

REVISIÓN DE SISTEMAS DE EVALUACIÓN DE IMPACTO NO ACADÉMICO DE LA INVESTIGACIÓN A NIVEL INTERNACIONAL

cnid Consejo Nacional
de Innovación
para el Desarrollo

DOCUMENTO DE TRABAJO N°7

Carla Alvial Palavicino



REVISIÓN DE SISTEMAS DE EVALUACIÓN DE IMPACTO NO ACADÉMICO DE LA INVESTIGACIÓN A NIVEL INTERNACIONAL

DOCUMENTO DE TRABAJO N°7

Carla Alvial Palavicino

La Serie Documentos de Trabajo de la Secretaría Ejecutiva del Consejo Nacional de Innovación para el Desarrollo busca abrir temas de discusión que permitan avanzar en el diseño consensuado de estrategias de largo plazo en materia de ciencia, tecnología e innovación para el desarrollo de nuestro país.

A continuación presentamos un documento que caracteriza y analiza distintos sistemas de evaluación del impacto no académico de la investigación para proveer insumos al diseño de sistemas de evaluación, que aportan a la instalación de la nueva institucionalidad en ciencia, tecnología, conocimiento e innovación.

Santiago, diciembre de 2018

Autor

Carla Alvial Palavicino

Edición

María José Menéndez Bass

Diagramación

Oriana Avilés M.

Esta obra está bajo una Licencia Creative Commons
Atribución— NoComercial-CompartirIgual 4.0
Internacional. Para ver una copia de esta licencia, visite
<http://creativecommons.org/licenses/by-nc-sa/4.0/>

Esta licencia significa que no se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original. Usted debe darle crédito a esta obra de manera adecuada, proporcionando un enlace a la licencia, e indicando si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo del licenciente.

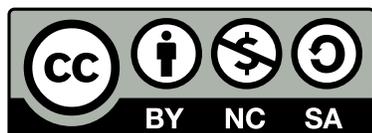


Tabla de contenidos

Introducción	7
1. Evaluación de impacto no-académico de la investigación	8
2. Revisión de metodologías internacionales de evaluación de impacto no académico	10
2.1. Evaluaciones de proyecto	10
2.1.1. National Science Foundation (NSF) - Broader Social Impacts (EEUU)	10
2.1.2. Payback Framework (Reino Unido)	12
2.1.3. Social Impact Assessment Methods (SIAMPI) Unión Europea (UE)	15
2.1.4. Research Quality Plus RQ+ (Instituto de Desarrollo Internacional de Canadá (IDRC) - Canadá	17
2.2. Evaluaciones de centros, organizaciones y/o programas	22
2.2.1. ASIRPA (INRA - Francia)	22
2.2.2. Commonwealth Scientific and Industrial Research Organisation CSIRO (Australia)	23
2.2.3. SAEI (Small Advanced Economies Initiative)	25
2.3. Evaluaciones a nivel nacional	28
2.3.1. Research Impact Framework REF14 (Reino Unido)	29
2.3.2. Piloto de Evaluación de Impacto y Vinculación (Australia)	32
2.3.3. Research quality evaluation in Sweden - FOKUS (Suecia)	34
2.3.4. Standard Evaluation Protocol - SEP (Países Bajos)	36
3. Conclusiones	39
4. Referencias	41

Introducción

Una de las propuestas de la Estrategia Nacional de Innovación y uno de los ejes de la política de innovación del Gobierno es orientar parte de los esfuerzos de CTCI en contribuir a abordar grandes retos o desafíos nacionales que se hacen cargo de las preocupaciones que compartimos los chilenos y que, además, pueden ser una fuente de creación de valor para el mundo.

El impacto de dicha contribución suele ir más allá de la creación de conocimiento de excelencia, buscando dar cuenta de su aporte a objetivos sociales y políticos, usualmente agrupados bajo la noción de impacto “no académico” de la investigación. Esto incluye la dimensión económica, social, cultural, ambiental y política.

En este contexto la Secretaría Ejecutiva del Consejo Nacional de Innovación para el Desarrollo¹ hizo una revisión de metodologías que se utilizan en distintos países para evaluar el impacto no académico de la investigación. Este documento sistematiza algunas de las mejores prácticas internacionales de este tipo de evaluación.

Es importante notar que este tipo de evaluaciones son relativamente nuevas en el mundo (menos de 10 años), por lo que la discusión sobre la metodología y su utilidad en general aún está abierta. El objetivo de esta revisión es contribuir a una discusión informada sobre el diseño de un sistema de evaluación que integre esta dimensión de impacto a objetivos sociales en Chile, la cual debe ser contextualizada a las características y competencias de las instituciones de política y gestión de la investigación en el país, y en función de los objetivos que la estrategia nacional de CTCI plantea.

El documento se estructura de la siguiente forma: la primera sección ofrece una breve introducción a los principios de la evaluación de impacto no académico de la investigación; la segunda parte describe distintas prácticas de evaluación a nivel internacional, divididas en tres categorías, según el tamaño de la unidad de evaluación, proyecto, programa y nacional; finalmente, la tercera sección entrega conclusiones.

¹ Órgano asesor de la Presidencia de la República que genera orientaciones estratégicas para fortalecer la contribución de la ciencia, la tecnología, el conocimiento y la innovación al desarrollo del país. Para ello, entre sus funciones está generar análisis prospectivos de las tendencias de desarrollo globales y nacionales, asesorar en la formulación de propuestas para fortalecer y desarrollar el sistema, y elaborar y revisar la estrategia nacional de ciencia, tecnología, conocimiento e innovación para el desarrollo.

1. Evaluación de impacto no-académico de la investigación

Las evaluaciones son un componente esencial de las políticas públicas, pues permiten recolectar evidencia sobre el nivel de cumplimiento de los objetivos estipulados, generar aprendizajes y buenas prácticas, y definir la distribución de recursos públicos entre distintos objetivos de interés público.

La evaluación de impacto de la investigación cumple al menos cuatro objetivos. Es importante tenerlos en consideración pues los métodos utilizados para la evaluación en cada caso deben alinearse con estos objetivos, y estar en constante revisión (Adam et al., 2018):

- » **Promoción:** promover y dar fundamentos sobre la importancia de la investigación, mostrando el valor de la ciencia (además del valor intrínseco de la producción de conocimiento), especialmente por parte de tomadores de decisiones en ciclos de definición de presupuesto, y en su validación en la opinión pública.
- » **Análisis:** el entender cómo funciona la investigación es esencial para mejorar sus procesos, resultados y potenciar su impacto. Por ejemplo, reconocer barreras, entender qué tipo de equipos y liderazgo son necesarios, qué mecanismos aseguran un mayor impacto, etc.
- » **Transparencia y Rendición de Cuentas:** demostrar a quienes pagan impuestos y a la sociedad en general cómo se gastan los fondos de investigación y con qué fines.
- » **Asignación de fondos:** en algunos casos, es un objetivo directo de la evaluación de impacto, que busca asignar fondos de acuerdo al impacto de la investigación y otros criterios relevantes, como la excelencia científica.

Esta diversidad de objetivos resulta en una gran variedad de propuestas metodológicas para la evaluación de impacto y diferentes usos de la evaluación. Por ejemplo, aquellas evaluaciones que buscan facilitar los impactos de la investigación a través de un mayor entendimiento de los procesos a través de los cuales se crean impactos tienen distinto énfasis que aquellas que buscan mejorar la transparencia en asignación de fondos públicos. Como se detalla en la sección siguiente, muchas veces las metodologías mezclan distintas formas de evaluar impacto, de manera de abordar sus distintas dimensiones con mayor profundidad.

A grandes rasgos, existen dos formas de utilizar la evaluación: la determinante y la formativa. El sistema de evaluación determinante promueve consecuencias inmediatas y sistemáticas: aquellas instituciones bien calificadas son compensadas, y los mal calificados, sancionados. El sistema formativo funciona inicializando un proceso de aprendizaje organizacional, en el que se decide caso a caso si aquellos más débilmente evaluados recibirán una oportunidad para mejora (Whitley & Gläser, 2007).

En el caso de la investigación realizada con fondos públicos, la evaluación ayuda a tomar decisiones de inversión en investigación, ya sea a través de proyectos, programas, centros de investigación o universidades, buscando conciliar los objetivos de la mejor investiga-

ción y el mayor aporte al interés público y social. Estas evaluaciones no están libres de controversia, pues definir que una investigación es mejor que otra, ya sea bajo criterios de excelencia académica, productividad, impacto y otros, puede variar mucho entre disciplinas y enfoques, de acuerdo a los intereses de investigación de cada uno de ellos (Molas-Gallart & Castro-Martínez, 2007).

Por su parte, la evaluación de los impactos no académicos de la investigación no es nueva. Los impactos económicos de la investigación han sido evaluados desde los años 50; lo novedoso hoy son las expectativas de que la investigación pueda contribuir a resolver problemáticas sociales y de interés público, y la evaluación de impacto en relación a estos aspectos (Joly et al., 2015; Kaldewey, 2017). Sin embargo, y a diferencia de las evaluaciones de excelencia y productividad académica, a la fecha no existe una metodología estándar de evaluación de impacto de la investigación a nivel internacional (Bornmann, 2013), aunque la gran mayoría de este tipo de evaluaciones se basan en los estudios de caso (Joly et al., 2015).

Frente a la complejidad de evaluar algo tan diverso como el impacto no académico, la Escuela Internacional de Evaluación de Impacto de la Investigación (ISRIA - International School of Research Impact Evaluation) considera que existen cinco desafíos metodológicos importantes para este tipo de evaluación (Adam et al., 2018):

1. Desfases temporales: ¿Cómo evaluar el impacto de la investigación cuando normalmente esta toma bastante tiempo para generar impacto? ¿Cuándo es el momento correcto?
2. Atribución y contribución: ¿Cómo atribuimos impactos particulares a proyectos de investigación e investigadores, si la investigación suele ser de naturaleza incremental y colaborativa?
3. Diferencias marginales: ¿Cómo distinguimos entre impactos fuertes y difusos si aún no existe un entendimiento compartido de los estándares de evaluación de impacto?
4. Costos de transacción: ¿Cómo aseguramos que los beneficios de la evaluación de impacto de la investigación no sobrepasen sus costos, especialmente si el proceso es costoso y engorroso?
5. Unidad de evaluación: ¿Cómo determinamos una unidad apropiada de evaluación si la investigación puede ser multi-disciplinaria y multi-impacto?

A estos desafíos es importante agregar que los tipos de impacto pueden variar considerablemente dependiendo de la disciplina y/o tipo de proyecto. La investigación en salud o en ciencias ambientales suele tener una relación mucho más cercana con los potenciales beneficiarios o usuarios, a diferencia de la investigación en ciencias exactas como física o matemática, o en artes y humanidades donde los impactos suelen ser más indirectos y de largo plazo (Spaapen, 2018)².

2. Revisión de metodologías internacionales de evaluación de impacto no académico

A continuación, se describen algunas metodologías utilizadas para el análisis de impactos no académicos de la investigación. Para la búsqueda de estas metodologías se privilegiaron manuales o guías metodológicas de agencias de educación superior, investigación e innovación a nivel internacional, las cuales fueron complementadas con referencias de literatura académica internacional en relación a evaluación de impacto. Las metodologías fueron agrupadas en tres tipos, de acuerdo al tamaño de la unidad de evaluación: evaluaciones de proyectos, evaluaciones de centros y programas y evaluaciones nacionales (o sistemas de financiamiento de la investigación basados en desempeño). Las metodologías de la primera y segunda categoría pueden ser, en algunos casos, intercambiables³.

2.1 Evaluaciones de proyecto

2.1.1. National Science Foundation (NSF) – Broader Social Impacts (EEUU)

Desde el año 1997, la National Science Foundation de Estados Unidos utiliza como criterio de evaluación de proyectos el criterio de impactos amplios (broader impacts), entendido como “el potencial de beneficiar a la sociedad y el potencial de alcanzar objetivos sociales específicos y deseables” (National Science Foundation (NSF), 2015). El objetivo de esta iniciativa es crear espacios para las comunidades poco representadas en ciencia y tecnología (como minorías, grupos con capacidades diferentes, etc.), además de potenciar las alianzas y efectos de la investigación realizada.

