

Julián Bertranou

Universidad Nacional de San Martín. Universidad Nacional de Cuyo. Argentina

jbertranou@hotmail.com

EL SEGUIMIENTO Y LA EVALUACIÓN DE POLÍTICAS PÚBLICAS

MODELOS DISPONIBLES PARA PRODUCIR CONOCIMIENTO Y MEJORAR EL DESEMPEÑO DE LAS INTERVENCIONES PÚBLICAS

Resumen: *El presente trabajo tiene como propósito introducir a los agentes públicos, a los docentes y estudiantes y al público interesado en general, en algunas de las especificidades del seguimiento y de la evaluación de las políticas públicas, así como establecer los alcances y limitaciones que tienen las distintas prácticas evaluativas en la producción de conocimiento y en su aplicación en las políticas públicas. El trabajo permite reconocer las fortalezas y potencialidades de las diferentes modalidades de seguimiento y evaluación, con el objeto de facilitar su comprensión e introducción en la trama de procesos de las políticas públicas. El artículo incluye recomendaciones para la institucionalización de procesos de evaluación en las organizaciones públicas.*

Palabras claves: *políticas públicas, evaluación, seguimiento*

Monitoring and evaluation of public policies. Available models to produce knowledge and to improve the performance of public interventions

Abstract: *The present article introduces public agents, teachers, students and interested people, to the characteristics of public policy monitoring and evaluation, and to the scope and limits of evaluation methods for knowledge production and for application in the improvement of public policies. It makes it possible to acknowledge the strengths and weaknesses of different monitoring methods in order to facilitate their comprehension and introduction in the public policy processes. The article includes recommendations for the institutionalization of policy evaluation processes in public organizations.*

Key words: *public policy, evaluation, monitoring*



1. Introducción

Las políticas públicas representan flujos de prácticas dirigidas intencionalmente a la resolución de algún problema, y en las cuales se transforman distintos tipos de recursos para la producción de bienes, servicios o regulaciones necesarios. Las políticas representan el instrumento principal que tienen los gobiernos para resolver las cuestiones que han problematizado, sin que esté garantizado *a priori* un resultado favorable. Los gobiernos enfrentan desafíos cada vez más complejos, principalmente por la creciente multicausalidad de los mismos y por las profundas interdependencias de los factores explicativos.

Los fracasos gubernamentales suelen tener muchas explicaciones. Abundan en estos tiempos las explicaciones asentadas en comportamientos ilegales de los agentes públicos, que encuentran siempre aliados o instigadores no estatales. Asimismo, los malos desempeños pueden deberse a la subordinación de los intereses públicos a los privados, sostenidos por prácticas de captura o cooptación estatal. Más allá de que estas razones existen y se expresan en distintos casos de desgobierno, una parte muy significativa de los malos desempeños gubernamentales se deben a la incertidumbre, a problemas de mala gestión, a negligencias o a la falta de capacidad institucional. En todos estos casos, ocurren importantes rupturas en la producción y el uso de conocimiento en los procesos de toma de decisiones.

En estos casos, es posible observar fallas en la producción de evidencia sobre los problemas públicos y sobre las posibles respuestas de política pública, así como problemas a la hora de incorporar esta evidencia en el núcleo mismo del proceso decisorio. La corriente reciente de las “políticas basadas en evidencia” (Siobhan Campbell y otros, 2007; Pedro Flores Crespo, 2013) pone de relieve el rol fundamental del conocimiento en la mejora del desempeño de las políticas. Unas de las prácticas más importantes tendientes a producir conocimiento son las relacionadas con el seguimiento y la evaluación de las políticas públicas.

El presente trabajo tiene como propósito introducir a los agentes públicos, a los docentes y estudiantes y al público interesado en general, en algunas de las especificidades del seguimiento y de la evaluación de las políticas públicas, así como establecer los alcances y limitaciones que tienen las distintas prácticas evaluativas en la producción de



conocimiento y su aplicación en las políticas públicas.

