

**COOPERACIÓN TÉCNICA
NO ATN/AA-15240-UR
BID-AGESIC**

**Cristina Zurbriggen
Mariana González Lago**

Entregable 4. Definición de indicadores y metodología de evaluación

“La evaluación se ha convertido una herramienta tan frecuente en el sector público que eclipsa la necesidad de resolución de problemas más rápidas, más experimentales, con miras al futuro. Las organizaciones del sector público destinan un 80% de sus energías a la comprensión del pasado y, en el mejor de los casos, la gestión del presente, y tal vez sólo un 20% de sus esfuerzos en explorar sistemáticamente las estrategias futuras para mejorar las políticas y los servicios” (Christian Bason, 2010, p. 18)

Introducción

El objetivo de este producto es aportar al Lab de AGESIC un modelo de evaluación y monitoreo del cambio sustentado en un paradigma epistemológico que permita reflexionar sobre su accionar y resultados del mismo para mejorar el quehacer de la organización en un proceso continuo de aprendizaje y experimentación.

Asimismo, se pretende con este producto brindar al equipo un marco conceptual que ayude a comunicar y encuadrar la naturaleza desafiante que supone intentar promover la innovación social dentro del sector público.

Para llegar a la propuesta de una metodología de evaluación y monitoreo que se adapte a la estrategia y objetivos del Lab, el documento se ha estructurado en dos partes partes. En la primera parte se exponen los retos que existen en el campo de la evaluación para la innovación pública y social dada la preponderancia de los modelos convencionales enraizados en la tradición positivista de pensamiento y toma de decisiones en el sector público. Posteriormente se expone la emergencia de un nuevo paradigma con la introducción de la “evaluación de desarrollo” cuyo objetivo es, entre otros, estimular el aprendizaje para innovar en contextos complejos cuando no se pueden definir a priori los resultados de una intervención. Luego se explicarán en forma más detallada las estrategias principales que existen en el ámbito de la evaluación, para así justificar la que se propone al Lab. Si bien la primera parte conceptual es amplia, consideramos relevante comprender los principios innovadores en evaluación y monitoreo para la innovación pública. En la segunda parte del documento se desarrolla concretamente una propuesta de evaluación y monitoreo orientada al cambio de carácter reflexivo y adaptativo

PRIMER PARTE - FUNDAMENTOS PARA EVALUAR Y MONITOREAR LA INNOVACIÓN PÚBLICA

1.1 Los desafíos de evaluar la innovación social y pública

La evaluación de la innovación social y pública es actualmente un reto importante y de debate debido a la naturaleza emergente del fenómeno cuya complejidad no se ajusta a los modelos convencionales de evaluar, y para los cuales aún no existen métricas definidas (Bund et al., 2015). Si bien la innovación social y pública ha logrado permear algunos discursos políticos, la comprensión sobre la escala y naturaleza del valor social que generan aún es limitada, y la investigación y evaluación en este ámbito no cuenta con un marco teórico ni definición concluyente (Dayson, 2016; Graddy-Reed & Feldman, 2015).

Sin embargo, se puede decir que existe un cierto acuerdo en que el propósito de evaluar estas innovaciones no es únicamente el de mejorar la gestión de una organización, sino además, el de comunicar el valor que éstas generan (Jeppesen, 2016; Dayson, 2016; Bund et al., 2015; Preskill & Beer, 2012; Bason, 2010; Phills et al. 2008). Tal como lo señala Bason, “*La innovación en el sector público tiene mucho más que ver con transformar la democracia que transformar políticas públicas y servicios. Tal vez no podamos tener uno sin lo otro*” (2010: 231).

El desarrollo de mecanismos de evaluación de la innovación en el sector público ha sido tema de interés tanto de la administración pública como la academia. Diversas metodologías y enfoques han sido generados a través del tiempo respondiendo, la mayoría de ellos, a la tradición positivista. En un contexto de constante de búsqueda por la operacionalización del concepto, un primer punto de inflexión hacia una mejor comprensión de lo que la innovación es, lo dieron Milbergs y Vonortas en el año 2004 con la introducción de la “cuarta generación de métricas de innovación”. Con este concepto los autores expresaron la necesidad de pasar de una concepción lineal de la innovación basada en la economía industrial, hacia una concepción no lineal mejor adaptada a la realidad que impone la economía en red basada en el conocimiento. De esta manera, promovieron mediciones para superar el viejo paradigma y para reflejar también los flujos y procesos de la innovación.

Para especialistas en la innovación social como Antadze y Westley (2012) el pensamiento de Milbergs y Vonortas fue una gran contribución en el análisis de las métricas de innovación al reconocerla como un proceso que es multidimensional, incierto e impredecible. Pero a diferencia de las innovaciones técnicas, una de las complejidades inherentes de la innovación social es la dificultad de capturar y calcular su impacto y resultados en términos positivistas de relaciones lineales de causa y efecto (Antadze y Westley, 2012).

No obstante, dada la prevalencia de este paradigma en el sector público tanto en la toma de decisiones como en la evaluación, aún se suele intentar medir el impacto social y público a través de los enfoques generados en este marco epistemológico (modelos de evaluación convencional). Es importante destacar, además, la fuerte influencia que ha tenido en el sector público, principalmente en el ámbito de la cooperación al desarrollo, los enfoques de carácter instrumental como el enfoque de marco lógico, que asume que el cambio puede ser

planificado de antemano, fundamentando en la planificación racional, así como el conocimiento experto predictivo sobre las causas y efectos que el comportamiento humano tiene en las dinámicas sociales (Arkesteijn et al. 2015: 104). Las críticas que suelen recaer sobre el enfoque de marco lógico se centran en que éste no reconoce la complejidad e invisibiliza el desacuerdo y el conflicto. Si bien apunta a que durante la implementación y entre los ciclos del proyecto haya una reflexión crítica sobre las suposiciones que relacionan las intervenciones y los cambios, en general, cree que la incertidumbre puede y debe ser reducida (Arkesteijn et al., 2015).

Bajo la lógica positivista, la innovación pública y social se definen de manera acotada como producto o servicio, en vez de entenderla como un proceso que se desarrolla dentro de sistemas complejos, caracterizados por la incertidumbre y el conflicto que deviene de la diversidad de percepciones y valores de los actores involucrados en una iniciativa, programa u organización. Es por esto que cuando la innovación pública y social se somete a la evaluación convencional, su naturaleza no es comprendida y el dato que se obtiene puede ser forzado, lo que puede llegar a afectar al propio proceso de la innovación en el afán por producir resultados (Morris, 2011).

De esta manera, al reto de evaluar un fenómeno nuevo y complejo como la innovación pública, se suma la dificultad de tener que superar modelos de evaluación de cohorte positivista fuertemente instalados en la administración pública que no son los apropiados para hablar del valor de su naturaleza en tanto capacidad transformadora de sistemas. Es preciso señalar como lo hacen Arkesteijn et al. (2015) que ya en la década de los 90s surgieron una serie de enfoques de evaluación en respuesta al marco lógico que buscaron reconocer tanto la diversidad de valores y percepciones de los actores y cómo estos interactúan con el sistema, así como la construcción colectiva de conocimiento y no restringido a los expertos. Algunos ejemplos de ello son el monitoreo y evaluación participativa, la investigación dialéctica, la evaluación receptiva, historias de aprendizaje, entre otros. Pero estos enfoques, en general, continuaron respondiendo a las lógicas de la evaluación convencional, y ninguno logró promover un cambio de paradigma.

1.2 Hacia un nuevo modelo de evaluación: la Evaluación de Desarrollo (ED)

El punto de inflexión hacia un nuevo modelo se encuentra en el trabajo de Michael Quinn Patton, reconocido experto en evaluación, quien ya en el año 1978 había introducido el concepto de “Evaluación Orientada al Uso -EOU” (*Utilization-Focused Evaluation*)¹ y que a lo largo de los años se fue actualizando con sucesivas ediciones (1986, 1997 y 2008). La EOU se