-
- 2 Esta generalización no siempre se cumple y no toma en consideración las actividades de divulgación y educación, donde las ciencias exactas y artes y humanidades generan impacto a corto plazo. Sin embargo, en general el impacto, y especialmente el involucramiento de usuarios en la investigación ya sea pacientes o comunidades, es más común en áreas como salud o conservación.
 - 3 La clasificación fue realizada por el autor luego de revisiones de casos. Esta revisión abarca algunos de los casos más importantes referenciados en la literatura especializada; sin embargo, no pretende ser exhaustiva

Corresponde a una evaluación *ex ante*, que es utilizada como uno de los criterios de selección de proyectos a financiar. Los investigadores, al presentar su propuesta de investigación para ser financiada por la NSF, deben indicar no sólo la relevancia científica de su trabajo, sino que también sus impactos amplios en la sociedad. De acuerdo a la NSF, estos impactos pueden ser en enseñanza, capacitación y aprendizaje; participación de grupos sub-representados; mejora de la infraestructura para investigación y educación; mejora del entendimiento científico y tecnológico; y/o beneficios directos o indirectos a la sociedad (Roberts, 2009).

Una de las principales críticas que ha recibido esta iniciativa, es que el criterio de impactos amplios no es claro, así como la forma en que es juzgado y su peso relativo en relación al criterio de mérito intelectual de la propuesta. Tampoco queda claro en qué medida las actividades orientadas a los impactos amplios deben corresponder con las actividades y propósitos del proyecto de investigación (Roberts, 2009). Las propuestas de impacto son evaluadas a través de paneles de revisores expertos de cada área, en el mismo proceso en que la propuesta de investigación es evaluada. Cada panel de expertos decide los criterios bajo los que se realiza la evaluación de la propuesta. Este aspecto ha sido objeto de críticas pues no hay claridad de cómo las propuestas de impacto amplio contribuyen a la selección de un proyecto.

Los impactos amplios son además evaluados durante la ejecución del proyecto y al final de este, con el objetivo de analizar como contribuyeron a la investigación. La NSF ha desarrollado una serie de guías metodológicas para abordar los impactos amplios (Allen et al., 2008). En ellas se destaca la importancia de considerar la evaluación como un elemento de apoyo al diseño de proyectos, más que un elemento que sólo se toma en consideración al final de un ciclo. Esto permite entender quién es la audiencia de un proyecto, cómo puede ser abordada, qué metodologías serán requeridas para conocer la efectividad de nuestras intervenciones, y qué efectos indeseados pueden ocurrir.

De todas maneras, la NSF propone un marco común respecto del tipo de impacto que estas iniciativas pueden lograr (Tabla 1). Estas categorías se basan en investigación en educación científica y en ciencias de la educación; además, muchas de ellas comúnmente ocurren en proyectos de educación científica informal y además se intersectan. Estas categorías se utilizan también para sistematizar los proyectos de impactos amplios (base de datos online) por lo que se sugiere que los proyectos mantengan sus evaluaciones de impacto dentro de las categorías predefinidas, y hagan uso de “otros” sólo cuando sea estrictamente necesario.

Tabla 1. Categorías de Impactos Amplios de la NSF

Categoría de Impacto	Definición general
Toma de conciencia, conocimiento y entendimiento	Demostración medible de una evaluación de un cambio en, o el ejercicio de, toma de conciencia, conocimiento y entendimiento de una temática científica particular, concepto, fenómeno o carreras centrales al proyecto.
Vinculación o interés	Demostración medible de evaluación de un cambio en, o el ejercicio de, vinculación o interés en una temática científica, concepto, fenómeno, teoría o carreras centrales al proyecto
Actitud	Demostración medible de evaluación de un cambio en, o el ejercicio de actitud hacia una temática científica, concepto, fenómeno, teoría o carreras centrales al proyecto, y en relación a las capacidades propias en estas áreas. Aunque similar a conciencia/interés/vinculación, las actitudes se refieren a cambios relativamente estables, constructos más intrincados como la empatía por los animales y sus hábitats, la apreciación por el rol de los científicos en la sociedad o actitudes hacia la investigación en células madre.
Comportamiento	Demostración medible de evaluación de un cambio en, o el ejercicio de un comportamiento relacionado con un tema STEM (ciencia, tecnología, ingeniería y matemáticas). Este tipo de impacto es particularmente relevante para proyectos medioambientales o con énfasis en ciencias de la salud dado que la acción es un resultado deseable.
Habilidades	Demostración medible del desarrollo y/o refuerzo de habilidades, ya sean completamente nuevas o la mejora, incluida la práctica, de desarrollo de habilidades. Esto suele ser considerado como un aspecto del proceso de conocer, en oposición a aspectos más declarativos de los impactos del conocimiento. Aunque algunas veces puede manifestarse como vinculación, las habilidades que se observan típicamente incluyen un nivel de profundidad y competencia, como involucrarse en indagación científica (observar, clasificar, explorar, cuestionar, predecir o experimentar), así como desarrollar o practicar habilidades muy específicas relacionadas con el uso de instrumentos o aparatos científicos (ej. Utilizar microscopios).
Otras	Específicas a cada proyecto

Fuente (Allen et al., 2008)

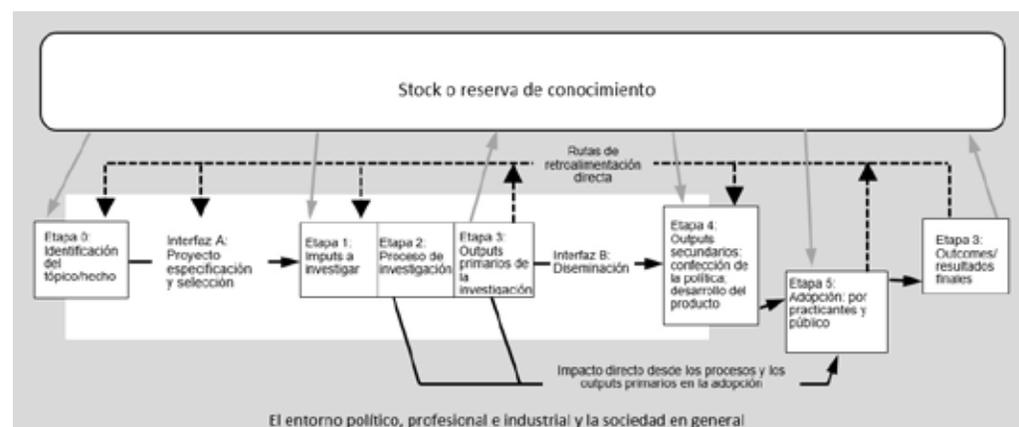
2.1.2 Payback Framework (Reino Unido)

Herramienta desarrollada originalmente para evaluar los impactos de proyectos de investigación en salud por Buxton & Hanney (1996): "Throughout the world there is a growing recognition that health care should be research-led. This strengthens the requirement for expenditure on health services research to be justified by demonstrating the benefits it produces. However, payback from health research and development is a complex concept and little used term. Five main categories of payback can be identified: Knowledge; research benefits; political and administrative benefits; health sector benefits; and broader economic benefits. Various models of research utilization together with previous assessments of payback from research helped in the development of a new conceptual model of how and where payback may occur. The model combines an input-output perspective with an examination of the permeable interfaces between research and its environment. The model characterizes

research projects in terms of Inputs, Processes, and Primary Outputs. The last consist of knowledge and research benefits. There are two interfaces between the project and its environment. The first (Project Specification, Selection and Commissioning, fue implementada a mediados de los años 90 para el Servicio Nacional de Salud del Reino Unido (National Health Service - NHS). Esta evaluación está compuesta de dos elementos: un modelo lógico que representa el proceso de investigación (no lineal), y una serie de categorías para clasificar los distintos impactos de la investigación.

El modelo de impacto propuesto se compone de 6 etapas, para cada una de las cuales es posible diseñar instrumentos de seguimiento y evaluación (Figura 1). En la **etapa 0** ocurren aquellas actividades previas al comienzo del proyecto, como la selección de prioridades por parte de agencias de financiamiento. La “interface” es el proceso de co-diseño en que investigadores y potenciales usuarios definen los objetivos, actividades y formas de comunicación de resultados del proyecto. En la **etapa 1** o inputs se consideran todos los elementos que se incluyen en el desarrollo del proyecto, desde recursos financieros, conocimiento, e interacciones y redes con otros actores del sector salud. La **etapa 2** o proceso, hace énfasis en la rigurosidad científica de la investigación, así como la constante interacción con usuarios. Algunos beneficios ocurren directamente en esta etapa, como cambios de conducta o creación de competencias. La **etapa 3** o de resultados primarios se refiere a la creación de conocimiento, especialmente en publicaciones académicas y de diseminación de resultados (talleres, seminarios, etc.). La **etapa 4** o de resultados secundarios es más compleja de evaluar pues corresponde a todos aquellos impactos entendidos como políticos, administrativos o sociales, como la influencia en la toma de decisiones como metas de salud pública, decisiones sobre contratos administrativos, etc. Son aquellos más difíciles de medir. La etapa 5 o aplicaciones se refiere a los servicios o beneficios económicos de la investigación. Puede ser evaluada en función de cambios de comportamiento. La **etapa 6** o impactos y resultados finales puede ser evaluada a través de análisis de reducción de costos, mejoras en salud etc., poniendo atención a la necesidad de probar la relación entre la investigación y los efectos observados.

Figura 1. Modelo lógico de Payback Framework



Fuente (Donovan & Hanney, 2011) traducción de la autora.

Se definen además cinco categorías de beneficios o “payback” que emergen de la investigación en salud (Buxton & Hanney, 1996) “container-title:”Journal of Health Services Research & Policy,”page:”35-43,”volume:”1,”issue:”1,”abstract:”Throughout the world there is a growing recognition that health care should be research-led. This strengthens the requirement for expenditure on health services research to be justified by demonstrating the benefits it produces. However, payback from health research and development is a complex concept and little used term. Five main categories of payback can be identified: Knowledge; research benefits; political and administrative benefits; health sector benefits; and broader economic benefits.Various models of research utilization together with previous assessments of payback from research helped in the development of a new conceptual model of how and where payback may occur. The model combines an input-output perspective with an examination of the permeable interfaces between research and its environment. The model characterizes research projects in terms of Inputs, Processes, and Primary Outputs. The last consist of knowledge and research benefits. There are two interfaces between the project and its environment. The first (Project Specification, Selection and Commissioning. Estos beneficios pueden ser evaluados y observados a partir de la etapa 3 en adelante del modelo.

1. Conocimiento, ya sea nuevo y/o de relevancia local. Evaluado a través de indicadores bibliométricos⁴ y evaluación por pares.
2. Beneficios para futuras investigaciones y los usos de estas investigaciones. Este aspecto es evaluado a través de análisis de citas, número de doctorados y análisis de redes. Estos beneficios pueden corresponder a:
 - a). Mejoras en la orientación de la investigación
 - b). Desarrollo de competencias, personal y capacidad de investigación
 - c). Desarrollo de capacidad crítica de utilizar apropiadamente la investigación ya existente a nivel internacional
3. Beneficios políticos y administrativos. Evaluado a través de encuestas con tomadores de decisiones⁵ y análisis de documentos.
 - a). Mejor base de información sobre la cual tomar decisiones políticas y administrativas
 - b). Otros beneficios
4. Beneficios al sector salud. Evaluado a través de reducción de costos, encuestas de satisfacción de pacientes, etc.
 - a). Reducción de costos en la entrega de servicios
 - b). Mejoras de calidad en la entrega de servicios
 - c). Aumento de efectividad de servicios y mejoría en la salud de las personas
 - d). Equidad, es decir, mejor asignación de recursos a nivel de área, mayor orientación de objetivos y accesibilidad.