Los contenidos permiten reconocer las fortalezas y potencialidades de los diferentes *ejercicios evaluativos*, con el objeto de facilitar su comprensión e introducción en la trama de procesos de las políticas públicas. El artículo incluye, primordialmente, una reflexión sobre el uso de los ejercicios evaluativos como fuente de conocimiento para las políticas públicas, en el contexto de otras motivaciones para la toma de decisiones.

El trabajo se organiza de la siguiente manera. En el apartado 2 se analizará el seguimiento y la evaluación en el marco de la planificación de políticas públicas. En el apartado 3 se analizarán los alcances y limitaciones del seguimiento y de la evaluación en términos de lo que significa el conocimiento, la validez de mismo, los usos de la evaluación y su institucionalización. En los apartados 4 y 5 se exponen algunos modelos de seguimiento y de evaluación de políticas públicas respectivamente. Finalmente, en el apartado 6 se presentan algunas reflexiones sobre el fortalecimiento de los procesos de evaluación en el sector público.

2. El seguimiento y la evaluación en el marco de la planificación de políticas públicas: necesidad, factibilidad

2.1 Introducción al porqué de la evaluación

Como una variante de la definición que inicia este trabajo, una política pública es una intención explícita de un gobierno para que cambie un estado de cosas en la realidad (un problema público) que contiene, organiza y coordina acciones de personas, quienes se desempeñan la mayor parte de las veces con el auxilio de procedimientos, tecnologías e infraestructuras.

Lo que frecuentemente desvela a quienes toman decisiones e implementan acciones públicas es saber si lo que hacen a través de las políticas tiene el efecto esperado sobre el problema que motivó su intervención. Y esta inquietud lleva de manera directa a la consideración sobre la *perfectibilidad* de cualquier iniciativa o emprendimiento humano (Julián Bertranou, 2015). ¿Fue precisa la descripción del problema público? Seguramente no de manera completa: la realidad social siempre permite diversas observaciones

e interpretaciones y a su vez, los métodos para conocer la realidad son necesariamente imperfectos. Entonces, podemos preguntarnos si nuestra caracterización del problema no es *perfecta* o inequívoca, con más razón y como fundamento para las evaluaciones, ¿podemos pretender que el diseño de nuestra intervención sea a su vez *perfecto*? Y si claramente no lo es, porque siempre es una respuesta que se aproxima a solucionar un problema, ¿qué medios tenemos para ir *corrigiendo* nuestra intervención a medida en que se va desplegando? Otra cuestión más: la realidad que describimos cuando percibimos un problema puede ir variando, pero no solo por el efecto de nuestra intervención, sino porque otros factores externos pueden estar gravitando. Entonces también por este motivo, ¿no se supone que tenemos que revisar el andar de una política (o programa) para también detectar cambios en el ambiente del problema que le dio origen?

En la base de cualquier ejercicio de seguimiento o evaluación de políticas está el reconocimiento de las limitaciones de cualquier emprendimiento humano y, de entre estos, de las políticas públicas. Revisar lo que hacemos, cómo los hacemos y para qué lo hacemos supone que nuestros diseños y modos de intervención y su implementación son siempre imperfectos y que un antídoto parcial contra la incertidumbre y las prácticas insuficientes es conocer y aprender sobre el alcance y los efectos de nuestra intervención. Hacer seguimiento y evaluar son aspectos específicos de un género de actividades que tienen como propósito conocer para aprender y mejorar nuestras intervenciones de política pública.

2.2 *Qué es el seguimiento y la evaluación*

El seguimiento y la evaluación de intervenciones públicas integran un conjunto de actividades sistemáticas con el propósito de conocer el alcance de estas intervenciones y perfeccionar futuras estrategias. En términos generales, denominaremos *actividades evaluativas*, en el sentido de equipararlas hasta cierto punto entre sí como modalidades a disposición de funcionarios y funcionarias, para indagar sobre los avances que se producen con la implementación de políticas públicas y para *generar un conocimiento pertinente para mejorar las intervenciones*.