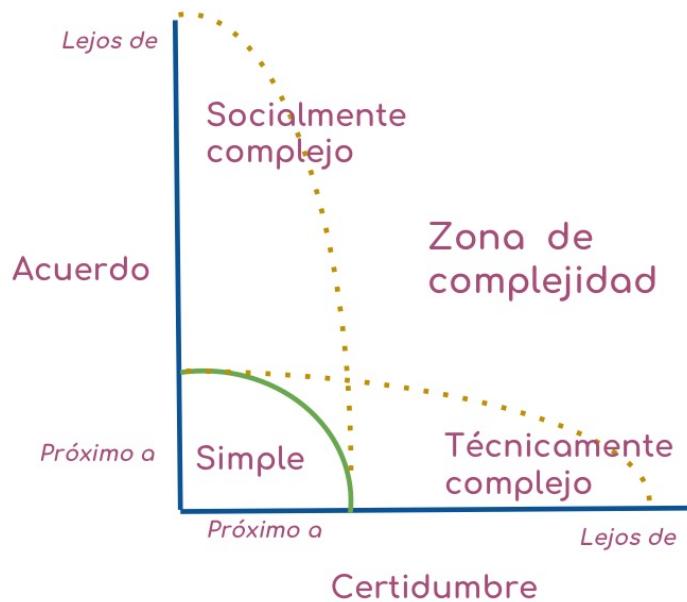
¹ De acuerdo a Patton (2015: 458) se basa en “...la premisa de que las evaluaciones deben juzgarse por su utilidad y uso real; por lo tanto, los evaluadores deben facilitar el proceso de evaluación y diseñar cualquier evaluación con la cuidadosa consideración de que cómo se hacen las cosas, de principio a fin, afectará el uso. El uso implica cómo las personas reales en el mundo real, aplican y experimentan los resultados y el proceso de la evaluación. Por lo tanto, el enfoque de la evaluación orientada al uso está en el <uso previsto por los usuarios previstos>. Dado que ninguna evaluación puede ser libre de valores, la evaluación centrada en el uso responde a la pregunta de qué valores enmarcan la evaluación al trabajar con los usuarios primarios claramente identificados que tienen la responsabilidad de aplicar los hallazgos de la evaluación e implementar recomendaciones”.

considera un marco orientador más que una metodología en sí misma, ya que no “recomienda un contenido, método o teoría particular”; de hecho, puede integrar distintos métodos de evaluación, pero bajo el precepto de la participación activa de quienes harán uso de la evaluación (Ramírez y Brodhead, 2013: 1)².

Pero según Patton (2008), es en el libro publicado en coautoría con Westley y Zimmerman en el 2006: “Getting to maybe: how the world is changed” que reflexiona sobre las implicancias de la teoría de la complejidad para la innovación social, donde se establecen las bases filosóficas de lo que denominará “**evaluación de desarrollo**” (**ED**) y que se constituirá en un nuevo paradigma de evaluación orientado al aprendizaje, la innovación y adaptación en sistemas dinámicos y complejos. En los años siguientes a esa publicación, se establecieron y fueron consolidando los principios que guían esta forma de evaluar (Gamble, 2008; Patton, 2010, 2011; Dozois et al. 2010; Preskill y Beer, 2012) siendo uno de los esenciales, el de aprender y no el de testear para emitir juicios definitivos.

El concepto nace por tanto fuertemente vinculado a la innovación social, y ante el reconocimiento de que los innovadores sociales suelen sentirse desalentados por las formas convencionales de evaluar, que las perciben más como un obstáculo que como un incentivo. De acuerdo a Dozois et al. (2010), la evaluación de desarrollo surge en respuesta a un círculo de frustración producto de las formas de evaluar que producen datos, pero ayudan poco a la comprensión, con evaluadores más preocupados por el rigor metodológico que por brindar recomendaciones útiles, así como por la vorágine de rendir cuentas y respaldar a financiadores más que ayudar a los usuarios de la evaluación.

Figura 1. Definición conceptual de problemas: simples, complicados, complejos



Fuente: Adaptado de Patton, 2011.

² Ver en anexo los 12 pasos resumidos por Ramírez y Brodhead que guían la EOU (pero que no deben ser considerados como parte de un proceso lineal).

De acuerdo a Patton, es necesario incorporar enfoques reflexivos en la evaluación de procesos donde existen diversos actores involucrados y que, como resultado de esas interacciones múltiples, no está claro cómo o si la intervención conducirá a un resultado específico. Concretamente, la define como una evaluación que “informa y apoya el desarrollo innovador y adaptativo en entornos dinámicos complejos. La ED trae a la innovación y la adaptación los procesos de preguntar preguntas evaluativas, aplicando la lógica de evaluación, y recopilando e informando datos evaluativos para apoyar proyecto, programa, producto y / o desarrollo organizacional con retroalimentación oportuna” (Patton, 2011)

Por lo tanto, la evaluación de desarrollo es apropiada cuando es imposible saber de antemano qué intervenciones funcionarán y cuáles serán los efectos de las mismas. Aún si los actores han acordado una solución, y en retrospectiva, entienden cómo se ha trabajado, no implica que la réplica del modelo conduzca a los mismos resultados, ya que una misma solución aplicada por segunda vez, puede fallar (Patton, 2010). Reconociendo la imprevisibilidad inherente de cualquier camino de cambio tomado, se supone que se requiere de una gestión integradora y adaptativa, de sondeo y aprendizaje, y una reflexión recurrente sobre los patrones emergentes (Jones, 2011; Patton, 2011; Snowden y Boone, 2007).

Al presentar la evaluación de desarrollo, Patton afirma lo siguiente: *“No espere tener todas las respuestas de cómo hacer esto al principio ni al final. Pero la evaluación de desarrollo le ayudará a tener claro dónde comenzó, qué se desvió en el camino que tomó y por qué, lo que aprendió en el camino, y donde terminó, al menos por un momento en el tiempo, antes de la próxima ráfaga de viento”* (Patton, en Gamble 2008: 6).

Una de las ideas principales de este enfoque es, por tanto, la de ayudar a guiar la acción colaborativa de iniciativas innovadoras que enfrentan una gran incertidumbre y que se caracterizan por su naturaleza experimental, de co-creación y de aprendizaje social (Arkesteijn et al., 2015).

Otra característica clave es que la unidad de análisis para el cambio y por tanto para la evaluación, ya no es el proyecto o programa (como en los modelos convencionales) sino el sistema. De acuerdo a Patton³, la evaluación nació y creció en el espacio cerrado y contenido de los proyectos, y por ese motivo tiene una fuerte mentalidad basada en el proyecto, pero en los últimos cuarenta años hay un reconocimiento creciente de que los proyectos por si solos no conducen al cambio. El cambio es sustentable cuando es un cambio sistémico de los cuales los proyectos son sólo una parte. Por tanto, si un proyecto quiere obtener determinados resultados tiene que afectar necesariamente al sistema (que es abierto, dinámico y con múltiples factores no controlables por la intervención) en el cual está inserto ya que estos suelen ser disfuncionales e interfieren en la obtención dichos resultados.

Es por ello, que Patton habla de la necesidad de intervenciones y gerenciamiento adaptativo, ya que no se trata de “testear” un modelo sino de generarla constantemente. Las técnicas tradicionales de evaluación no sirven para el cambio sistémico y hacer frente ante lo

³ Traducido y adaptado de Patton:

http://www.betterevaluation.org/resources/guide/Planning_Evaluating_Social_Change

inesperado e impredecible. Estas se basan en modelos probados a través de la evidencia y las buenas prácticas, y que se supone pueden ser escalables. La evaluación de desarrollo no ofrece un modelo prescripto, sino principios de cambio que son adaptables a distintos contextos y que se identifican en el gerenciamiento adaptativo de abajo hacia arriba.

Cuadro I: Paradigmas de evaluación

EVALUACIÓN CONVENCIONAL	EVALUACIONES DE DESARROLLO BASADAS EN LA COMPLEJIDAD ...
Hace juicios definitivos de éxito o fracaso.	Proporciona retroalimentación, genera aprendizajes, apoya la dirección o afirma la dirección de los cambios.
Mide el éxito contra metas predeterminadas.	Genera nuevas medidas y mecanismos de seguimiento a medida que los objetivos emergen y se desarrollan.
El evaluador se posiciona fuera para asegurar independencia y objetividad.	Posiciona la evaluación como una función interna del equipo integrada a la acción y al proceso continuo de interpretación.
La evaluación es diseñada en base a modelos lógicos lineales de causa-efecto.	La evaluación es diseñada para capturar la dinámica del sistema, así como sus interdependencias y emergentes interconexiones.
Tiene como objetivo producir resultados generalizables a través del tiempo y el espacio.	Tiene objetivo producir entendimientos específicos al contexto que informen a la innovación continua.
La rendición de cuentas está enfocada y dirigida a autoridades externas y financieros.	La rendición de cuentas está centrada en el sentido profundo de los innovadores de los valores fundamentales y compromisos.
La rendición de cuentas se hace para controlar y localizar la culpa del fracaso.	Aprender a responder a la falta de control y estar en contacto con lo que se está desarrollando para responder estratégicamente.
El evaluador controla la evaluación y determina el diseño basado en su perspectiva de lo que es importante.	El evaluador colabora en el esfuerzo de cambio para diseñar un proceso que coincida filosófica y organizacionalmente.
La evaluación genera miedo al fracaso.	La evaluación apoya al deseo de aprender.