4 Por indicadores bibliométricos se entiende el número de publicaciones, citas, entre otros.

5 Por “tomador de decisiones” se entiende a todos aquellos involucrados en el proceso donde se toman decisiones políticas, ya sea parlamentarios, miembros de ministerios u otras agencias del estado, etc. Decisiones políticas son aquellas tomadas dentro del gobierno con el fin de maximizar el bien social, incluyendo la asignación de presupuestos, aprobación de leyes y regulaciones, etc.

5. Beneficios económicos más amplios. Evaluado a través de metodologías de retorno de la inversión como las utilizadas en negocios.
 - a). Beneficios económicos de la explotación comercial de la innovación que ocurre como consecuencia de I+D
 - b). Beneficios económicos como resultado de una fuerza de trabajo saludable y reducción de los días no trabajados por enfermedad.

Este marco ha sido aplicado a una serie de proyectos de investigación en salud en el Reino Unido. Algunos de los aprendizajes más relevantes son respecto a los periodos de tiempo en los cuales se puede observar impacto (largos); las limitaciones de un modelo “lineal” que no siempre calza con lo que ocurre en la realidad; problemas de atribución, es decir, qué investigación dio origen a cierto impacto y cómo influyeron factores adicionales y la necesidad de categorizar los impactos para evitar dobles conteos.

2.1.3. Social Impact Assessment Methods (SIAMPI) Unión Europea (UE)

Esta metodología fue desarrollada por el programa Social Impact Assessment Methods en el contexto del Marco de Investigación de la Unión Europea FP7 (2007-2013) (Spaapen et al., 2013). Reconoce dos desafíos principales en la evaluación de impacto social: **atribución** de los impactos observados específicamente a actividades de investigación, y **temporalidad**, es decir, el periodo de tiempo entre la investigación y su transformación en productos, procesos o prácticas sociales.

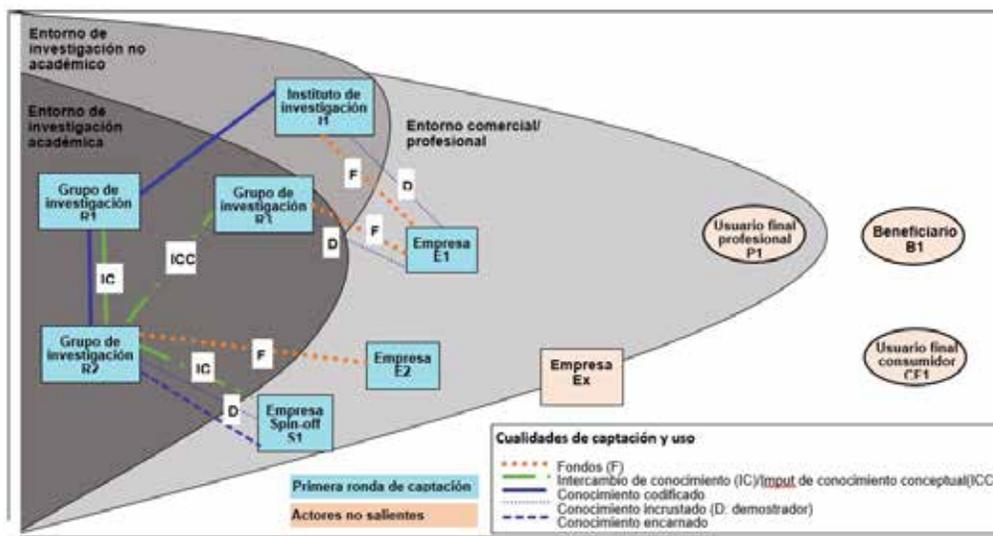
La metodología se basa en la identificación de “interacciones productivas”, es decir, intercambios entre investigadores y otros actores⁶ en los cuales el conocimiento producido y valorado es tanto científicamente robusto como socialmente relevante (Spaapen & Van Drooge, 2011). Este concepto parte del supuesto de que usualmente no existe una clara distinción entre impacto social y una “interacción productiva”, porque la transición entre interacción e impacto es muchas veces gradual. Esto conlleva algunas consideraciones metodológicas:

1. El objeto de la evaluación cambia: en vez de evaluar sólo la entidad de investigación, se evalúa el proceso de investigación en sí.
2. El número de actores crece.
3. Los evaluadores tienen un desafío mayor, pues no sólo deben discutir las temáticas de proyecto con pares, sino que deben incluir un grupo mayor de actores sociales, incluyendo potenciales usuarios de la investigación, intermediarios u otros. Además, se deben incluir evaluaciones cualitativas que capturen el proceso de interacción.

⁶ Por actores se refiere a cualquier grupo o persona involucrada en alcanzar impacto social: investigadores, industria, organizaciones públicas, gobierno, el público en general.

La metodología define tres tipos de interacciones: **directas** [interacciones personales, cara-a-cara o mediada por teléfono, correo electrónico, u otro]; **indirectas** [establecidas a través de algún gestor material, como textos, artefactos, modelos, exhibiciones, etc.] y **financieras** [mediadas por un intercambio económico, por ejemplo, contrato de investigación, contribución financiera o valorizada]. Para cada uno de estos tipos de interacciones, se pueden agregar indicadores cuantitativos y estudios de caso narrativos. Las interacciones pueden ocurrir de forma top-down (coordinadas) o bottom-up (espontáneas) (Figura 2).

Figura 2. Figura explicativa de interacciones productivas para un caso de nanotecnología



Fuente (Spaapen et al., 2013) traducción de la autora

Esta metodología fue utilizada en la evaluación de una serie de casos de los Países Bajos, España, Inglaterra y Francia, y en cuatro disciplinas: nanotecnología, tecnologías de información y comunicación (TICs), ciencias sociales y humanidades y salud. Los autores recomiendan tener en consideración que no son las interacciones productivas en sí mismas las que importan, sino que su rol en el proceso de lograr impacto social. Es decir, cómo las interacciones, ya sean directas, indirectas o financieras, construyen secuencialmente impacto social. No se trata, por ejemplo, de sólo tener muchas instancias de diálogo con usuarios si es que estas no generan coherencia entre sí que permite generar impactos sociales en el largo plazo. Por tanto, es importante evaluarlas en el contexto del propósito y estrategia del centro o proyecto, y de acuerdo al tipo de actores con quienes se generan estas interacciones.

2.1.4. Research Quality Plus RQ+ (Instituto de Desarrollo Internacional de Canadá (IDRC) - Canadá

Esta metodología fue desarrollada por el Instituto de Desarrollo Internacional de Canadá (IDRC), que financia proyectos de investigación científica e innovación que mejoren las condiciones de vida de personas en países en desarrollo. La gran mayoría de esa investigación es llevada a cabo por investigadores pertenecientes a países en desarrollo, en sus países de origen, e incluye disciplinas que van desde la economía a la neurociencia, incentivando el trabajo inter y transdisciplinario (McLean & Sen, 2018). Esta metodología es una forma de evaluar la calidad de la investigación contextualizada y relevante para la realidad local. La metodología reconoce que la excelencia académica es necesaria pero no es el único elemento a evaluar para entender el impacto de un proyecto. Considera las visiones de actores sociales y usuarios para evaluar la legitimidad y relevancia de una investigación, poniendo atención a la forma en que los investigadores posicionan su investigación para hacerla utilizable (Lebel & McLean, 2018).

La aplicación de esta metodología se basa en la revisión por pares expertos de una muestra de productos (publicaciones, reportes, cuadernos de laboratorio, etc.), así como encuestas y entrevistas con informantes claves representantes, incluyendo usuarios de la investigación. El proceso de implementación parte con la selección de la muestra de productos, representativa de todos los años del proyecto. En segundo lugar, los evaluadores caracterizan la investigación en función de sus factores contextuales utilizando una rúbrica. En tercer lugar, los evaluadores determinan la calidad de la investigación; en este caso es importante que la rúbrica utilizada sea clara para los evaluadores, especialmente en relación a cómo será estandarizada para comparar distintos proyectos. Por último, se sintetizan los rankings, primero a nivel de proyecto y luego a nivel de programa.

Esta nueva forma de evaluar los proyectos financiados por IDRC nace de la preocupación por entender y relevar los contextos en que se realiza la investigación, entendiendo que el éxito de esta no es sólo función de su excelencia científica, y que además depende de un grupo de actores mucho más amplio. Más aún, si la evaluación tiene como objetivo entender el impacto de una investigación debe tomar en cuenta el grado de control e influencia de parte de los investigadores. La influencia que pueda tener una investigación sobre hacedores de políticas u otros actores sociales puede no estar del todo en la esfera de control de los investigadores. Por su parte, el impacto a largo plazo, aunque relevante, está muy por sobre la esfera de control de los investigadores (Figura 3) (Ofir, Schwandt, Duggan, & McLean, 2016).

Figura 3. Esferas de investigación e impacto



Fuente: Ofir et al., 2016. Traducido por la autora

La metodología se basa en tres postulados⁷:

1. Aceptar una visión multidimensional de calidad basada en los valores y objetivos que guían la agenda de investigación. El rigor científico es un elemento central y no negociable de la evaluación de la investigación, sin embargo, existen dimensiones subjetivas y propias de cada agencia o programa de financiamiento.
2. La investigación ocurre en un contexto organizacional, territorial y social específico, es necesario aceptarlo y aprender de ello. Las evaluaciones tradicionales buscan aislar a los proyectos de su contexto de desarrollo; al contrario, al considerar el contexto es posible realizar una evaluación sistémica de la investigación y de la producción de conocimiento.
3. Las evaluaciones deben estar basadas en evidencia empírica, no sólo en opiniones. Esta evidencia puede ser recolectada a través de consultas con beneficiarios de la investigación, indicadores bibliométricos, expertos del área, etc.

⁷ Cada uno de estos postulados se relaciona con el Manifiesto de Leiden, una guía de 10 principios de buenas prácticas para evaluaciones de la investigación (Hicks, Wouters, Waltman, De Rijcke, & Rafols, 2015).

La evaluación en sí se compone de tres elementos (1) Dimensiones y sub dimensiones de calidad de la investigación (2) factores contextuales y (3) rúbrica de evaluación (Figura 4) (Ofir et al., 2016).

Figura 4. Componentes de la evaluación RQ+



Fuente: Ofir et al., 2016. Traducido por la autora.

