y que abarque aspectos no sólo de la producción científica, sino también sus aportes a la construcción colectiva, a la construcción institucional, o sus aportes a la enseñanza, o al relacionamiento con el medio, (...) entonces cobra mucha importancia quién evalúa eso. Porque se ha hablado de quién es capaz de entender el paper o la investigación... Pero para apreciar y evaluar estas otras cosas, no necesariamente es la misma persona que es capaz de evaluar el paper (...) Por ejemplo, para la evaluación sustantiva de las actividades de relacionamiento con el medio, deberían estar también los actores no académicos. O sea que quienes tienen que formar parte de las y los evaluadores, no son los mismos que tenemos hoy. Y eso tiene que ver también con, por ejemplo, la actividad integral del docente en su entorno de trabajo. En mi opinión, la única forma de evaluar ese desempeño adecuadamente es de cerca, no de lejos: el Consejo (...), sus estudiantes, sus colegas, ésta es la gente que sabe lo que pasa. (...) Entonces, este tema de dónde se evalúa y quiénes participan en la evaluación, es crítico en este asunto”*.

Otras intervenciones también coincidieron en la relevancia de la pregunta sobre quiénes deben evaluar, al tiempo que destacaban la importancia de la transparencia de los procesos para asegurar las máximas garantías a las y los evaluados, para aprovechar las virtudes de ciertas cercanías en los procesos de evaluación por ejemplo en términos de los aprendizajes individuales y colectivos.

En este sentido, resulta elocuente la intervención que hacía un integrante de una CTA del SNI: *“(...) hoy los sistemas de evaluación son distantes, muchas veces se evalúa a la gente que no está cercana institucionalmente, en la investigación. Pero si hacemos una evaluación cualitativa tenemos que asumir que la cercanía no es algo que esté mal necesariamente. Sí se mezcla con cuestiones de poder, conflicto de intereses y [por lo tanto] hay que asegurar los derechos de los evaluados. Pero también es cierto que, en una comunidad de investigadores seria, hay que fortalecer las comunidades sobre temas específicos, y eso también implica a la evaluación. (...) evaluar no en el sentido de decir: no estás aprobado o sí estás aprobado, sino evaluar en el sentido de elaborar un conocimiento crítico que le pueda aportar al evaluado; quien mejor puede*

hacerlo es alguien que sabe lo que está haciendo esa persona. Creo que ahí hay un tema sobre la cercanía y sobre los derechos de los evaluados, que es positiva”

7.3 Desafíos para la evaluación plural

Otro desafío para la evaluación, es el reconocimiento de la diversidad de las variadas formas de producción académica existente, sus tradiciones y prácticas. En el intercambio se planteaba la importancia de que la evaluación cualitativa colabore a reconocer y recuperar esa diversidad, que los sistemas actuales de evaluación tienden a homogeneizar. Para ello se planteaba la necesidad de pluralizar los criterios de calidad, integrando actividades y productos de acuerdo con los objetivos del trabajo académico y atendiendo siempre a los contextos en que se producen. A la vez se señalaba la relevancia de generar espacios para la reflexión sistemática en torno a la evaluación y sus criterios.

En su intervención, un integrante de una CTA del SNI decía: *“el primer desafío es que la ciencia o el pensamiento académico no es una categoría homogénea, sino que es ultra diversa, las formas de producir son diversas, los productos que hacemos son diversos, las formas de intervenir en la sociedad son diversas, entonces las evaluaciones también son diversas, que es un gran problema que incluso tenemos en la ANII, lo vemos constantemente, incluso lo vemos entre diferentes disciplinas. Tenemos esta lógica que evalúa papers y artículos, pero muchas veces para algunos es más fácil publicar que para otros... Entonces, esta cuestión de la heterogeneidad me parece que es un tema central si pensamos en una evaluación cualitativa, reconocer esta heterogeneidad. En general todos los sistemas de evaluación han tendido a homogenizar, usan un lenguaje que viene más de las ciencias duras y lo uniformizan para todo el mundo, y en general los que venimos de las humanidades actuamos tímidamente diciendo sí, sí, bueno, pero me parece que hay que reconocer la heterogeneidad para pensar la evaluación y la evaluación cualitativa”*.

Desde el INIA se complementaba este asunto así: *“¿qué es calidad? Y ahí pienso que tenemos que pluralizar los criterios de calidad, me parece. Y lo otro es tener un espacio para la reflexión. Tanto si la evaluación es cuantitativa como cualitativa se requiere de reflexión y análisis (...) porque hecha la ley, hecha la trampa”* haciendo referencia a las estrategias adaptativas que desarrollan las personas, colectivos e instituciones para desempeñarse mejor –lograr mejores posiciones relativas- en los sistemas de evaluación.

7.4 Desafíos para la evaluación más humana y con perspectiva de género

Otro desafío para la evaluación en general y la cualitativa en particular es cómo incorporar la perspectiva con equidad de género, reconociendo, valorando y no castigando las carreras académicas de las mujeres, y en particular de las que están realizando el posgrado. En el intercambio se anotaba este desafío derivado de los rezagos que acumulan como resultado de sus mayores cargas de cuidados, *“¿cómo considerar los cuidados en la evaluación?”*. Asimismo, se señalaba que es importante la formación a las y los evaluadores, porque *“una cosa es hablar de la perspectiva de género, generar reglamentos, pero si después quienes evalúan no tienen una formación, la aplicación de estos criterios es nula o es muy a voluntad”*.