Fuente: Traducido de Patton, 2006, p. 30

1.3 Sobre las tres estrategias de evaluación

Actualmente en el ámbito de la evaluación se señala la existencia de tres grandes estrategias⁴ cuya selección e implementación depende para algunos del propósito, así como del grado de maduración de un proyecto, programa u organización, y para otros, también del reconocimiento del paradigma epistemológico que sustenta la propia existencia de los mismos. Como se indicó en el producto 3 de esta consultoría, estas estrategias son:

- Evaluación de desarrollo (ej. Evaluación Orientada al Uso, Monitoreo Reflexivo en Acción)
- Evaluación formativa (ej. Evaluación de Implementación, Evaluación de Procesos)
- Evaluación sumativa (ej. Evaluación de Impacto, Evaluación de Resultados, Evaluación de Costo-Efectividad, Evaluación de Costo-Beneficio)

Figura 2. Estrategias de evaluación



Fuente: adaptado de Preskill y Beer, 2012

La **evaluación de desarrollo** como se ha mencionado anteriormente, tiene como propósito motivar la innovación apoyando la exploración y la adaptabilidad en contextos inciertos y dinámicos considerando la complejidad de los sistemas en que tienen lugar las iniciativas. Este tipo de evaluación, por tanto, sirve de apoyo a la innovación social/pública vinculando el pensamiento creativo de la innovación con el crítico de la evaluación (Gamble, 2008).

⁴ Desde la década de los 60s el campo de la evaluación ha evolucionado enormemente, emergiendo diversas técnicas y enfoques como se ha mencionado anteriormente, así como clasificaciones (House, 1978; Stufflebeam & Webster, 1980; Chen, 1996); sin embargo, la categorización más utilizada continúa siendo la de la evaluación formativa y sumativa (Scriven, 1967).

La evaluación de desarrollo (ED) es un enfoque para comprender las actividades de un programa en interacción con que opera en un entorno dinámico y novedoso con interacciones complejas. Se centra en la innovación y el aprendizaje estratégico en lugar de los resultados estándar, y es una forma de pensar sobre los programas en su contexto y los comentarios que producen. Ontológicamente, se basa en la opinión de que "las conclusiones incluyen razonamiento, pensamiento crítico, juicio y argumento, y no se puede reducir a métodos y datos" (Patton, 2011: 15).

De acuerdo a Patton (2011) la evaluación del desarrollo puede ser utilizada con los siguientes propósitos: para el desarrollo continuo de un proyecto, programa, estrategia, política u otra iniciativa adaptándose a las condiciones que imponen los sistemas dinámicos complejos; para adaptar principios generales a un nuevo contexto y desarrollarlos dentro de un nuevo entorno; para desarrollar una respuesta rápida (en tiempo real) frente a cambios importantes repentinos; para desarrollar el rendimiento de una potencial innovación escalable; y para evaluar los principales cambios de sistemas y escalas cruzadas, proporcionar comentarios sobre el despliegue de los cambios en el sistema, evidenciar puntos de inflexión emergentes y/o en cómo se puede tomar una innovación en el esfuerzo de generar impacto. La evaluación de desarrollo trata de persistentemente hacer preguntas y buscar respuestas creíbles a tiempo para ser usadas (Patton, 2011: 288).

Preskill y Beer (2012) señalan que la evaluación de desarrollo tiene cinco características fundamentales que la diferencian de otras formas de evaluación. Estas son, su propósito de desarrollar y adaptar iniciativas innovadoras (y no el de mejorar, emitir juicios o rendir cuentas); la intencionalidad del aprendizaje a lo largo de la evaluación es la de elucidar las innovaciones y los procesos adaptativos, rastrear las trayectorias tomadas en la intervención y ver sus implicancias y resultados para facilitar la toma continua de decisiones basadas en datos tomados en tiempo real; la naturaleza emergente y receptiva del diseño de la evaluación que se convierte en parte de la intervención; el papel y la posición del evaluador como parte del equipo que realiza la intervención, trabajando colaborativamente para ayudar a conceptualizar y diseñar nuevos abordajes; y el énfasis en usar un lente sistémico para recolectar y analizar datos, así como también para generar ideas.

En síntesis, es una evaluación emergente y adaptativa que se realiza en la acción y en tiempo real. A diferencia de los modos convencionales, la ED no evalúa resultados pre-establecidos, sino que evalúa resultados a medida que se van produciendo. No es apropiada para cualquier tipo de intervención, si lo es para aquellas que son emergentes y se insertan en contextos dinámicos y con altos niveles de incertidumbre, y son innovadoras y buscan generar cambios sistémicos.

Pero históricamente la evaluación de proyectos, programas y organizaciones ha estado dominada por dos tipos de evaluación: las llamadas evaluaciones formativas que tienen como propósito mejorar, refinar, y estandarizar la intervención; y las evaluaciones sumativas cuyo propósito es determinar si una intervención ha sido exitosa o ha fracasado en relación a los objetivos propuestos ex ante.

Actualmente la evaluación continúa estando fuertemente asociada a estos modelos convencionales de evaluación cuya clasificación fue introducida por Scriven en 1967. Esta terminología que surgió en el ámbito educativo ante la necesidad de diferenciar los roles de la evaluación entre lo que son procesos (formativa), de resultados (sumativa)⁵ logró calar hondo en las administraciones públicas como esquemas funcionales a la medición de resultados y rendición de cuentas. Es decir, tanto la evaluación orientada a la mejora, como la evaluación orientada a hacer juicios, involucran el uso “instrumental”⁶ de los resultados (Patton, 2001).

Preskill y Beer (2012: 4) señalan que tanto en la evaluación formativa como sumativa subyacen las siguientes suposiciones y principios:

- Testear el modelo con una cadena hipotética clara de causa y efecto.
- Medir el éxito contra objetivos predeterminados.
- El evaluador debe posicionarse como un observador externo, independiente y objetivo.
- Las evaluaciones deben basarse en modelos de predicción lógica.
- Las evaluaciones deben seguir un plan fijo.
- La finalidad de la evaluación es refinar el programa o modelo y luego emitir juicios definitivos de éxito o fracaso.

La **evaluación formativa** (a veces denominada interna) se desarrolla generalmente durante el proceso de implementación de un programa o proyecto con el fin de mejorarlo. La evaluación formativa se aplica en organizaciones o proyectos con mayor grado de consolidación y está centrada en lo que hacen los actores y como ellos se perciben como líderes de cambio en la organización. Trata de obtener información para poder decidir cómo debería continuar desarrollándose el proceso o qué cambios en el diseño deben realizarse para obtener mejores resultados. Es una evaluación, por tanto, que busca comprender y monitorear el trayecto del cambio y entender qué funciona, qué no, y por qué. Su propósito es refinar y mejorar el modelo de la intervención para alcanzar los objetivos planteados.

Dentro de este enfoque de evaluación, Hellström (2013, 2015) identifica dos dimensiones centrales cuando se aplica evaluación formativa a centros de innovación: 1) Capacidad operacional o administrativa que es lo que la organización debe responder todos los días, 2) Capacidad de orientarse al aprendizaje y propia reflexión sobre las condiciones fundamentales de aprendizaje y co-creación transdisciplinaria. Adicionalmente a las dos dimensiones (operaciones y aprendizaje) estos espacios aspiran a tener impacto hacia afuera (redes, cooperación con otros centros) por lo tanto agrega a la dimensión interna, una externa.

⁵ Frente a la crítica recibida a esta categorización como una rígida dicotomización, Scriven responderá más tarde, que ésta corresponde a un tipo de clasificación, pero que no es la única, solo referencia a un rol particular de la evaluación. Como el autor expresa: “*Existen muchos otros sistemas de clasificación para los roles que desempeña la evaluación, y no son necesariamente exclusivos el uno del otro. Existe, por ejemplo, el contraste entre el proceso y la evaluación de resultados y entre el rol de investigación y el rol práctico (...) Otras clasificaciones de evaluación se relacionan con la metodología utilizada más que con el rol de la evaluación o la función prevista, por ejemplo, cualitativa vs. cuantitativa; distanciado vs. interactivo; autoría del evaluador vs. autoría colaborativa; caja negra vs. programa basado en la teoría; basado en objetivos vs. libre de objetivos. Sin embargo, otros se refieren a la calidad, es decir, amateur vs. profesional, o al tipo de evaluación, como en producto vs. proceso, programa vs. política*” (Scriven, 1996: 151).

⁶ En su visión sobre la evaluación Patton (2001) propone el uso “conceptual” de los resultados, y por tanto no se espera ninguna acción o juicio que la suceda, sino que sirva para promover el pensamiento y profundizar la comprensión, mejorando así el conocimiento.

Este tipo de evaluación suele utilizar técnicas de investigación cualitativa, las preguntas que se formulan son más abiertas y conducen a la exploración de procesos, tanto desde el punto de vista del personal del proyecto como de los participantes y otras partes interesadas⁷.

La **evaluación sumativa** se orienta a mejorar una organización, programa o una iniciativa que está relativamente bien desarrollada y el propósito es emitir juicios en el continuo éxito-fracaso en base al cumplimiento de los objetivos del modelo de la intervención. La evaluación sumativa (a veces referida como externa) mide, por tanto, los resultados contra objetivos y marcos predeterminados. Los conceptos claves aquí son ajuste, eficacia y eficiencia en base a preguntas tales como: *¿Cómo podemos dar resultados con más eficacia? ¿Cómo podemos refinar el programa para aumentar los efectos, reducir costos o facilitar la implementación? ¿Qué podemos confirmar sobre nuestra iniciativa que nos ayudará a dar resultados con más consistencia en el futuro?* (Gamble, 2008).