1. Dimensiones y sub dimensiones de calidad de la investigación

1.1. Integridad: considera la calidad técnica, pertinencia y rigor del diseño y ejecución de la investigación, juzgado en base a estándares y métodos aceptados comúnmente para estas dimensiones. Son evaluadas en función de documentos del proyecto y productos de la investigación seleccionados, con énfasis en el diseño de la investigación, el rigor metodológico, la revisión de la literatura, y la relación entre la evidencia obtenida y las conclusiones de una investigación

-
- 1.2. Legitimidad:** considera en qué medida el proceso de investigación fue justo y tomó en consideración todas las preocupaciones, conocimiento, valores y perspectivas de actores relevantes. Tiene que ver con quién participó y quién no, el proceso de toma de decisiones, cómo fue producida y diseminada la información, cómo fue embebido localmente el conocimiento y si fueron respetadas las tradiciones y sistemas de conocimiento locales. Tiene cuatro sub dimensiones (i) abordar consecuencias negativas (ii) perspectiva de género (iii) inclusividad y (iv) involucramiento con el conocimiento local.
- 1.3. **Importancia:** considera la importancia y el valor para los usuarios del conocimiento y la comprensión generada por una investigación, en relación a los productos o procesos que serían relevantes para estos usuarios, y la contribución a la teoría o práctica. Tiene dos sub dimensiones (i) originalidad y (ii) relevancia.
- 1.4. Posicionamiento para el uso:** considera el grado en que el proceso de investigación ha gestionado los productos o resultados para aumentar su probabilidad de uso e impacto. Esta dimensión fue incorporada bajo el entendimiento de que la investigación es un proceso inherentemente político, que requiere establecer relaciones y seleccionar las mejores plataformas para poner a disposición los productos de una investigación.
2. Factores contextuales. Son factores que pueden influenciar de manera positiva o negativa el proceso y los resultados de investigación. Estos son la madurez del área de investigación [el grado en que los marcos teóricos y conceptuales y las hipótesis de un área han sido testeadas y establecidas], fortalecimiento de la capacidad de investigación [el grado en que el proyecto fortalece las capacidades para concebir, conducir, gestionar y comunicar una investigación], riesgo en el ambiente de investigación [es decir, el grado en que el contexto organizacional facilita la investigación], riesgo en el ambiente político [inestabilidad política, conflicto armado, crisis humanitarias, u otro riesgo externo que pueda emerger debido al contexto político y de gobernanza]; y riesgo en el ambiente de datos [el grado en que existe instrumentación adecuada y es posible realizar una toma de datos, incluyendo si el proyecto se desarrolla en un contextos rico o pobre en datos]
3. Rúbrica de evaluación, es decir, un set de criterios para los evaluadores que describe cómo debe ser evaluada cada dimensión y subdimensión. Las subdimensiones de la dimensión de calidad de la Investigación se evalúan de 1 a 8, donde 1-2 [inaceptable] 3-4 [menos que aceptable], 5-6 [aceptable o bueno] y 7-8 [muy bueno]. Los puntajes para cada subdimensión son agregados para llegar a una calificación total de la dimensión. Esta rúbrica se describe en detalle para cada subdimensión (Tabla 2). En el caso de la dimensión de factores contextuales, se utiliza una escala de 1 a 3, donde 1 es bajo, 2 es medio y 3 es alto [entendido como grado de influencia en la investigación].

Tabla 2. Ejemplo de rúbrica de Evaluación para la dimensión Legitimidad de la investigación: involucramiento del conocimiento local

No aplica	Inaceptable	Menos que aceptable	Aceptable a Bueno	Muy bueno
	1 - 2	3-4	5-6	7-8
La naturaleza de la investigación es tal que el conocimiento y la vinculación local no necesitan ser tomadas en consideración	El involucramiento con los contextos locales ha sido desatendido en la investigación. Se pueden encontrar múltiples debilidades, relacionadas con la forma en que las necesidades de investigación y preguntas fueron identificadas, la forma en que se involucró a las comunidades locales, en que los contextos y conocimientos locales fueron incluidos, y en la forma de asegurar los beneficios locales generados a partir del proceso de investigación	Los contextos locales y el involucramiento han sido considerados en la investigación, pero persisten algunas debilidades relacionadas con la forma en que las necesidades de investigación son identificadas, la forma en que se involucró a las comunidades locales, en que los contextos y conocimientos locales fueron incluidos, y en la forma de asegurar los beneficios locales generados a partir del proceso de investigación	Los contextos locales e involucramiento han sido parte del proceso de investigación. Existen debilidades menores relacionadas con la forma en que las necesidades de investigación son identificadas, la forma en que se involucró a las comunidades locales, en que los contextos y conocimientos locales fueron incluidos, y en la forma de asegurar los beneficios locales generados a partir del proceso de investigación	Los contextos locales e involucramiento han sido un foco claro y sistemático en el proceso de investigación. Las necesidades y preguntas de investigación fueron identificadas de forma apropiada, las comunidades locales fueron involucradas de forma apropiada, los contextos y conocimientos locales fueron incluidos, y se aseguraron los beneficios locales generados a partir del proceso de investigación

Fuente: McLean & Sen, (2018), traducido por la autora.

Esta metodología fue utilizada para evaluar los proyectos financiados por IDRC. Se analizaron 130 proyectos (2010-2015) individualmente bajo la metodología RQ+, y luego se condujo un análisis estadístico de los resultados generales. De ellos se concluye que existe una alta correlación entre la integridad científica y la utilidad de una investigación, contrario a la creencia común. Esto también se asocia a la creación de capacidades. Por otro lado, no se encontró una correlación importante entre los factores contextuales, por lo que se sugiere tomar en consideración caso a caso el contexto en que se asignan fondos y evalúan proyectos. Sin embargo, conocer el contexto en que se desarrolla una investigación ayuda a aminorar los riesgos asociados a éste. Por último, demuestra que la investigación más innovadora está siendo realizada por aquellos que son nuevos en un área (McLean & Sen, 2018).

2.2 Evaluaciones de centros, organizaciones y/o programas

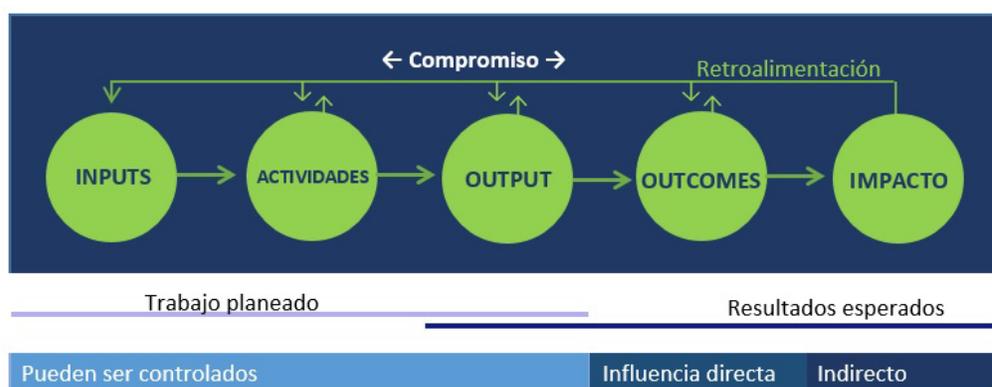
2.2.1. ASIRPA (INRA - Francia)

Esta metodología, denominada Análisis Socio-económico de los Impactos de Investigación Pública en Agricultura, ha sido desarrollada por el Instituto Francés de Investigación en Agricultura (INRA), y consiste en estudios de casos de impacto *ex post* con el objetivo de promover el aprendizaje y rendición de cuentas dentro de la institución. Los estudios de caso son estructurados en función de la teoría de innovación de Callon (1986), de la “traslación” y con bases en la teoría del actor red.⁸ Los supuestos utilizados para implementar esta metodología son que el impacto es (i) multidimensional (ii) basado en el involucramiento de una red de actores (iii) que están presentes en distintas etapas y juegan distintos roles (iv) en un proceso de creación de impacto que es no-lineal (Joly et al., 2015).

La metodología fue implementada utilizando una base de datos de todos los proyectos realizados por INRA en el periodo 1996-2011. Estos proyectos fueron seleccionados de acuerdo a criterios que reflejaran la misión y actividades de investigación de INRA, además de alto impacto y estar basados en ciencia de excelencia. Se realizó una cronología de cada proyecto, en la que se describen inputs, outputs y el tipo de interacciones que se van creando a través del tiempo. Uno de los objetivos de esta cronología es evitar la “falacia de proyecto”, es decir, la expectativa de las agencias de financiamiento de atribuir los resultados a un solo fondo o contrato, siendo que la mayoría de las veces un fondo es uno de múltiples contribuciones a proyectos de mayor tamaño y duración (Georghiou & Clarysse, 2006). Esto también aborda el problema de que una parte importante de los impactos de la investigación ocurren a largo plazo, por tanto, no son capturables en una evaluación. Es necesario entonces evaluar impactos intermedios.

Basado en el *impact pathway* desarrollado por Walker et al., (2008), se agregan los conceptos de configuración productiva [todas las inversiones financieras, humanas y físicas realizadas por INRA y sus socios académicos y socio-económicos para la producción de conocimiento científico y tecnológico] e intermediarios [que operan a distintos niveles para hacer productivo el conocimiento, como oficinas de transferencia y licenciamiento (OTL), start-ups, agencias de estandarización, etc.]. Además, considera no sólo los impactos inmediatos y más visibles, sino que se enfoca en reconocer impactos visibles en el grupo objetivo (Impacto 1) e impactos mayores que ocurren con el escalamiento, horizontal o vertical, de los primeros impactos (Impacto 2). A su vez, definen 5 tipos de impacto: económico, sanitario, socio-territorial, medioambiental y político (Figura 5).

8 La teoría del actor o sociología de la translación es enfoque de los estudios de ciencia, tecnología y sociedad que considera la agencia de humanos y objetos no-humanos y discursos, y con atención a las redes que se establecen en la producción de conocimiento. Traslación es el proceso por el cual una “red”, es decir, una combinación de actores humanos, tecnologías, discursos se presenta como una entidad única. Se compone de cuatro procesos (i) definir un problema común que tiene que ser resuelto o problematización (ii) interés e involucramiento de actores (interés) (iii) definir los roles de los distintos actores involucrados (enrolamiento) e (iv) ir más allá del primer grupo de actores (movilización) (Latour, 2005).

Figura 5. Modelo de impacto ASIRPA

Fuente: Joly et al., 2015. Traducción de la autora.

La metodología ASIRPA enfatiza el rol del contexto (científico, regulatorio y económico) en cada uno de los estadios del proceso de generación de impacto, la importancia de la creación de redes sólidas, la existencia de al menos dos esferas de impacto (referidas como Impacto 1 e Impacto 2), y la gran diversidad de tipos de impacto que se pueden observar en un proyecto.

2.2.2. Commonwealth Scientific and Industrial Research Organisation CSIRO (Australia)

Las evaluaciones de impacto de CSIRO tienen como objetivo principal el demostrar, con evidencia, a actores relevantes [gobierno, directores de CSIRO, investigadores y área de negocios de CSIRO, y el público australiano], que el objetivo de la institución está siendo cumplido a través de sus actividades de investigación e innovación.

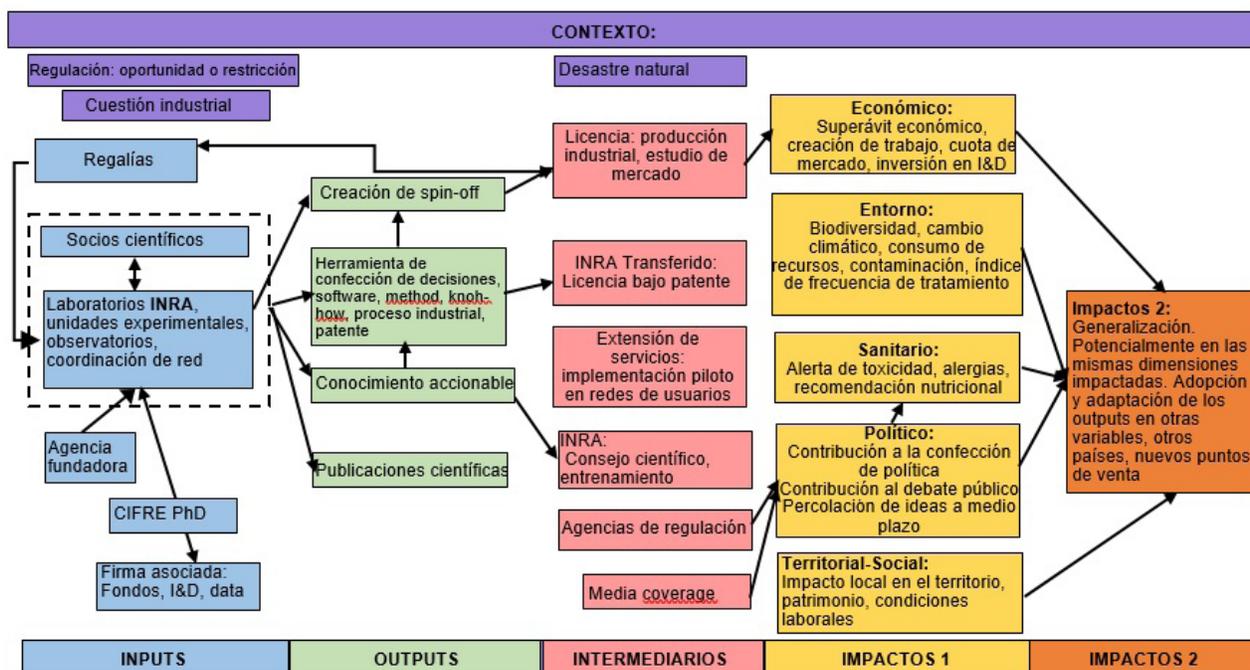
La definición de impacto utilizada es “un efecto, cambio o beneficio en la economía, sociedad y medio ambiente, más allá de las contribuciones del conocimiento académico”. Para las evaluaciones de impacto de CSIRO, se considera impacto los efectos del trabajo de esta agencia que han sido generados *después* de que dicho trabajo ha sido adoptado (CSIRO, 2015). La metodología está compuesta de seis pasos

Paso 1: Establecer el propósito y la audiencia. Establecer el propósito de la evaluación, pues de ello depende la selección de métodos, datos y escala de la evaluación.