La variedad de estas actividades evaluativas es muy



grande, cuestión que analizaremos con más detalle en los apartados 4 y 5. Pero por lo pronto, es importante comenzar a hacer algunas distinciones entre el seguimiento de actividades y la evaluación propiamente tal.

La evaluación ha sido definida como una “actividad programada de reflexión sobre la acción, basada en procedimientos sistemáticos, con la finalidad de emitir juicios valorativos fundamentados y comunicables, y formular recomendaciones” (Olga Nirenberg y otras, 2000). La evaluación implica un proceso que indaga sistemáticamente la intervención que realiza una actividad pública sobre la realidad social, con el objeto de generar un conocimiento que facilite luego la mejora de esta actividad en términos de eficacia, eficiencia y equidad.

Estas definiciones nos permiten identificar algunas características de la evaluación de políticas públicas:

- Es una actividad sistemática, que apela a metodologías diversas.
- Que busca determinar el mérito, la relevancia y el valor de una política o un programa.
- Que este mérito o valor remiten al diseño, la gestión, los productos, los resultados o los impactos y la distribución de beneficios de esta política o programa.
- El propósito final es generar un conocimiento acumulado y producir un aprendizaje organizacional necesario que permita perfeccionar la calidad, la pertinencia, la eficacia y la eficiencia de las políticas y los programas.

La *evaluación de políticas* es una actividad investigativa, con rasgos muy similares a los desarrollados por la investigación académica, pero con la particularidad de que debe generarse un aprendizaje en el nivel de las organizaciones públicas (y no gubernamentales) para poder aplicar los nuevos conocimientos a la mejora de las intervenciones. Representa una actividad investigativa con fines prácticos, y que busca consolidar saberes profesionales en los hacedores de políticas públicas.

La *evaluación* requiere que las intervenciones públicas hayan cumplido un determinado ciclo para ponderar los cambios acaecidos en la realidad social objeto de la intervención. No es necesario que el programa haya concluido, pero sí que haya desplegado de manera completa sus líneas de intervención y haya transcurrido un tiempo prudencial (pertinente a cada tipo de política), para captar los avances,

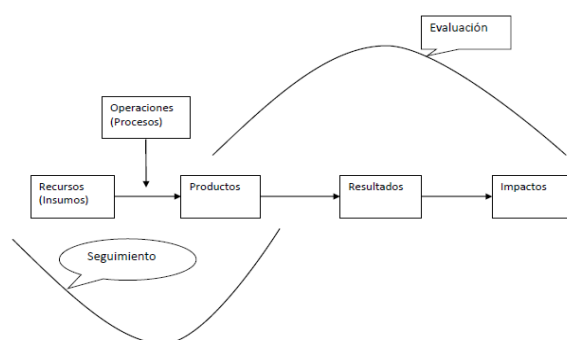
los resultados, y los aprendizajes.

En ese sentido, el *seguimiento de políticas y programas* (también mencionado muchas veces como el monitoreo) es una actividad que se desarrolla desde el comienzo y durante todo el proceso de ejecución de una intervención, porque su principal propósito es suministrar evidencia sobre el proceso de materialización del programa: la adquisición/provisión de los insumos necesarios, la producción de los bienes y servicios intermedios y finales previstos en el diseño del programa, y la calidad y oportunidad de todos los procesos productivos necesarios. En otras palabras, el seguimiento es una actividad concomitante a la ejecución del programa que tiene como propósito verificar el cumplimiento del diseño operativo, generando información utilizable para producir adaptaciones o ajustes inmediatos que permitan asegurar un buen desempeño general.

La actividad de seguimiento puede contemplar también indagaciones sobre la situación del problema que dio origen a la intervención, habilitando determinadas adaptaciones en el suministro de insumos y productos que permitan enfrentar de manera más efectiva las nuevas manifestaciones.