La evaluación de tipo sumativa responde a los formatos tradicionales de evaluación, que como sostiene Gamble (2008), pretende ser “una prueba definitiva de efectividad” y está fuertemente vinculada a conceptos como estandarización, linealidad y cumplimiento, impacto, predictibilidad y control. Esta forma de evaluar no es, según el autor, la más apropiada para abordar la complejidad de la innovación, ya que no se trata de modelos establecidos sino en constante transformación y adaptabilidad.

Las evaluaciones formativas y sumativas son las más extendidas y generalmente su uso es complementario bajo el entendido que de que la primera ayuda a mejorar el desarrollo, a refinar metas y generar estrategias precisas para alcanzar los objetivos, y es una etapa previa a la segunda en la que se determina si los objetivos propuestos se han alcanzado o no y de esta manera testean la teoría subyacente a la intervención. Estas evaluaciones son relevantes e importantes en contextos estables, con variables claramente definidas, y donde el problema es bien comprendido.

1.4. La selección de la estrategia para el LAB-AGESIC: Evaluar para el Cambio

Mulgan (2009) sostiene que los modelos tradicionales de evaluación que responden al paradigma positivista naturalmente conducen hacia una lectura causal, lineal, y de testeo empírico de relaciones hipotéticas. En este sentido señala que la propia definición de la innovación social, que enfatiza más en los procesos complejos que en los resultados, da la respuesta a la evaluación que necesita. Siguiendo el pensamiento de Gunderson y Holling (2002) sobre los ciclos adaptativos aplicados a la innovación social, Westley et al. (2006), sostienen la idea de que estas innovaciones son procesos que no tienen un inicio y un fin determinado, sino que más bien es un ciclo con períodos de continuidad y cambios rápidos caracterizados por la incertidumbre, la no linealidad y auto-organización. Por tanto, sostienen que, en el caso de la innovación social, se debe cambiar el foco de medirla en términos de productos o servicios (como pueden hacerlo los modelos convencionales), a evaluarla como un

⁷ <http://evaluationtoolbox.net.au/>

proceso con impacto, como en el modelo propuesto por Patton. La evaluación de desarrollo ofrece el mejor enfoque para medir la innovación social.

De acuerdo a Patton (2011), la evaluación de desarrollo se diferencia de los modelos tradicionales de evaluación formativa y sumativa, en tanto estos últimos buscan “testear un modelo” y la primera busca adaptarse a realidades dinámicas y emergentes moviéndose adelante y hacia atrás entre el problema y las soluciones. Los elementos que deben estar presentes para aplicar las evaluaciones tradicionales tales como metas previas claramente definidas, un ambiente estable, y soluciones óptimas, no suelen estar presentes en la innovación social en tanto “intervención emergente” que demanda su desarrollo continuo (Patton 2011, p. 41).

La naturaleza de la innovación pública y social se acerca por tanto a la estrategia que propone la evaluación de desarrollo. Concretamente para el caso del Lab, esta puede aportar el marco conceptual para abordar uno de sus principales desafíos como es la evaluación de los procesos de aprendizaje y de co-creación transdisciplinaria⁸ de conocimiento, que son la base de su estrategia. Tal como se estableció en el producto 1 de esta consultoría, su objetivo general el de “promover y facilitar procesos de innovación en el sector público basados en metodologías de co-creación que contribuyan a la resolución de problemáticas complejas en gobierno digital para la generación de valor público y, por tanto, para mejorar la calidad de vida de las personas”.

Figura 5. El desarrollo de la estrategia de acuerdo a Henry Mintzberg



Fuente: Adaptado de Patton, 2012

⁸ **Interdisciplinario** es una forma de investigación que incluye dos o más disciplinas y trabajan para resolver un problema. **Multidisciplinario** es cuando varias disciplinas trabajan en conjunto, pero independientemente una de otra en un tema común. **Transdisciplinario** combina el modo interdisciplinario (integra diferentes disciplinas académicas) y el conocimiento práctico en el proceso de investigación y busca aplicar el conocimiento en un área concreta de intervención. Las disciplinas y el interés social son integrados en nuevo producto de conocimiento (Huutoniemi, 2010). La investigación transdisciplinaria (IT) está orientada a la creación de soluciones eficaces y legítimas a problemas sociales complejos y requiere el intercambio de conocimientos y experiencias entre una diversidad de disciplinas (interdisciplinario) y prácticas de trabajo (Gibbons et al. 1994; Lang et al. 2012; Nowotny et al. 2001; Polk, 2014; Robinson, 2008).

No se puede pensar en el modelo de evaluación del Lab como algo estático, sino que hay que entenderlo como una estrategia viva que debe incorporarse como parte integral de su quehacer en tanto informa y apoya a la innovación. La evaluación de desarrollo es adaptativa porque las estrategias de intervenciones innovadoras que apuntan al cambio sistémico también lo son, en el proceso de desarrollo de una estrategia deliberada, habrá cometidos planificados que no podrán realizarse mientras que otros aparecerán en el camino, dando como resultado la “estrategia realizada” que es lo que efectivamente pudo hacerse en el contexto.

Lo que distingue a la innovación pública y social es que si bien los objetivos de largo plazo se pueden definir claramente, no hay certezas en torno al camino para alcanzarlos y se sabe poco sobre “lo que funcionará, dónde, bajo qué condiciones, cómo y con quién” (Preskill y Beer, 2012: 4). Por tanto, quienes están vinculados a la intervención deben estar atentos y generar información de manera oportuna para guiar el proceso.

En este sentido, y como señala Patton, la ED no ofrece recetas de cómo hacer las cosas, sino un marco para reflexionar y aprender en la intervención, aportando al proceso de innovación preguntas evaluativas, aplicando la lógica de la evaluación y recogiendo y comunicando datos evaluativos para apoyar al desarrollo de la iniciativa. Es un marco riguroso de evaluación siempre que se sea consciente sobre la intencionalidad de las acciones, y se documente constantemente sobre las decisiones que se toman y las que no, el por qué y para qué.

La evaluación del Lab tiene como finalidad contribuir al aprendizaje estratégico, entendiendo como sostienen Preskill y Beer, que existe una relación recíproca entre estrategia y evaluación “que lo que se evalúa debe estar relacionado e informado por las estrategias de la organización, y esa evaluación debería alimentar el desarrollo y el refinamiento de la estrategia. Cuando se concibe e implementa de esta manera, la organización está mejor preparada para aprender, crecer, adaptarse y cambiar continuamente de formas significativas y efectivas” (2012: 4).

Figura 4. Bucle de aprendizaje estratégico



Fuente: Adaptado de Preskill y Beer

Preskill y Beer (2012) afirman que este ciclo de aprendizaje es clave para todo tipo de iniciativa de cambio social, pero es imprescindible para la innovación ya que, si no está presente, tanto los tomadores de decisiones como quienes deben implementar la iniciativa, no disponen de la información necesaria en torno a patrones emergentes y nuevas condiciones que requieren adaptación. Como la innovación suele tener lugar en la incertidumbre, las acciones que se llevan adelante son experimentales, de prueba, y demandan la reflexión constante para saber qué es lo que parece estar sucediendo en el proceso, pero como afirman las autoras “pocos enfoques de evaluación son adecuados para apoyar este tipo de ensayo y error, y de hecho, muchos operan en contra” (2012: 4).

En esta línea, investigadores de la Universidad de Wageningen y la Universidad Libre de Amsterdam (Vrije Universiteit Amsterdam) han desarrollado un modelo de evaluación reflexiva denominado “Monitoreo Reflexivo en Acción (MRA)” (2010) que comulga con muchos de los principios postuladas por Patton. Esta metodología interactiva de evaluación facilita la reflexión en torno a las metas de largo plazo en base a los resultados intermedios, fomentando así el aprendizaje interactivo y colectivo en iniciativas innovadoras, con el fin de crear nuevas reglas, prácticas y relaciones en la red de actores involucrados y contribuir así al cambio sistémico (van Mierlo et al., 2010).

Como establecen van Mierlo et al. (2010a, 2010b), el MRA⁹ se basa en la suposición de que la reflexión colectiva recurrente sobre el sistema actual (barreras y oportunidades) ayuda a estimular el aprendizaje colectivo y a diseñar y adaptar intervenciones sistémicas específicas. En ese proceso, estas iniciativas de innovación generan o cambian las reglas locales, prácticas y relaciones dentro de la red de actores involucrados. Esto tiene lugar en la confusión de las luchas cotidianas de las trayectorias de cambio. Según Regeer et al. (2016) en este enfoque, la reflexividad de un experimento que persigue el cambio sistémico se considera como su propiedad sistémica emergente en términos de su capacidad para interactuar y afectar el contexto institucional en el que opera.