Paso 2: Identificar los impactos. Para que un impacto sea considerado, el flujo que le da origen debe estar claramente relacionado con las actividades de CSIRO. Los programas de CSIRO emergen en respuesta a un desafío [necesidad local o nacional, nueva tarea o competencia, una innovación requerida por empresas, un objeto de investigación de parte del grupo de CSIRO], y el impacto descrito debe ponerse en este contexto, con el detalle adecuado. Cada unidad de evaluación debe haber producido uno o más impactos (efectivos o anticipados), los que deben ser determinados de manera separada. Se reconocen las categorías de impacto económicos, sociales y ambientales. Luego, es necesario identificar el

flujo que genera un impacto. Esta identificación se basa en el Marco de Impacto de CSIRO (Figura 6). El flujo de impacto puede ser complejo, e idealmente debe haber sido planificado utilizando alguna herramienta de gestión de programas adecuada al comienzo del proyecto que tome en consideración impacto, como árbol de decisiones o diagramas de causa y efecto.⁹ Esto permite además evaluar que información es necesario coleccionar con anterioridad para la evaluación.

Figura 6. Impact Pathway



Fuente: CSIRO, 2015

Paso 3: Clarificar los impactos. En esta etapa la narrativa del impacto es profundizada ¿qué habría pasado si CSIRO no estuviese involucrado?, las contribuciones realizadas por otras entidades, y cuánto del impacto anticipado aun ocurre. Todos los impactos son evaluados en comparación con un “contrafactual”, es decir, un punto de referencia que representa lo que habría ocurrido en la ausencia de intervención de CSIRO. Luego, es necesario evaluar cuanto del impacto es atribuible a CSIRO, en comparación con otras organizaciones que hayan participado del proyecto. Para ello se utiliza como referencia el porcentaje del costo total del proyecto por cada institución. También en esta etapa se considera la adopción de

9 http://www.cyta.com.ar/biblioteca/bddoc/bdlibros/herramientas_calidad/causaefecto.htm; <https://www.lucidchart.com/pages/es/que-es-un-diagrama-de-arbol-de-decision>

las innovaciones desarrolladas por CSIRO, incluso si los sistemas aún no han madurado¹⁰.

Paso 4: Evaluar los impactos. CSIRO usa métodos mixtos de evaluación que incluyen evaluación costo-beneficio, para aquellos impactos monetarios; evaluación no-monetaria para otros tipos de impacto; cuantificación no monetaria para aquellos impactos donde existen los datos adecuados (ej. Ambientales); y métodos cualitativos (encuestas, entrevistas, focus group) para los impactos restantes. Es necesario estimar los costos reales del proyecto, los flujos económicos incluyendo externalidades y los efectos distribucionales (grupos que se benefician positiva y negativamente). Para las evaluaciones de CSIRO se utiliza una tasa de descuento del 7%. Es importante que todos los procesos, supuestos y metodologías de evaluación queden bien documentadas.

Paso 5: Agregación y comparabilidad de impactos a través de distintos programas. En el caso de que sea posible, entregar indicadores de costo-beneficio. A nivel de agregación de proyectos es especialmente importante que los supuestos de trabajo se mantengan constantes de manera de poder agregar y/o comparar impactos, según sea el caso.

Paso 6: Análisis de sensibilidad y reporte. Se recomienda en general compartir y comentar los métodos de evaluación con todos los actores interesados. Se recomienda además realizar un análisis de sensibilidad que permita estimar la incertidumbre asociada a las estimaciones de costo-beneficio. Los reportes de evaluación son estructurados de la misma forma en que las evaluaciones fueron llevadas a cabo. Todas las evaluaciones son entregadas y comunicadas a la audiencia interesada en un reporte, por lo que se sugiere también contar con un plan de comunicación de resultados.

2.2.3. SAEI (*Small Advanced Economies Initiative*)

La Asociación de Pequeñas Economías Avanzadas (SAEI) es una colaboración entre Irlanda, Dinamarca, Nueva Zelanda, Finlandia e Israel, con foco en Ciencia, Tecnología e Innovación¹¹, entre otras áreas. En el año 2015 fue presentada una propuesta relacionada con evaluación de impacto, relevante para programas, centros y agencias de investigación de estos países, con el objetivo de comparar experiencias y compartir aprendizajes para países de tamaño y condiciones económicas similares (Harland & O'Connor, 2015).

El foco de esta propuesta son los centros y/o programas de cierta escala (nivel de inversión medio o alto) y duración (continuidad en el tiempo), con interés político demostrado y donde la atribución de dicho impacto es clara. En este contexto, impacto es definido como “La ‘influencia’ directa o indirecta de la investigación o sus ‘efectos sobre’ un individuo, comuni-

10 En estos casos se utiliza el análisis de opciones reales, para evaluar el valor opcional de una investigación en vez de su impacto. Corresponde al valor presente de la investigación, pero no en función de impacto, sino en base a las opciones futuras de utilizarlo en un periodo más avanzado.

11 Incluye a la Fundación de Ciencia de Irlanda (Science Foundation Ireland), la Agencia para la Ciencia, Tecnología e Innovación de Dinamarca (Danish Agency for Science), La Comisión de Educación Superior del Ministerio de Economía, Innovación y Empleo de Nueva Zelanda, la Academia de Ciencias de Finlandia y el Comité de Planificación y Presupuesto del Consejo de Educación Superior de Israel.

dad o la sociedad en general, incluyendo los beneficios a nuestro capital económico, social, humano y natural” (Harland & O’Connor, 2015; pg.6, traducción del autor).

Esta amplia definición busca incluir todos los efectos de la investigación, incluyendo la creación de conocimiento y capacidades, y tomando en consideración a distintas disciplinas. Impacto también es distinto de relevancia, pues considera la implementación de los resultados de investigación para su uso. El impacto generalmente se asocia a un objetivo, que está sujeto a una evaluación valórica que define su importancia para la sociedad. Este objetivo además depende de la función de cierto programa o agencia: una agencia dedicada a la conservación tendrá impactos esperados muy distintos a una dedicada a la transferencia tecnológica.

Se definen seis áreas o pilares de impacto con tres tipos de impacto transversales. Las áreas de impacto son:

1. Económico, incluyendo la creación de empleo
2. Salud y bienestar
3. Medio ambiente
4. Política pública, servicios y regulación
5. Capacidades humanas
6. Social y vinculación internacional, incluyendo el rol cultural y de identidad nacional asociado a la ciencia y la investigación

Las áreas transversales definidas son

1. Creación de nuevos productos, procesos, políticas y comportamientos
2. Mejorar la eficiencia y eficacia de prácticas, políticas y procesos existentes
3. Mejorar la resiliencia, sustentabilidad y reducción de riesgos de sistemas (ya sean económicos, naturales, de salud o de la sociedad en general).

Tabla 3. Los 6 pilares de impacto propuestos, y las 3 temáticas transversales

Económico	Salud y Bienestar	Capital Natural y ambiente construido	Política pública y servicios	Capacidades y competencias futuras	Sociedad y relaciones internacionales
Crear nuevos productos, procesos, políticas o comportamientos					
Mejorar la eficiencia y Eficacia de prácticas existentes					
Investigación que mejora la resiliencia y/o sostenibilidad					

Fuente: Harland & O’Connor, 2015.

La metodología propuesta busca generar una serie de prácticas comunes para las distintas agencias que participan de esta iniciativa, entendiendo que cada una de ellas tiene distintos procesos y objetivos. El propósito de esta metodología es motivar a los investigadores a considerar el impacto en todas las dimensiones, más allá de las propuestas por el proyecto o aquellas que sean más evidentes, y apoyar a las agencias en la generación de medios que permitan generar estrategias de impacto en los proyectos y evaluarlo.

La metodología propuesta sugiere analizar estos componentes, a través de una matriz, en base a las características y objetivos de cada centro o programa. No todas las áreas de impacto tendrán la misma importancia para un proyecto. Por ejemplo, un programa en biotecnología ambiental tendrá mayor posibilidad de generar impacto en las áreas medio ambiente y económica, y esto permite enfocar los esfuerzos de impacto en dicha dirección. El resto de los pilares no quedan excluidos, pero se especifica que los impactos pueden ocurrir de manera más indirecta y son menos relevantes para los objetivos del proyecto, por ejemplo, la creación de capacidades relevantes. Un ejemplo de estas distintas dimensiones se detalla en la Tabla 4. Para cada uno de los impactos especificados se sugiere declarar también el periodo de tiempo en que estos impactos pueden ser esperados (corto, medio o largo plazo).

Tabla 4. Ejemplos de uso de matriz para identificar potenciales impactos de distintos centros de investigación

Centro (ficticio)	Económico	Salud y Bienestar	Capital Natural y ambiente construido	Política pública y servicios	Capacidades y competencias futuras	Sociedad y relaciones internacionales
Centro de Biocombustibles	Potencial de la creación de un nuevo biocombustible	Reducción de emisiones por transporte; mejora en resiliencia y seguridad energética		Mayor comprensión del ciclo de vida de las emisiones, evidencia para el uso en política en estándares e incentivos para mezclas de combustibles	Conocimiento del público en relación al uso apropiado de los biocombustibles y sus mezclas en vehículos	Contribución al desafío global del cambio climático
Centro de Biodiversidad		Mejorar la resiliencia de especies en riesgo			Vinculación con el público en la prevención de pestes y especies invasoras	Proteger la herencia nacional y las especies de importancia cultural

Fuente(Harland & O'Connor, 2015). El color verde indica aquellos impactos más relevantes para cada centro, luego le siguen los en color azul y los impactos más indirectos se indican en color celeste.

Es importante que asociadas a cada una de las declaraciones de impacto generales, se indique claramente cuáles serían impactos específicos y cómo medirlos. La ventaja de este método es que permite agrupar y/o comparar los impactos de distintos centros o programas, incluso aquellas evaluaciones de carácter cualitativo, al poder agruparlas en distintas dimensiones.

2.3 Evaluaciones a nivel nacional

Este tipo de evaluaciones a nivel nacional o sistemas de financiamiento basados en el desempeño (Performance-based research funding systems - PRFS) se utiliza para decidir el financiamiento que se asignará a distintas instituciones de investigación dependientes de fondos públicos como universidades y centros e institutos de investigación. La mayoría de los países europeos han implementado este tipo de evaluación, aunque sólo los Países Bajos y el Reino Unido consideran el criterio de impactos no académicos de forma rutinaria en las evaluaciones ya realizadas¹². La Comisión Europea actualmente busca implementar una metodología común a todos los países miembros para evaluaciones de este tipo (Spaapen, 2018).

De acuerdo a Sivertsen, (2017) los países europeos pueden ser divididos en cuatro tipos, de acuerdo a la relevancia que tienen los indicadores cuantitativos y bibliométricos, versus evaluaciones cualitativas, en este tipo de ejercicios.