A partir de los diferentes focos de las actividades evaluativas, el seguimiento volcado hacia los insumos, procesos y productos, y la evaluación propiamente tal, volcada también hacia los resultados e impactos, conviene recordar la relación entre todos estos componentes lógicos de una intervención, dentro de lo que comúnmente se denomina, la cadena de valor público (Figura 1).

Figura 1: Cadena de valor público y actividades evaluativas



Fuente: Elaboración propia.



2.3 Argumentos adicionales sobre el para qué y el qué evaluar

Las reflexiones hasta el momento se han orientado a justificar de manera inicial la necesidad de realizar actividades evaluativas y a definir y distinguir estas actividades. En este punto se agregaran otras justificaciones al desarrollo de las actividades evaluativas, y en especial, de las evaluaciones. Es posible organizar y describir estas justificaciones de acuerdo a ciertas dimensiones (Tabla 1).

Tabla 1: Dimensiones en la justificación de la evaluación

Dimensiones	Justificaciones
Política	<ul style="list-style-type: none"> • La evaluación es una herramienta de la legitimación del accionar estatal. • La evaluación es una herramienta de fortalecimiento de la democracia y de control de los ciudadanos. • La evaluación es una oportunidad para estimular la participación ciudadana y el involucramiento no gubernamental.
Político - programática	<ul style="list-style-type: none"> • La evaluación es una oportunidad para conocer y aprender sobre las intervenciones de política pública. • La evaluación es una oportunidad para perfeccionar las intervenciones públicas (modelos, estrategias, integraciones).
Técnica	<ul style="list-style-type: none"> • La evaluación es una herramienta para analizar y generar un conocimiento sobre la pertinencia, calidad y eficacia de ciertos instrumentos de política y ciertas tecnologías. • La evaluación es una oportunidad para consolidar los saberes profesionales de los agentes de la administración pública.

Fuente: Elaboración propia.

La evaluación es una actividad orientada a la generación de conocimiento sobre distintos componentes de una intervención pública, aunque no todos estos componentes son objeto de todas las evaluaciones que se realizan.

A continuación, se detallan estos posibles componentes:

- La pertinencia del diagnóstico del problema.
- La pertinencia y calidad del modelo de intervención.
- La calidad y pertinencia de la teoría de la intervención.
- La calidad, cantidad y oportunidad de los insumos.
- Los procesos de abastecimiento y provisión de estos insumos.
- Los procesos productivos de reconversión de insumos en productos intermedios y finales.
- La calidad, cantidad y pertinencia de los productos

(bienes y servicios).

- La calidad y suficiencia de los efectos/resultados.
- La generación de otros efectos no buscados.
- El logro de impactos o resultados distinguibles de otras variables.

2.4 Restricciones a la práctica de actividades evaluativas

Si bien las actividades evaluativas se justifican por las razones expuestas, su desarrollo no es tan frecuente como el que podría esperarse. Hay diversos factores ligados a las dinámicas políticas y organizacionales que ejercen una influencia tendiente a limitar la realización de evaluaciones o a acotar su relevancia y aplicabilidad.

Es importante conocer estos factores para abordarlos debidamente y, así, limitar su efecto obstaculizador. A continuación, se exponen algunos de los factores más significativos:

a) Opacidad deliberada como estrategia para reducir riesgos políticos: la actividad evaluativa ofrece generar conocimiento sobre las políticas públicas y sus efectos en la sociedad. Es frecuente que las autoridades políticas no promuevan las evaluaciones para evitar confrontarse con los eventuales fracasos de los programas, y que la publicidad de estos últimos afecte negativamente su continuidad en los cargos.