Por otra parte, Arkesteijn et al. (2015) sostienen que el MRA se basa en la premisa que de que si bien la contribución de una única iniciativa de innovación a los procesos de sistemas de innovación largos y caprichosos no puede ser evaluada, es posible caracterizar las acciones de un proyecto y sus resultados en términos de relevancia para la innovación del sistema. La innovación en curso se evalúa con la ayuda de teorías de rango medio sobre los procesos de cambio social, incluyendo comunicación, aprendizaje, construcción de redes y manejo de conflictos, así como también teorías sociológicas e institucionales teorías sobre la innovación de sistemas y prácticas sociales.

Articulando el pensamiento tanto de Patton (2008, 2011), de Bason (2010) como de van Mierlo et al. (2010a, 2010b, 2016) concebimos la evaluación del Lab como una evaluación cuyo propósito no es la rendición de cuentas como en la evaluación tradicional, sino que su propósito es el de generar aprendizaje para la innovación sistémica. No trata, por tanto, de evaluar los resultados del Lab enfocado en conceptos de eficacia y eficiencia. Busca generar

⁹ Ver <https://www.wur.nl/en/show/Reflexive-Monitoring-in-Action.htm>

miradas, ideas, conciencia, apoyar la reflexividad para mejorar la acción mientras se desarrolla, y facilitar el aprendizaje colectivo.

Cuadro 2. Comparación de la evaluación para la rendición de cuentas y evaluación para el aprendizaje para la innovación sistémica.

	Evaluación para la rendición de cuentas	Evaluación para el aprendizaje para la innovación sistémica
Meta de la evaluación	Juzgar los resultados del programa enfocado en la eficacia y eficiencia.	Generar perspectivas para mejorar la acción, facilitar el aprendizaje mutuo entre actores involucrados en el programa y apoyar la reflexividad.
Resultado de la evaluación	Retroalimentación a los financiadores sobre cómo su dinero fue gastado.	Un programa que es capaz de reconocer y responder al cambio, así como dirigir el cambio en su contexto institucional.
Momento de la evaluación	Evaluación ex-post de las metas.	Evaluación ex-durante incluyendo cambios en las metas del programa y medidas.
Metodología de la evaluación	Aplica una metodología establecida.	Implementa un enfoque emergente, específico al contexto, apoyando a los gerentes del proyecto en la toma de decisiones.
Involucramiento de actores relevantes	Poca/No hay participación de actores involucrados	Abordaje multi-actoral basado en la participación de los actores involucrados

Fuente: Traducido de Regeer, de Wildt-Liesveld, van Mierlo, y Bunders, 2016: 11

SEGUNDA PARTE – PROPUESTA DE EVALUACIÓN Y MONITOREO PARA EL LAB-AGESIC

“No vamos a engordar un cerdo pesándolo” (Mulgan, 2009)

2.1. La hoja de ruta para una evaluación para el cambio

En base a la discusión teórica en torno a los paradigmas de evaluación y su pertinencia para la innovación pública y social, propondremos al Lab la implementación una estrategia de evaluación para el cambio, epistemológicamente fundada en el paradigma de las evaluaciones de desarrollo basadas en la complejidad. A esta hoja de ruta que brinda la ED, combinaremos elementos del monitoreo reflexivo en acción (MRA) y el pensamiento sobre la innovación en el sector público de Bason. El principio fundamental de una evaluación de este tipo, es el de generar aprendizajes colectivos que apoyen al equipo en la toma de decisiones y mejorar el desarrollo de su estrategia mientras se va desplegando.

La evaluación de resultados, que es la más usada para proyectos y organizaciones, no garantiza un cambio de comportamiento, sino que proporciona información sobre la gestión de corto y largo plazo y sirve como elemento de rendición de cuentas y transparencia de los fondos públicos, así como de responsabilidad política. En el proceso de innovación de los Labs, lo relevante es aprender a medir los “drivers” del aprendizaje con el fin de mejorar la performance de la organización habilitando un proceso continuo de innovación y experimentación.

La evaluación en un laboratorio de innovación social/pública tiene que apuntar a generar una reflexión constante para su sustentabilidad, así como la de las innovaciones que propone. Por tal motivo, la propuesta de evaluación se validará con el personal del Lab ya que esta tiene que ser incorporada a la organización como parte integral.

Hablar de esta forma de evaluar en el sector público es, sin lugar a dudas, un reto muy importante ya que supone instalar nuevas lógicas y formas de diálogo a las que se está habituado, pero es justamente un Lab de innovación quien debe asumir el desafío de conducir este proceso para legitimar un espacio de gobierno experimental en la órbita pública. El lab no es solamente un espacio “para” la experimentación, sino que debe pensarse a sí mismo como organización experimental. Para apoyar al equipo en esta tarea hemos brindado las bases filosóficas y conceptuales necesarias para facilitar ese diálogo, que a su vez esperamos se irá enriqueciendo con los aprendizajes emergentes de la evaluación.

Esta propuesta es una “hoja de ruta” donde se presentan los principios que según Patton deben guiar la evaluación de desarrollo, pero no hay aquí orden o causalidad. La evaluación para el cambio en tanto parte del quehacer del Lab, tiene que ser desarrollada por el equipo del Lab. Como señalan Schippers et al. (2012) los “equipos reflexivos” se asocian positivamente con la innovación. A partir de un trabajo empírico, los autores han demostrado que cuanto mayor es la capacidad de los equipos de trabajo para reflexionar consciente y colectivamente

sobre sus formas de trabajo y funcionamiento, mayor es la capacidad de predecir resultados, principalmente en innovación.

Como sostiene Gamble, la ED es una investigación rigurosa para el desarrollo. Es ser intencional sobre el uso de datos de una manera significativa para informar la innovación en el proceso, El producto o resultado de un proceso de evaluación de desarrollo exitoso son cambios informados en lo que se está evaluando (2008: 19).

Figura 5. Hoja de ruta para la evaluación para el cambio



Fuente: Elaboración propia en base a Preskill y Beer 2012, van Mierlo et al. 2010, y Bason 2010

En esta hoja de ruta se plantean dos grandes fases que tienen lugar en el ecosistema de la innovación donde se inserta el Lab: la primera relacionada con la evolución de la estrategia, en la cual tras la ejecución de las acciones planificadas, el equipo observa, para dar lugar a una segunda fase donde analiza y reflexiona los procesos que se fueron desarrollando (que partes de lo planificado se pudieron ejecutar, cuales no y que cosas nuevas emergieron en el contexto) para ajustar y adaptar futuras acciones generando aprendizajes que conduzcan al impacto social y el cambio sistémico.

2.2. Evaluar las capacidades de aprendizaje para el cambio sistémico

La evaluación y monitoreo tiene como propósito apoyar un sistema de aprendizaje y cambios institucionales, es decir, reflexionar mientras las acciones se van desplegando y en consecuencia ir generando los ajustes y adaptaciones necesarias para cambiar las reglas del juego. Pensar en torno a las condiciones y capacidades para el aprendizaje debe ser, por tanto, el primer y constante ejercicio del equipo del Lab.

Si bien una iniciativa como el Lab no puede cambiar un sistema, si es posible caracterizar los patrones emergentes de innovación que se traducen en la estrategia de gestión del Lab, las acciones de los proyectos y sus resultados en términos de relevancia para la innovación del sistema.

Algunos disparadores para la reflexión sobre las condiciones y capacidades para el aprendizaje en el Lab son los siguientes:

Monitoreo y Evaluación de Capacidades para el Aprendizaje
¿Se usa regularmente el pensamiento crítico y las prácticas reflexivas al interior del equipo?
¿Cómo se conecta el pensamiento crítico y la reflexión a la toma de decisiones?
¿Qué evidencia indicaría que el proceso planificado está funcionando? ¿O que no funciona?
¿Cuáles son los mecanismos de retroalimentación en tiempo real de la organización para rastrear cambios en el desarrollo de nuestra estrategia planificada?
¿Qué podría salir mal y cómo lo sabríamos?
¿Qué factores imprevistos emergieron y cómo nos adaptamos?
Si las cosas han salido bien ¿Por qué tuvimos éxito? ¿A qué lo atribuimos? ¿Lo sabemos?
¿Cómo podemos aprender de nuestro éxito/fracaso?

Lo central es ir más allá de las teorías de uso frecuente sobre el aprendizaje organizacional, para pasar los límites de las organizaciones y avanzar a un aprendizaje del sistema. Se necesitan analizar si, y en qué medida, el aprendizaje colectivo del Lab hacia la innovación del sistema está teniendo lugar entre el equipo del Lab y sus socios, es decir, la red de actores en términos de nuevos conocimientos, así como nuevas acciones y prácticas. Por lo tanto, el aprendizaje del sistema necesita evaluar si el conjunto actual y relativamente estable de estructuras sociales está siendo desafiada (Van Mierlo, 2010)¹⁰.

2.3. Evaluar el desarrollo del Lab en el marco del ecosistema de innovación

Para evaluar el desarrollo del Lab, necesitamos monitorear:

- a) el contexto de la innovación,
- b) la cultura de la innovación,
- b) cómo se desarrolla la estrategia y qué sucede en el proceso entre lo planificado y lo realizado,
- c) cómo se aprende de los experimentos y qué institucionales sistémicos tienen lugar mientras las nuevas prácticas de innovación se están introduciendo por el Lab.