1. El propósito de la asignación de fondos se combina con una evaluación de la investigación, la cual es organizada en intervalos de varios años y se basa en juicio de pares expertos. Los datos cuantitativos se usan para informar la evaluación [Italia, Lituania, Portugal y el Reino Unido].
2. La asignación de fondos se basa en una serie de indicadores cuantitativos de investigación, los cuales se usan anualmente para orientar la inversión en investigación [Croacia, República Checa, Polonia y Suecia].
3. Similar al caso anterior, pero en este caso los indicadores incluyen otras actividades universitarias, además de la investigación [Flandes (Bélgica), Dinamarca, Estonia, Finlandia, Noruega y Eslovaquia].
4. Similar al caso (iii) pero no se incluyen indicadores bibliométricos dentro del set de indicadores [Austria, Países Bajos].

12 Otros países han desarrollado propuestas para incluir este tipo de evaluación, como es el caso de Suecia y Australia, pero no han sido incluidos en sus evaluaciones a la fecha. Spaapen (2017) menciona que Croacia, Estonia, Italia y Turquía incluyen indicadores de impacto en algunas de sus evaluaciones, pero sólo cuantitativos.

En todos estos casos, se combinan las funciones de evaluación de la investigación y de distribución de fondos públicos, pero varía el grado de relevancia de cada uno de estos componentes. En algunos casos, como el del Reino Unido, estos dos objetivos están estrechamente relacionados y la evaluación de la investigación se traduce directamente en fondos públicos; sin embargo, en casos como el de los Países Bajos, la evaluación de la investigación tiene un rol formativo, pero no se traduce directamente en distribución de fondos públicos hacia a las instituciones de investigación, los cuales son asignados por otras vías.

A continuación, se describen los sistemas de evaluación PRFS, habiendo seleccionado aquellos casos de países donde se tome en consideración el criterio de impacto no académico.

2.3.1. Research Impact Framework REF14 (Reino Unido)

El marco de evaluación de impacto de la investigación (REF) es una evaluación nacional, organizada por el Consejo de Educación Superior de Inglaterra y Gales (HEFCE), el Consejo Escocés de Financiamiento (SFC), y el Departamento de Empleo y Aprendizaje de Irlanda del Norte (DEL). Se realiza desde el año 1986, y cada seis años, con el objetivo de asignar fondos de investigación basado en el desempeño. La evaluación incluye las dimensiones calidad de los resultados de investigación (65%), ambiente donde se desarrolla la investigación (15%) e impacto (20%). El año 2014¹³ fue el primer año que el criterio impacto fue incluido en la evaluación.

Para esta evaluación, impacto fue definido como “un efecto, cambio o beneficio en la economía, la sociedad, la cultura, las políticas o servicios públicos, la salud, el medio ambiente o la calidad de vida, más allá de la academia. El impacto incluye un efecto, cambio o beneficio en:

- » La actividad, actitud, toma de conciencia (awareness), comportamiento, capacidad, oportunidad, desempeño, política, práctica, proceso o entendimiento
- » de una audiencia, beneficiario, comunidad, organización o individuo
- » en cualquier ubicación geográfica ya sea local, regional, nacional o internacional

El impacto incluye la reducción o prevención de daños, riesgos, costos u otros efectos negativos” (Hill, 2016). Los impactos fueron evaluados en dos dimensiones: alcance (cuan ampliamente son los impactos) y significancia (cuan transformativos han sido los impactos).

La evaluación se realizó en función de 36 áreas de investigación divididas en 4 paneles (Tabla 5), en las cuales se definieron “unidades de evaluación” de acuerdo al número de investigadores por áreas de investigación de cada institución. Para cada evaluación se consideraron dos elementos narrativos, un “formulario de impacto” que describe la aproximación y estrategia de impacto de la unidad de investigación. El segundo elemento es un número de estudios de caso que ilustraran ejemplos de impacto ocurridos durante el periodo de referencia, que incluyera (i) resumen del impacto (ii) investigación asociada (iii) referen-

13 El periodo de referencia utilizado para esta evaluación es 2008-2013.

cias de la investigación (iv) detalles del impacto y (v) fuentes para corroborar el impacto. El número de casos por unidad dependió el tamaño de la unidad, con 1 caso por cada 10 investigadores.

Tabla 5. Áreas de investigación para la evaluación REF

Panel Principal	Unidad de Evaluación
A	1 Medicina Clínica
	2 Salud Pública, Servicios de Saludo y Cuidados primarios
	3 Profesiones de la salud asociadas, odontología, enfermería y farmacia
	4 Psicología, psiquiatría y neurociencia
	5 Ciencias Biológicas
	6 Ciencias Agrícolas, Veterinarias y de los alimentos.
B	7 Ciencias Ambientales y de la tierra
	8 Química
	9 Física
	10 Matemáticas
	11 Informática y Ciencias de la computación
	12 Ingeniería
C	13 Arquitectura, urbanismo y planificación
	14 Geografía y Estudios Medioambientales
	15 Arqueología
	16 Economía y Econometría
	17 Administración y Negocios
	18 Leyes
	19 Estudios Políticos e Internacionales
	20 Trabajo social y política social
	21 Sociología
	22 Antropología y estudios del desarrollo
	23 Educación
	24 Ciencia del deporte y estudios de Turismo

Panel Principal	Unidad de Evaluación
D	25 Estudios de área
	26 Lingüística y lenguajes modernos
	27 Lengua y literatura inglesa
	28 Historia
	29 Estudios clásicos
	30 Filosofía
	31 Teología y estudios religiosos
	32 Arte y Diseño: Historia, práctica y teoría
	33 Música, Teatro, Danza, Artes performativas, Cine
	34 Estudios de la comunicación, medios y culturales; bibliotecnología y gestión de la información

Fuente: REF 2014

El impacto fue evaluado por un grupo de expertos de cada área, compuesto de 1052 expertos, 77% de los cuales eran académicos y 23% usuarios (individuos que hacen uso de la investigación universitaria y colaboradores fuera de la academia en el sector privado o público). Las evaluaciones fueron graduadas en una escala de cuatro puntos, donde 4 corresponde a liderazgo mundial, 3 excelencia internacional, 2 reconocido internacionalmente y 1 reconocido nacionalmente. También se incluyó la evaluación “inclasificable”. Es importante recordar que los resultados de esta evaluación se traducen en financiamiento público para las universidades evaluadas.

Los resultados de esta evaluación fueron hechos públicos el año 2015¹⁴, y desde entonces ha tenido una serie de reflexiones, impresiones y propuestas de mejora. Algunos de los aspectos más discutidos fueron:

1. La ambigüedad de la definición de impacto y sus criterios de evaluación (alcance y significancia), lo que significó una dificultad para los paneles evaluadores y en la definición de criterios comunes. Se enfatiza la importancia de las instancias de preparación y deliberación de los grupos evaluadores previo a la evaluación misma (Derrick & Samuel, 2017)
2. Los tipos de incentivos para la labor de investigación que implica una evaluación de impacto, y cómo evitar que ésta no genere incentivos perversos, como por ejemplo privilegiar investigación con impactos a corto plazo. De la misma forma, tomar en consideración qué tipos de impactos son deseables, y para quién (Hill, 2016; Smith & Stewart, 2017).
3. Incluir casos de impacto a nivel institucional para destacar el impacto generado por el trabajo multi e interdisciplinario (Stern, 2016).

14 Los resultados pueden ser accedidos en <https://www.ref.ac.uk/2014/>

2.3.2. *Piloto de Evaluación de Impacto y Vinculación (Australia)*

A comienzos del 2017, el Consejo de Investigación de Australia (ARC) condujo un piloto de una nueva metodología de evaluación de unidades de investigación públicas, que incluye las dimensiones de vinculación e impacto no académico. Esta metodología fue inspirada por REF14, pero a diferencia de REF evalúa de forma independiente las dimensiones de impacto y vinculación. La definición de vinculación utilizada es “la vinculación de la investigación es la interacción entre investigadores y usuarios finales de la investigación para intercambio de conocimiento, tecnologías y métodos, y recursos” (Australian Research Council, n.d.). La definición de impacto utilizada es “el impacto de la investigación es la contribución que la investigación hace a la economía, sociedad, medio ambiente y cultura, más allá de la contribución a la investigación académica”. Se entiende como usuario a “un individuo, comunidad u organización externa a la academia que usará directamente o se beneficiará directamente del output, producto o resultado de la investigación”¹⁵.

El piloto fue diseñado en función de una evaluación retrospectiva, que incluía a todas las universidades y centros de investigación alojadas bajo el Higher Education Act (42 universidades en total), todas las disciplinas presentes. El piloto, desarrollado entre Diciembre del 2016 y Agosto del 2017 contó con la participación de 39 universidades, y se basó en 94 casos de vinculación en 4 disciplinas¹⁶, y 200 casos de impacto, en 6 disciplinas además de estudios¹⁷ indígenas e interdisciplinarios.

El piloto buscó testear tanto la metodología narrativa como indicadores cuantitativos tanto para vinculación como para impacto. Cada evaluación de vinculación se basa en una serie de indicadores cuantitativos, una declaración explicativa de los indicadores (750 palabras) y una narrativa de vinculación (1000 palabras). Los indicadores de vinculación seleccionados tras el piloto son: apoyo financiero de parte de usuarios finales, gasto en I+D de las universidades (HERD) por investigadores jornada equivalente (FTE) (esquemas específicos),¹⁸ los fondos auspiciados por usuarios finales (como proporción de HERD) y los ingresos por comercialización de la investigación. Además, se incluyen los indicadores opcionales: co-autoría de productos de investigación con usuarios finales, co-financiamiento de investigación con usuarios finales, patentes otorgadas, citaciones a patentes, apoyo no pecuniario de parte de usuarios, resultados disponibles vía acceso abierto.

15 Es importante notar que esta definición excluye a organizaciones financiadas con fondos públicos como CSIRO, a otras organizaciones de educación superior, organizaciones afiliadas, controladas o filiales de una organización de educación superior (como hospitales dependientes de una universidad) y a sus equivalentes nacionales e internacionales.

16 Ciencias Químicas, Ciencias Médicas y de la Salud, Historia y Arqueología y Filosofía y Estudios de la Religión.

17 Ciencias Ambientales, Ciencias Agrícolas y agropecuarias, Ingeniería, Educación, Estudios de artes creativas y escritura, y Lenguaje, comunicación y cultura.

18 HERD (HIGHER EDUCATION EXPENDITURE ON R&D) es el gasto en I+D de las Instituciones de Educación Superior; la sigla FTE (Full-Time Equivalent) es una forma de generar equivalencia de horas trabajadas por empleados partime con contratos de tiempo completo.

El piloto de impacto consistió en evaluación de una narrativa de impacto por unidad de análisis, que incluye información sobre la forma en que la unidad aborda el impacto (mecanismos), un ejemplo de impacto (la naturaleza y el grado del impacto) y la investigación asociada a dicho impacto. El periodo de referencia utilizado fue de 6 años (01/2011 a 12/2016).

Para evaluar estos casos de impacto se consideraron los siguientes aspectos: la forma en que la institución apoya a las unidades de análisis, cómo los investigadores interactúan y se involucran con usuarios, la evidencia de que los mecanismos de traslación de la investigación fueron integrados a las prácticas académicas, y la articulación entre la aproximación a impacto y el ejemplo presentado. La estrategia de impacto y el caso de impacto se evalúan de forma separada.

En ambos casos, la evaluación es asignada en una escala de tres puntos: bajo, medio y alto, estando a cargo de paneles divididos por disciplinas, similares a los utilizados en REF (ver Tabla 5). El objetivo de este piloto fue entender qué aspectos de la evaluación funcionan y las modificaciones requeridas para aplicar la metodología como parte del sistema de evaluación y asignación de fondos de investigación de Australia ERA (Excellence in Research for Australia). Uno de los principales resultados fue identificar aquellos indicadores cuantitativos adecuados para apoyar las evaluaciones narrativas de impacto y vinculación.