¹⁰ En la implementación de estas estrategias se han utilizado autoevaluaciones cualitativas y se han introducido en varios países con el fin de evaluar adaptación al cambio climático (GIZ) y otras políticas de innovación, la misma tiene la ventaja de incorporar una dimensión reflexiva de monitoreo. En países como Finlandia se ha establecido esta forma de evaluación en todas las políticas públicas estratégicas elegidas por el gobierno y la sociedad (Torfing et al., 2011).

Estos procesos de monitoreo nos muestran la interacción entre una iniciativa de innovación del LAB y su contexto. Con otras palabras, ser conscientes que iniciar cambios en el entorno institucional implica aprovechar las oportunidades proporcionadas e ir adaptándose a un entorno adverso muchas veces a la innovación. Los Labs necesitan, por tanto, crear/construir una cultura acerca de la innovación, y ayudar al personal público para que sean capaces de reconocer, aceptar y asumir las implicaciones que ello supone. En resumen, los drivers son educación, comunicación y reflexión sobre la propia práctica de innovación.

Las dimensiones relevantes a monitorear y evaluar son: Contexto de la Innovación, Cultura de la Innovación, Estrategia Creativa de la Innovación y forma de trabajo.

a. Contexto de la innovación

Como señala la ED, la unidad de análisis para el cambio y por tanto para la evaluación del Lab, ya no es el proyecto o programa sino el sistema. El cambio es sustentable cuando es un cambio sistémico de los cuales los proyectos son sólo una parte. Por tanto, si un proyecto quiere obtener determinados resultados tiene que afectar necesariamente al sistema (que es abierto, dinámico y con múltiples factores no controlables por la intervención) en el cual está inserto ya que estos suelen ser disfuncionales e interfieren en la obtención dichos resultados.

Por tanto, la evaluación del Lab debe incluir una evaluación del contexto (político, arreglos institucionales, cultura de la innovación) (Bason, 2010). Como sostiene van Mierlo et al. (2010) los factores contextuales no deberían verse simplemente como factores estabilizadores, sino que pueden ser susceptibles de cambio.

Con el fin de implementar este proceso de evaluación se propone realizar auto-evaluaciones cualitativas organizadas por el Lab una vez al año invitando a actores públicos y privados involucrados en la generación de proyectos de innovación. Las dimensiones a evaluar son:

Monitoreo y Evaluación del Contexto Político

¿Cómo juega el contexto político en la consolidación del LAB? (identificar donde están los apoyos y la resistencia al cambio, posibilidades de inversiones en el Lab, etc.)

¿Cómo juega el contexto legal? (donde están las limitantes y donde las oportunidades para el desarrollo y adaptación, posibilidades para desarrollar legislación flexible, etc.)

¿Con qué actores vamos a trabajar? (mapeo de actores internos y externos)

¿Cómo está organizada la innovación para trabajar colaborativamente con otros actores, funcionarios de gobierno, ciudadanos y sector privado?

¿Cuáles son los obstáculos identificados para la innovación pública y social (en la AGESIC y con otros sectores de gobierno)?

¿Cuáles son los incentivos y motivaciones para la innovación pública y social (en la AGESIC y otros sectores del gobierno)?

b. Cultura de la Innovación

Las dimensiones a evaluar son:

Monitoreo y Evaluación de la Cultura de la Innovación

¿Qué espacios se desarrollan para el diálogo y reflexión sobre las prácticas del Lab? Se propicia la discusión/reflexión en torno a lo que significa la innovación para el LAB?

¿Tenemos un lenguaje común de la práctica de la innovación para el Lab? (se ha construido una narrativa/lenguaje común o persisten visiones diferentes, sino qué podemos hacer)

¿Los gerentes y el equipo de Lab tienen la misma visión de la innovación?

¿Se desarrollan estrategias para comunicar o educar en la terminología de la innovación pública (a nivel interno y/o externo)?

¿Cómo comunicamos a nivel del equipo? ¿y con otros actores? (canales, contenidos)

¿Cómo gestionamos las competencias del equipo para trabajar en una organización innovadora, en valores como la diversidad, involucramiento, etc.?

¿Cómo convive el lenguaje de la innovación del Lab con el de la AGESIC? ¿Cuáles son las barreras y desafíos culturales internos?

¿Cómo convive el lenguaje de la innovación del Lab con otros sectores de la administración pública con la que trabaja? ¿Cuáles son las barreras y desafíos culturales externos?

c. Estrategia creativa de innovación

Como ya se ha mencionado, la evaluación de desarrollo es adaptativa porque las estrategias de intervenciones innovadoras que apuntan al cambio sistémico también lo son, en el proceso de desarrollo de una estrategia creativa deliberada, habrá cometidos planificados que no podrán realizarse mientras que otros aparecerán en el camino, dando como resultado la “estrategia realizada” que es lo que efectivamente pudo hacerse en el contexto.

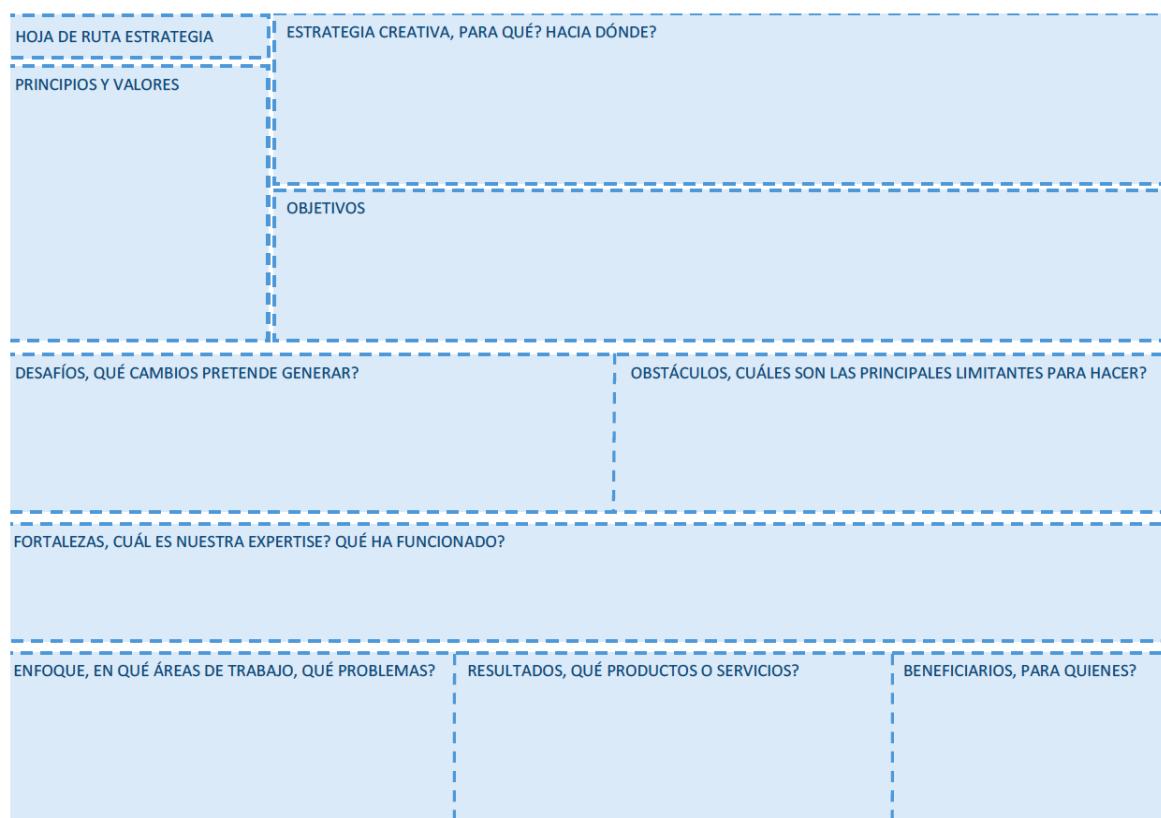
La estrategia creativa refiere al “para qué” de la innovación, es decir, implica, por un lado, identificar, priorizar y volver acción los desafíos concretos que necesitan ser abordados con soluciones innovadoras y canalizar recursos en esa dirección; y por otro, elegir los abordajes conceptuales y metodológicos, así como construir capacidades y habilidades internas para innovar (producto 3).

Monitoreo y Evaluación de la Estrategia Creativa

- ¿Tenemos el qué y el para qué de la innovación?
- ¿Tenemos una estrategia clara para el Lab en términos de objetivos a largo y mediano plazo?
- ¿Cómo estos objetivos y medios están relacionados con una teoría del cambio?
- ¿Qué acciones específicas se piensan desarrollar para alcanzar los objetivos planteados?
- ¿Tenemos una estrategia de cómo vamos a trabajar con el entorno?
- ¿Qué aspectos de lo planificado se pudo ejecutar? ¿Qué hubo que descartar y por qué?
- ¿Qué aspectos no planificados surgieron en el proceso de desarrollo de la acción planificada?
- ¿Cómo se actuó frente a esos emergentes?
- ¿Con que metodología vamos a trabajar? ¿qué adaptaciones o transformaciones se han realizado?

En el siguiente cuadro se presenta una plantilla para sistematizar la información de la reflexión anual en torno a la estrategia creativa, con el objetivo que sirva de base para monitorear y comparar al año siguiente:

Figura 6. Monitoreo anual de la estrategia creativa



Fuente: elaboración propia

d. Evaluación de los experimentos del Lab

Para contribuir y facilitar el proceso de aprendizaje, es útil incorporar una pauta de evaluación sencilla por cada proyecto (o experimento), que basándonos en la experiencia del Mindlab, denominaremos “Protocolos de experimentos”.

Una opción es realizar una evaluación externa con dos cuestionarios obligatorios a las contrapartes de cada experimento (funcionarios de gobiernos) que colaboran con el LAB. El Mindlab propone dos cuestionarios:

Un primer cuestionario se realiza inmediatamente del primer trabajo conjunto con las siguientes preguntas:

Cuestionario 1

- ¿Hemos generado el input correcto (por ej. conocimiento, competencias, habilidades ,etc.)?
- ¿Hemos gestionado el proceso de innovación profesionalmente?
- ¿Fueron los métodos de co-creación y su combinación adecuados?
- ¿Cómo funcionó el involucramiento de los ciudadanos u otros actores en el trabajo?
- ¿La solución que ayudamos a generar es posible de realizarla?

Y un segundo cuestionario 6 meses después:

Cuestionario 2

- ¿Con qué extensión las ideas y los conceptos generados se han podido implementar?
- ¿Cuáles han sido las barreras y los desafíos en su implementación? Y si es así ¿cuáles fueron y cómo fueron abordados o sorteados?
- ¿Se ha generado otro tipo de valor desde su implementación, en términos de productos, servicios, experiencias, democracia, etc.?
- ¿Hay algo que se está haciendo de forma diferente?
- ¿Hay algo más que podemos hacer o ayudar para avanzar hacia delante en el proyecto?

A esto, se le suma de forma complementaria un formulario de auto-evaluación de aprendizaje que hace el equipo del Lab:

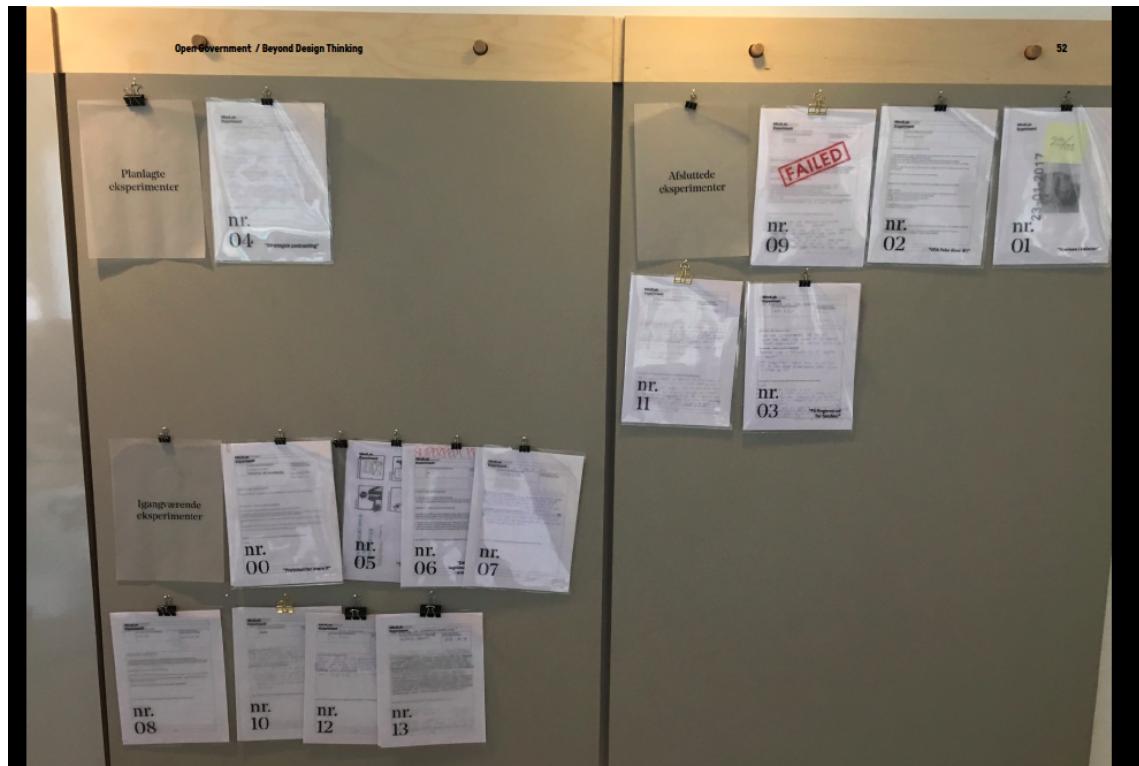
PROTOCOLO PARA LOS EXPERIMENTOS (ex-ante)
Nombre del experimento
Sector/es con los que se desarrolla
Nombre de los participantes
Período de tiempo en el que se desarrolla
¿Qué es lo que estamos testeando y por qué necesitamos testearlo?
Hipótesis: ¿Que esperamos que suceda?
¿Qué queremos aprender del experimento?
Diseño del experimento: ¿Cómo vamos a ejecutar el experimento paso a paso?
Riesgos: ¿Qué puede salir mal?

Fuente: Adaptado de Mindlab, 2017

PROTOCOLO PARA LOS EXPERIMENTOS (ex-post)
Nombre del experimento
Sector/es con los que se desarrolla
Nombre de los participantes
Período de tiempo en el que se desarrolla
Luego del experimento ¿Qué aprendimos?

Fuente: Adaptado de Mindlab, 2017

Estas evaluaciones luego se exponen en el Mindlab:



3. Midiendo el valor público

Parte del ejercicio de evaluar la gestión del Lab debe necesariamente conllevar a una reflexión en torno a su capacidad para **crear de valor público**, entendido como capacidad transformadora y de cambio, ya que éste es su fin último.

De acuerdo a Bason (2010), la innovación social y pública **genera valor público** cuando a partir de la implementación de nuevas ideas a través de un proceso de co-creación se crea valor en al menos una de las siguientes dimensiones: productividad (eficiencia), experiencia de servicios (satisfacción de los usuarios con el servicio), resultados (nueva política o servicio), democracia (participación, transparencia, accountability).

- **Productividad:** expresa la habilidad de alcanzar una relación más favorable entre los inputs y los outputs en la producción de servicios públicos, por ejemplo, cómo una forma más inteligente de trabajo reduce costos, siendo que el resto de los valores se mantiene igual (ej. introduciendo una nueva tecnología o proceso).

Instrumento - Productividad	
Indicador	Fuente/Metodología
Cantidad de talleres realizados por trámite/cliente	Lab/Número total de talleres realizados por cada trámite o para cada cliente
Prototipos diseñados	Lab/Número total de prototipos diseñados
Prototipos en ejecución	Lab/Número total de prototipos en ejecución
Simplificación de procesos	Lab – cliente/% de tiempo reducido con la implementación de la solución
Reducción de costos	Lab – cliente/% de costos reducidos con la implementación de la solución

- **Resultados:** es una expresión del grado en el cual los outputs producto de los esfuerzos del sector público de hecho obtienen resultados deseables. Los resultados son en última instancia, lo que las organizaciones públicas son llamadas a hacer. Por ejemplo, productos tangibles e intangibles, tecnologías, nuevos servicios, cambios estructurales y acciones, planes implementados, decisiones tomadas, nuevos puestos de trabajo, estructura organizativa involucrando a los ciudadanos en la co-producción, etc.

Instrumento - Resultados	
Indicador	Fuente/Metodología
Tramites en línea (numero de trámites disponibles en la web)	Lab - Organismos del Estado/ Número de trámites disponibles en la web
Nuevos servicios	Lab - cliente/Número de nuevos servicios creados
Servicios ajustados/adaptados	Lab - cliente/Número de servicios ajustados o adaptados
Nuevos procedimientos	Lab - cliente/Número de nuevos procedimientos creados
Procedimientos ajustados/adaptados	Lab - cliente/Número de procedimientos ajustados o adaptados

- **Experiencia de servicio:** cómo los ciudadanos y empresarios califican su experiencia subjetiva en el uso de servicios público. Esta una medición que tiene que ver con las subjetividades de los usuarios y el foco es de afuera hacia adentro colocando a los ciudadanos en el centro. No se trata de una medición operacional de la performance sino iluminar sobre calidad subjetiva de la experiencia. Lo importante no es sólo alcanzar el nivel deseado de satisfacción del usuario (encuestas de satisfacción), sino

entender a qué le dan valor los usuarios como clave para identificar el tipo de intervenciones que conducir, finalmente, a mejores resultados.