Tabla 6. Indicadores de vinculación seleccionados a partir del piloto

Obligatorio	Opcional
Apoyo en dinero por parte de usuarios finales	Co-autoría de resultados de investigación con usuarios finales
Total de ingresos en I+D de las universidades por jornada equivalente	Co-financiamiento de outputs de investigación con usuarios finales
Fondos de investigación apoyados por usuarios finales: proporción de HERD categoría 1	Patentes generadas
Ingresos por la comercialización de la investigación	Citas de patentes
	Aporte no pecuniario de usuarios finales
	Outputs disponibles vía acceso abierto

Fuente: Australian Research Council, n.d.

2.3.3. Research quality evaluation in Sweden – FOKUS (Suecia)¹⁹

Esta metodología corresponde a una propuesta comisionada por el Gobierno de Suecia al Consejo de Investigaciones de Suecia, en relación a un modelo para la asignación de recursos a universidades que involucrara revisión por pares de la calidad y relevancia de la investigación (Monaco et al., 2015). Esta propuesta reemplazaría al modelo actual de asignación de fondos basado en indicadores,²⁰ con el objetivo de relevar y promover aspectos como el impacto de la investigación.

Para el desarrollo de este modelo fueron estudiados los casos de Australia, Reino Unido y los Países Bajos. El modelo incluye evaluaciones en tres aspectos: calidad científica/artística (70%), factores que promueven la calidad (15%), e impacto más allá de la academia (15%). Para el primer aspecto, los paneles se dividen en 24 áreas de investigación de acuerdo a la clasificación sueca,²¹ compuestos por expertos (investigadores) en cada una de las áreas. Para los dos aspectos siguientes, los paneles evaluadores se agrupan en cuatro grandes áreas de investigación: Ciencias naturales y ciencias agrícolas, ciencias de la ingeniería, medicina, ciencias sociales, e investigación humanista y artística. Estos paneles están compuestos en 50% por usuarios o beneficiarios de la investigación, incluyendo empresas y otros actores sociales, y 50% de expertos en gestión de la investigación. Las unidades de evaluación representan la suma de la investigación de una institución de educación superior en alguna de las 24 áreas de investigación, y se definen de la misma forma en que estadísticas de educación superior e investigación son reportadas al sistema de estadísticas nacional de Suecia SwePub²².

Para la evaluación, cada unidad debe presentar una pequeña descripción del perfil de investigación, visión, estrategia y gestión (liderazgo, reclutamiento y carrera académica), que formará parte de la información de contexto. También se debe incluir datos cuantitativos como financiamiento externo, y cuantificación del personal de investigación y enseñanza, programas de doctorado, y perfil de publicaciones.

La **calidad de la investigación** (en términos científicos o artísticos) es evaluada en relación a su novedad y originalidad, significancia de la investigación para su área, y rigor académico. Esto se basa en una declaración de investigación de la unidad, una muestra de los resultados de investigación, y análisis de citas, cuando sea posible, y/o revisión por pares externa, según corresponda a la disciplina (como en el caso de las artes y humanidades, donde las citas son poco frecuentes y no representan la calidad de la investigación).

19 El sistema Sueco fue diseñado inspirado por REF, pero se decidió que no sería implementado, principalmente porque las universidades prefirieron mantener su autonomía respecto de las evaluaciones (Sivertsen, 2017)

20 El modelo actual, implementado en el 2009, asigna un porcentaje de financiamiento (20%) para educación doctoral e investigación en base a dos indicadores: publicaciones/citas y financiamiento externo.

21 La clasificación *Standard för svensk indelning av forskningsämnen* (National Standard for Research Subject Classification) puede ser encontrada aquí <https://www.scb.se/contentassets/10054f2ef27c437884e8cdeod38b9cc4/standard-for-svensk-indelning--av-forskningsamnen-2011-uppdaterad-aug-2016.pdf>

22 Repositorio de publicaciones y productos de la investigación realizada en Suecia. <http://swepub.kb.se/>

Los **factores que promueven la calidad** son evaluados a nivel de área de investigación en relación a dos criterios: potencial de renovación y sostenibilidad, es decir, un ambiente de investigación dinámico, con flujos de nuevas personas e ideas, pero a la vez un buen balance y estabilidad que crea las condiciones para investigación de excelencia. Para lograr una buena evaluación en estas dimensiones se debe tener un buen puntaje en cada una de las siguientes sub dimensiones: educación doctoral y de investigadores jóvenes, colaboración y movilidad dentro de la academia (nacional e internacional), colaboración, asociación y movilidad más allá de la academia (nacional e internacional), la integración entre la investigación y la educación, y la equidad de género. Para ello se evalúan métricas cuantitativas y una descripción de las estrategias para desarrollo de la promoción de la calidad.

El **impacto más allá de la academia** es evaluado a nivel de área. Impacto se entiende como “los efectos de la investigación más allá de la academia que en algunos casos y a través del tiempo puede generar influencia concreta en la sociedad a través de la aplicación de los resultados de investigación para lograr efectos sociales, económicos, ambientales o culturales. El impacto más allá de la academia entonces se refiere a la diseminación, refinamiento, comercialización, patentamiento, licenciamiento o cualquier otro uso práctico de los resultados de investigación” (Monaco et al., 2015; pg.36, traducción propia). Se basa en estudios de caso y la descripción de cada unidad de su estrategia de diseminación, colaboración y promoción del uso de investigación más allá de la academia.

Los criterios de evaluación son alcance y significancia. Alcance se refiere a el rango de empresas, autoridades públicas, otras organizaciones y/o individuos que han aprendido o han sido afectados por el impacto de la investigación. No existe en la evaluación diferencias jerárquicas entre local, regional, nacional o internacional; la evaluación debe ser entendida dentro de las especificidades de cada grupo. Significancia se refiere al grado en el cual el impacto de la investigación ha enriquecido, influenciado, valorado o puesto en uso productos, servicios, cambios en política, métodos de trabajo, condiciones médicas descubiertas, problemas y soluciones para empresas, autoridades públicas, otras organizaciones e individuos. Un impacto significativo puede ocurrir tanto en un rango amplio como reducido que afecte sólo a un número de organizaciones, o dentro de un área geográfica limitada. Estos criterios son los mismos utilizados en REF14 (ver sección 2.3.1.).

Los estudios de caso a presentar como evidencia se basan en un formato específico y tienen un límite de extensión de 4 páginas. Contienen la descripción de una investigación claramente delimitada, los resultados de ella, su impacto más allá de la academia y material que permita corroborar ese impacto. El tipo de actividades que pueden incluir distintos aspectos del impacto incluyen: diálogo y diseminación de resultados [ej. Seminarios, conferencias, participación en debates, etc.], colaboración [ej. empresas, consorcios no académicos, financiamiento empresarial o de otros sectores, publicaciones conjuntas, etc.], nuevos productos y procesos [ej. patentes, IP, nuevas políticas, normativas o guías, nuevos servicios, etc.], uso [ej. productos, procesos o servicios usados por organizaciones o empresas, políticas o guías que conlleven cambios en los métodos de empresas u organizaciones, etc.] y efectos positivos documentados [ej. beneficios económicos por comercialización, mejoras en la salud de las personas, beneficios sociales y culturales, mejoras en educación y conocimiento, etc.].

Además, se entrega una breve descripción de la estrategia de impacto de la unidad, que incluye información más específica relacionada con las funciones existentes de apoyo a investigadores en relación a los resultados de diseminación más allá de la academia, puntos de contacto entre investigadores y otros actores no académicos, etc. Esta información apoya el proceso de evaluación, pero el principal componente es el estudio de caso.

La evaluación de cada uno de estos criterios se basa en una escala única de 5 puntos, con 5 como el puntaje más alto. Se deben además adjuntar una declaración explicativa de la evaluación. En este caso, el piloto de evaluación no tuvo consecuencias para las unidades evaluadas, pues su objetivo fue proponer un sistema de distribución de fondos de investigación basado en performance. Es importante notar que la decisión de la autoridad sueca fue no implementar este sistema finalmente, principalmente por un tema de costos, pero además porque las universidades prefieren mantener su autonomía de investigación y organizar las evaluaciones de la investigación ellas mismas (Sivertsen, 2017)²³.

2.3.4. Standard Evaluation Protocol - SEP (Países Bajos)

El Protocolo de Evaluación estándar o SEP (Standard Evaluation System) corresponde al sistema de evaluación de la investigación realizada en Universidades de los Países Bajos y de los institutos de investigación dependientes de NWO (Nederlandse Organisatie voor Wetenschappelijk) la principal agencia de financiamiento de la investigación de los Países Bajos y la Real Academia de las Artes y ciencias de los Países Bajos (KNAW). Es realizado cada 6 años, pero a diferencia de los otros casos de evaluación nacional, no es realizado de una sola vez, sino que se realiza en distintos tiempos para cada grupo de investigación, durante el periodo comprendido entre estos 6 años. Es un sistema de evaluación cuya finalidad no es la asignación de fondos públicos

El principal objetivo de SEP es “revelar y confirmar la calidad y la relevancia de la investigación para la sociedad, y realizar mejoras donde fuese necesario” (Monaco et al., 2015). Esta evaluación se rige bajo los siguientes principios

1. SEP, como un sistema de gestión de calidad sirve para garantizar, revelar y confirmar la calidad y relevancia de la investigación académica. La evaluación incluye los aspectos científicos, sociales y estratégicos de la investigación.
2. Los directorios de las universidades, NWO (Nederlandse Organisatie voor Wetenschappelijk) y la Real Academia toman toda la responsabilidad de la evaluación y de las acciones que emerjan de ella dentro de sus propias instituciones. Son responsables de asegurar que cada unidad sea evaluada una vez cada 6 años. Los directorios deben tomar decisiones procedurales específicas, por ejemplo, respecto del nivel de agregación de las evaluaciones, si son o no necesarias evaluaciones de medio-término, y con respecto a los requerimientos de indicadores de producto (detallados más adelante).

23 La propuesta presentada para un sistema de evaluación incluye además una revisión de los sistemas de asignación de fondos de acuerdo a puntaje, y una estimación de los costos de la implementación de la propuesta (Monaco et al., 2015)

3. La estrategia y objetivos propios de cada una de las unidades son los principios guía cuando se diseña un proceso de evaluación. Esto incluye la composición del comité de evaluación y los términos de referencia, así como el contenido requerido para la auto-evaluación, los que son definidos por cada unidad de evaluación.

Son los directorios de cada institución quienes deciden el calendario de evaluaciones para cada unidad, los términos de referencia para cada evaluación, el nivel de agregación [un grupo de investigación, un instituto, un clúster de investigación o una facultad], y los benchmark adecuados, en consulta con la unidad de investigación. La unidad provee de una auto-evaluación que incluye su estrategia futura, y documentos de apoyo. El comité de evaluación es asignado por el directorio de la organización de educación superior a la que pertenece la unidad y está compuesto por expertos nacionales e internacionales en el área. La unidad debe nominar candidatos para presidente del comité y miembros. Antes de invitar al comité, la composición final del comité evaluador debe ser entregada a la unidad a evaluar, quien debe dar su aprobación. Es importante notar que el comité debe ser capaz de evaluar la importancia social de la investigación, así como conocer el sistema de investigación de los Países Bajos, por lo que no puede incluir expertos no académicos.

Este comité revisa estos documentos, y además realiza una visita a la unidad de evaluación y conduce entrevistas con algunos de sus miembros. Para la evaluación, se consideran tendencias internacionales ciencia y sociedad, así como los objetivos que la unidad se ha planteado. La evaluación se divide en tres dimensiones: calidad de la investigación, relevancia para la sociedad, y viabilidad [el grado en que una unidad está preparada para el futuro].

1. Calidad de la investigación: se refiere a la calidad de la contribución que hace la unidad para el avance del conocimiento científico, incluyendo resultados (publicaciones científicas, instrumentos e infraestructura desarrollada por la unidad, etc.)
2. Relevancia para la sociedad: se refiere a la calidad, escala y relevancia de contribuciones orientadas a grupos específicos, ya sean económicos, sociales o culturales.
3. Viabilidad: se refiere a la estrategia que la unidad busca implementar en los siguientes años, y en qué grado es capaz de lograr sus objetivos, tanto en investigación como en impacto en la sociedad. También considera las competencias de gobernanza y liderazgo de la unidad.

Estas dimensiones se evalúan en cuatro categorías: excelente, muy bueno, bueno e insatisfactorio. También se consideran los programas de doctorado, la integridad de la investigación (aspectos éticos, gestión de datos, y otros elementos de la cultura interna de investigación) y la diversidad de la unidad. Con respecto al criterio de relevancia para la sociedad, algunos indicadores propuestos son:

Tabla 7. Listado de indicadores propuestos de relevancia social

Indicadores del criterio de relevancia para la Sociedad		
Productos Demostrables	Utilización de los productos de la investigación por grupos sociales	Marcas de reconocimiento por grupos sociales
Reportes (por ejemplo para políticas públicas)	Licencias o patentes	Premios públicos
Artículos en revistas para lectores no académicos	Uso de los laboratorios de investigación por grupos sociales	Financiamiento para valorización (innovación, generalmente empresarial)
Instrumentos, sets de datos, software desarrollado para público no académico	Proyectos desarrollados en conjunto con grupos sociales	Número de participaciones financiadas por grupos sociales
Actividades de divulgación, como charlas abiertas, exhibiciones.	Investigación por contrato (Contract research)	Membresía en consejos de instituciones u organizaciones de la sociedad civil

Luego de la revisión de la autoevaluación y documentos asociados, así como la visita a la unidad, el comité prepara un informe de evaluación. Este informe es entregado al directorio de la organización de investigación, quien entrega sus comentarios, luego de lo cual el informe es publicado en el sitio web de la institución. La implementación de las medidas indicadas en el informe de evaluación es responsabilidad de la institución albergante, pero debido a que el informe es de dominio público, funciona en base a *accountability* pública.

3. Conclusiones

Este documento ofrece una breve revisión de los sistemas de evaluación del impacto no académico de las actividades de investigación. Los sistemas de evaluación fueron clasificados en tres tipos: evaluación de proyecto, evaluación de programa o centro y evaluaciones nacionales. Las evaluaciones por proyecto y programa están mayormente orientadas a entender los determinantes y procesos de la generación de impacto en la investigación, poniendo énfasis en la forma en que se crean redes y relaciones virtuosas entre usuarios de la investigación e investigadores. Este tipo de evaluaciones también aporta a que las agencias de apoyo a la investigación científica puedan demostrar el cumplimiento de sus objetivos y permite una mayor sostenibilidad de los fondos públicos disponibles.

Por otra parte, las evaluaciones nacionales, cumplen una función de estandarización y comparación de la investigación a nivel nacional, y de transparencia y rendición de cuentas. En aquellos casos en que las evaluaciones están asociadas a asignación de bloques de financiamiento público, estas pueden llegar a ser muy costosas y crear incentivos perversos y malas prácticas en la presentación de evaluaciones, como ocurre en cierta medida hoy con las publicaciones científicas y/o citas como métricas de evaluación de excelencia. Es importante entonces que las evaluaciones a nivel nacional tomen en consideración las particularidades de cada área de investigación (por ejemplo: ciencias exactas versus artes y humanidades) y de cada institución donde se realiza la investigación (grandes universidades nacionales versus institutos de investigación locales o del estado), entendiendo que sus capacidades y objetivos responden a distintos contextos y funciones.

La evaluación de impacto no académico es cada vez más frecuente y una variedad de sistemas están siendo implementados en diversos países. Es, por tanto, una tendencia que sólo irá en aumento dado que se alinea con una nueva forma de construir políticas públicas que fortalezcan la relación ciencia-sociedad. La experiencia internacional, especialmente en aquellos países donde estos sistemas han sido implementados a mayor escala, como Reino Unido u Holanda, es inmensamente valiosa para entender bajo qué condiciones un sistema equivalente podría ser implementado en Chile. En Latinoamérica, a la fecha, no han sido implementadas evaluaciones de impacto propiamente tal, pero si existe un interés cada vez mayor en considerar otras dimensiones de la actividad de investigación, esto se ha manifestado en documentos como la Política de Ciencia e Innovación 2030 de Colombia (libro verde) que toma el enfoque de innovación transformativa.

Por último, las evaluaciones de impacto, especialmente aquellas relacionadas a programas o proyectos, son particularmente relevantes para políticas CTCI orientadas por misión, donde el objetivo es articular redes que a través del tiempo puedan generar transformaciones sociales a través del conocimiento y la innovación. Sin embargo, las políticas de CTI por misión requieren de sistemas aún más diversos que consideren también la evolución de los ecosistemas de innovación, sus diversos actores y componentes²⁴. Entendiendo que uno de los componentes de los ecosistemas de CTCI es el contexto local (político, económico, ambiental y social) en el que este se desenvuelve, se requieren formas de evaluar los cambios en ese contexto. Por ejemplo, el entender cómo cambios en la política pública inciden en la innovación y viceversa o bien como la innovación afecta y se ve afectada por el mercado. Se requiere entonces sumar, dependiendo del caso y del interés del programa o política CTCI, evaluaciones de mercado, de políticas públicas, ambientales, sociales, etc.

²⁴ Ver este enfoque en el documento: CNID 2018 “Ampliando la comprensión del ecosistema CTCI: desafíos en la forma de monitorear y evaluar: Aplicación al desafío país de sostenibilidad de recursos hídricos”.

4. Referencias

- Adam, P., Ovseiko, P. V., Grant, J., Graham, K. E. A., Boukhris, O. F., Dowd, A.-M., ... Chor-zempa, H. (2018). ISRIA statement: ten-point guidelines for an effective process of research impact assessment. *Health Research Policy and Systems*, 16(8). <https://doi.org/10.1186/s12961-018-0281-5>
- Allen, S., Campbell, P. B., Dierking, L. D., Flagg, B. N., Friedman, A. J., Garibay, C., & Ucko, D. A. (2008). Framework for evaluating impacts of informal science education projects. Presented at the Report from a National Science Foundation Workshop. The National Science Foundation, Division of Research on Learning in Formal and Informal Settings.
- Australian Research Council. (n.d.). Engagement and Impact Assessment Pilot.
- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the Association for Information Science and Technology*, 64(2), 217-233.
- Buxton, M., & Hanney, S. (1996). How Can Payback from Health Services Research Be Assessed? *Journal of Health Services Research & Policy*, 1(1), 35-43. <https://doi.org/10.1177/135581969600100107>
- Callon, M. (1986). The sociology of an actor-network: The case of the electric vehicle. In *Mapping the dynamics of science and technology* (pp. 19-34). Springer.
- CSIRO. (2015). Impact Evaluation Guide. CSIRO.
- Derrick, G. E., & Samuel, G. (2017). The future of societal impact assessment using peer review: pre-evaluation training, consensus building and inter-reviewer reliability. *Palgrave Communications*, 3(17040). <https://doi.org/10.1057/palcomms.2017.40>
- Donovan, C., & Hanney, S. (2011). The 'payback framework' explained. *Research Evaluation*, 20(3), 181-183.
- Georghiou, L., & Clarysse, B. (2006). Introduction and synthesis in Government R&D Funding and Company Behaviour. *Measuring Behavioural additionality*. OECD, 243.
- Harland, K., & O'Connor, H. (2015). Bradening the Scope of Impact. Defining, assessing and measuring impact of major public research programmes with lessons from 6 small advanced economies. Small Advanced Economies Initiative.
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429.
- Hill, S. (2016). Assessing (for) impact: future assessment of the societal impact of research. *Palgrave Communications*, 2(16073). <https://doi.org/10.1057/palcomms.2016.73> | www.palgrave-journals.com/palcomms

-
- Joly, P.-B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., & Matt, M. (2015). ASIRPA: A comprehensive theory-based approach to assessing the societal impacts of a research organization. *Research Evaluation*, 24(4), 440-453.
- Kaldewey, D. (2017). The Grand Challenges Discourse: Transforming Identity Work in Science and Science Policy. *Minerva*, 1-22.
- Latour, B. (2005). *Reassembling the social* (Oxford, Ne). Oxford University Press.
- Lebel, J., & McLean, R. (2018). A better measure of research from the global south.
- McLean, R., & Sen, K. (2018). Making a Difference in the real world. A meta-analysis of research for development (No. IDL-56969). Ottawa, Canada: IDRC. Retrieved from <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56969/IDL-56969.pdf?sequence=5&isAllowed=y>
- Molas-Gallart, J., & Castro-Martínez, E. (2007). Ambiguity and conflict in the development of 'Third Mission' indicators. *Research Evaluation*, 16(4), 321-330.
- Monaco, S., Söderberg, S., Quist, M., Tumpane, J., Eklund, M., Augustsson, A., ... Sandberg, B. (2015). Research Quality Evaluation in Sweden - FOKUS. Report of a government commission regarding a model for resource allocation to universities and university colleges involving peer review of the quality and relevance of research. Stockholm, Sweden: Swedish Research Council.
- NABI. (2018). The current state of Broader Impacts: advancing science and benefiting society. The National Science Foundation. Retrieved from <https://broaderimpacts.net/wp-content/uploads/2018/01/nabi-current-state-of-bi-011118.pdf>
- National Science Foundation (NSF). (2015). Perspectives on broader impacts. NSF, 15-008.
- Ofir, Z., Schwandt, T., Duggan, C., & McLean, R. (2016). RQ+ Research Quality Plus. A Holistic approach to evaluating research. Ottawa, Canada: IDRC. Retrieved from *A Holistic Approach to Evaluating Research - International ...* <https://www.idrc.ca/.../Research-Quality-Plus-A-Holistic-Approach-to-Evaluating-Rese...>
- Roberts, M. R. (2009). Realizing Societal Benefit from Academic Research: Analysis of the National Science Foundation's Broader Impacts Criterion. *Social Epistemology*, 23(3-4), 199-219.
- Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective. *Palgrave Communications*, 3, 17078.
- Smith, K., & Stewart, E. (2017). We Need to Talk about Impact: Why Social Policy Academics need to Engage with the UK's Research Impact Agenda. *Journal of Social Policy*, 45(1), 109-127. <https://doi.org/10.1017/S0047279416000283>

- Spaapen, J. (2018). Third Stream Metrics in PRFS - MLE on Performance-based Research Funding Systems (PRFS) (No. Thematic Report No 4). Brussels: European Commission.
- Spaapen, J., & Van Drooge, L. (2011). Introducing 'productive interactions' in social impact assessment. *Research Evaluation*, 20(3), 211-218.
- Spaapen, J., Van Drooge, L., Propp, T., van der Meulen, B., Shinn, T., Marcovich, A., ... Castro-Martínez, E. (2013). SIAMPI final report. Retrieved from http://www.siampi.eu/Content/SIAMPI_Final%20report.pdf
- Stern, N. (2016). Building on success and learning from experience: An independent review of research excellence framework. London: Department for Business, Energy and Industrial Strategy.
- VSNU, NWO, & KNAW. (2016). Standard Evaluation Protocol 2015-2021. Protocol for Research Assessment in the Netherlands. Retrieved from Standard Evaluation Protocol 2015 - 2021 - Vsnu <https://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/SEP2015-2021.pdf>
- Walker, T., M., M., Kelley, R., L. R., R., Templenton, D., T., D., & Thiele, G., D., B. (2008). Strategic guidance for ex post impact assessment of agricultural research. (No. Report prepared for the Standing Panel on Impact Assessment.). Rome, Italy: GIAR Science Council.
- Whitley, R., & Gläser, J. (2007). *The Changing Governance of the sciences*. Dordrecht: Springer Science + Business Media. Retrieved from <https://doi.org/10.1007/978-1-4020-6746-4>





cnid | Consejo Nacional
de Innovación
para el Desarrollo