Instrumento - Experiencia de servicio	
Indicador	Fuente/Metodología
Nivel de conformidad de los participantes con el proceso de co-creación	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)
Nivel de conformidad de los participantes con el tratamiento por parte del equipo del Lab a la pluralidad de ideas	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)
Nivel de conformidad de los participantes con el tratamiento de la información por parte del equipo para encuadrar el problema	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)
Nivel de conformidad de los participantes con las herramientas utilizadas para trabajar en el problema	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)
Nivel de conformidad de los participantes con las decisiones tomadas	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)
Nivel de conformidad de los participantes con la solución acordada	Lab/Escala de Likert (sugerido 1 a 5 donde 1 es muy disconforme y 5 es muy conforme)

Democracia: es una expresión global para los diferentes tipos de valor generados por innovaciones que pueden guiar a incrementos de factores como la participación ciudadana, empoderamiento y transparencia, rendición de cuentas y equidad (Bason, 2010). La innovación en el sector público tiene mucho más que ver con transformar la democracia que transformar políticas públicas y servicios. Tal vez no podamos tener uno sin lo otro (Bason, p. 231).

Instrumento - Democracia	
Indicador	Fuente/Metodología
Participación ciudadana	Lab/Número de ciudadanos que participaron de las actividades
Participación de organizaciones no gubernamentales	Lab/Número de ONGs y asociaciones civiles que participaron de las actividades
Participación de funcionarios	Lab/Número de funcionarios que han participado de las actividades
Participación de otros sectores del Estado	Lab/Número de organizaciones del sector público que han participado
Transparencia y Rendición de cuentas	Lab/ Número de informes disponibles en el sitio web por año
Equidad	Lab/% de población beneficiaria con las soluciones co-creadas en el Lab
Rendición de cuentas	Lab/% cumplimiento de metas anuales
Empoderamiento	

Bibliografía

- Antadze, N., & Westley, F. R. (2012). Impact metrics for social innovation: barriers or bridges to radical change?. *Journal of Social Entrepreneurship*, 3(2), 133-150.
- Argyris, C., & Schon, D. (1978). *Organizational learning: A theory of action approach*. Reading, MA: Addison Wesley.
- Arkesteijn, M., van Mierlo, B., & Leeuwis, C. (2015). The need for reflexive evaluation approaches in development cooperation. *Evaluation*, 21(1), 99-115.
- Bason, C. (2010). *Leading public sector innovation: Co-creating for a better society*. Policy Press.
- Bund, E., Gerhard, U., Hoelscher, M., & Mildenberger, G. (2015). A methodological framework for measuring social innovation. *Historical Social Research/Historische Sozialforschung*, 48-78.
- Chen, H. (1996). A comprehensive typology for program evaluation. *Evaluation Practice*, 17(2), 121.
- Dayson, C. (2016). Evaluating social innovations and their contribution to social value: the benefits of a 'blended value' approach. *Policy & Politics*, 45(3), 395-411.

Dozois, E., Blanchet-Cohen, N., & Langlois, M. (2010). *DE 201: A practitioner's guide to developmental evaluation*. International Institute for Child Rights and Development, University of Victoria.

Gamble, J. A. (2008). *A developmental evaluation primer*. Montreal: JW McConnell Family Foundation.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. Sage.

Graddy-Reed, A., & Feldman, M. P. (2015). Stepping up: an empirical analysis of the role of social innovation in response to an economic recession. *Cambridge Journal of Regions, Economy and Society*, 8(2), 293-312.

Hellström, T. (2015). Formative evaluation at a transdisciplinary research center. Co-producing knowledge for sustainable cities: joining forces for change. Routledge, London, 146-165.

Hellström, T. (2013). Centre of Excellence as a tool for capacity building. Draft Synthesis Report. Programme on Innovation, Higher Education and Research for Development (IHERD), OECD.

House, E. R. (1978). Assumptions underlying evaluation models. *Educational researcher*, 7(3), 4-12.

Huutoniemi, K. (2010). Evaluating interdisciplinary research. *The Oxford handbook of interdisciplinarity*, 309-320.

Jeppesen, L. K. (2017). Evaluating public sector innovation in Denmark. National Centre for Public Sector Innovation, Denmark.

Jones, H. (2011). Taking responsibility for complexity: How implementation can achieve results in the face of complex problems.

Lang, D. J., Wiek, A., Bergmann, M., Stauffacher, M., Martens, P., Moll, P., ... & Thomas, C. J. (2012). Transdisciplinary research in sustainability science: practice, principles, and challenges. *Sustainability science*, 7(1), 25-43.

Mayne, J., & Stern, E. (2013). Impact evaluation of natural resource management research programs: a broader view. Canberra: Australian Centre for International Agricultural Research.

Milbergs, E., & Vonortas, N. (2004). Innovation metrics: measurement to insight. Center for Accelerating Innovation and George Washington University, National Innovation Initiative 21st Century Working Group, 22.

Morris, L. (2011). The innovation master plan: the CEO's guide to innovation. Innovation Academy.

Nowotny, H., Scott, P., Gibbons, M., & Scott, P. B. (2001). Re-thinking science: Knowledge and the public in an age of uncertainty (p. 12). Cambridge: Polity.

Patton, M. Q. (2015). The sociological roots of utilization-focused evaluation. *The American Sociologist*, 46(4), 457-462.

Patton, M. Q. (2011). Essentials of utilization-focused evaluation. Sage Publications.

Patton, M. Q. (2010). Developmental evaluation: Applying complexity concepts to enhance innovation and use. Guilford Press.

Patton, M.Q. (2008). Utilization-focused evaluation, 4th edition. Thousand Oaks, CA: Sage.

Patton, M. Q. (2006). Evaluation for the way we work. *Nonprofit Quarterly*, 13(1), 28-33.

Patton, M. Q. (2001). Evaluation, Knowledge Management, Best Practices, and High Quality Lessons Learned. *American Journal Of Evaluation*, 22(3), 329.

Phills, J. A., Deiglmeier, K., & Miller, D. T. (2008). Rediscovering social innovation. *Stanford Social Innovation Review*, 6(4), 34-43.

Polk, M. (2014). Achieving the promise of transdisciplinarity: a critical exploration of the relationship between transdisciplinary research and societal problem solving. *Sustainability Science*, 9(4), 439-451.

Preskill, H., & Beer, T. (2012). Evaluating social innovation. Center for Evaluation Innovation. Disponible en: https://guelph.ca/wp-content/uploads/Contract_12-191_Description.pdf

Ramírez, R., & Brodhead, D. (2013). Evaluaciones orientadas al uso: guía para evaluadores. Southbound Sdn. Bhd, Malasia. Disponible en:
<https://evaluationinpractice.files.wordpress.com/2013/04/ufeprimerspanish29aug2013.pdf>

Regeer, B. J., de Wildt-Liesveld, R., van Mierlo, B., & Bunders, J. F. (2016). Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. *Evaluation*, 22(1), 6-28.

Robinson, K. (2008). Changing paradigms—how we implement sustainable change in education. RSA/Edge Lecture. London. 16th June.

Schippers, M. C., West, M. A., & Dawson, J. F. (2015). Team reflexivity and innovation: The moderating role of team context. *Journal of Management*, 41(3), 769-788.

Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151.

Scriven, M. (1967). The methodology of evaluation. R.W. Tyler, R M. Gagne, M. Scriven (eds.), *Perspectives of curriculum evaluation*, pp.39-83. Chicago, IL: Rand McNally.

Sinclair, S., & Baglioni, S. (2014). Social Innovation and Social Policy – Promises and Risks. *Social Policy and Society*, 13(3), 469-476. doi:10.1017/S1474746414000086

Stufflebeam, D. L., & Webster, W. J. (1980). An analysis of alternative approaches to evaluation. *Educational Evaluation and Policy Analysis*, 2(3), 5-20.

van Mierlo, B. C., Janssen, A.P., Leenstra, F.R., and Van Weeghel, H.J.E. (2013). Encouraging system learning in two poultry subsectors. *Agricultural Systems*, 115: 29–40

van Mierlo, B. C., Arkesteijn, M., & Leeuwis, C. (2010a). Enhancing the reflexivity of system innovation projects with system analyses. *American Journal of Evaluation*, 31(2), 143-161.

van Mierlo, B. C., Regeer, B., van Amstel, M., Arkesteijn, M. C. M., Beekman, V., Bunders, J. F. G., de Cock Buning, T., Hoes, A.C., & Leeuwis, C. (2010b). *Reflexive monitoring in action. A guide for monitoring system innovation projects*. Communication and Innovation Studies, WUR; Athena Institute, VU.

Voima, P., Heinonen, K., & Strandvik, T. (2010). Exploring customer value formation: a customer dominant logic perspective. Hanken School of Economics, Department of Marketing, Finlandia.

Westley, F., Zimmerman, B., & Patton, M. (2009). Getting to maybe: How the world is changed. Vintage Canada.