b) Debilidad de la cultura evaluativa y ausencia de tecnologías adecuadas: si la evaluación no es considerada una actividad necesaria y productiva para la organización, es muy probable que se desestime la implementación de actividades evaluativas. A esta debilidad se suma la falta de tecnologías o metodologías adecuadas para la realización de las evaluaciones. Si bien no debe sobreestimarse esta necesidad de metodologías, que pueden hacer de la evaluación una actividad excesivamente sofisticada para la media de las organizaciones públicas, es cierto que es preciso contar con un núcleo de profesionales que dominen dimensiones centrales en la implementación de actividades evaluativas.

c) Resistencia al cambio en general: las organizaciones son proclives a resistirse a los cambios que se introduzcan en sus rutinas y en sus propósitos, en la medida en que afecten espacios de poder o de prestigio para algunos miembros. La evaluación es una fuente importante de eventuales cambios



en la organización, por lo que puede ser resistida en su implementación. Por otro lado, las organizaciones expresan de distintas maneras un temor a eventuales sanciones a partir del juicio que se haga de la evolución de los programas bajo evaluación. Es imperioso disipar el miedo y asociar la evaluación con procesos de aprendizaje organizacional y mejora de la ejecución de programas.

d) Desconocimiento general sobre el accionar de las acciones públicas: la sociedad, los ciudadanos y, entre ellos, los comunicadores y periodistas tienden a tener una visión simplista del accionar de las organizaciones públicas y de la complejidad de la acción de gobierno. La ausencia de logros por parte de un gobierno, además del carácter controversial de los mismos, es atribuida principalmente a la incapacidad de las organizaciones o a la corrupción. Sin descartar estas causas, las principales fuentes de fracaso están en la incertidumbre, la complejidad de la situación social, la incompletitud de los modelos de intervención y la posibilidad de error en cualquier emprendimiento humano. La evaluación, nuevamente, es una herramienta para el aprendizaje y el perfeccionamiento, y esto debe comunicarse a todos los ciudadanos y comunicadores.

2.5 Variedad de enfoques, metodologías y técnicas en las actividades evaluativas

La variedad de enfoques, metodologías y técnicas de evaluación serán el objeto principal en los apartados 4 y 5. En ese punto, basta con señalar esta diversidad que se manifiesta en distintos planos.

Guerrero Amparán (1995) distingue cinco tipos de evaluaciones de acuerdo a cómo se fija el criterio de evaluación:

- *Descriptivas*: no asumen ningún criterio *a priori*, limitándose a registrar dos fotos (comienzo y final de una intervención) e identificar las diferencias.
- *Clínicas*: se realizan de acuerdo a un criterio definido endógenamente. Se trata de la mayoría de las evaluaciones, donde los términos de la evaluación se fijan desde el propio programa.
- *Normativas*: se realizan de acuerdo a un criterio definido exógenamente. La medida de la evaluación es establecida por un actor externo. Las actividades relacionadas con la normalización y la certificación pertenecen a este rubro.

- *Pluralistas*: se realizan a partir de la conjunción de distintos criterios evaluativos, aportados por distintos actores (beneficiarios, técnicos, otros), tanto internos como externos. Se trata de evaluaciones participativas que involucran en su diseño e implementación a una gran variedad de actores.

- *Experimentales*: requieren de la identificación/construcción de un grupo control para contrastarlo con un grupo de destinatarios a partir de criterios definidos por el propio programa.

Las evaluaciones se diferencian también a partir del objetivo de su indagación: descriptiva (relevando procesos y cambios), normativa (contrastando con ideales o estándares) o de profundización de la relación causa-efecto (comprobando la teoría del programa y sus mecanismos). Las evaluaciones pueden hacer un uso diferenciado de abordajes cuantitativos y cualitativos, así como requerir diseños de carácter semiexperimental.

En suma, el campo de las actividades evaluativas es muy diverso, enriqueciendo los instrumentos para conocer la realidad y actuar sobre la misma (Banco Mundial, 2004).

3. Alcances y limitaciones del seguimiento y de la evaluación: conocimiento, validez, usos e institucionalización

3.1 Introducción

La evaluación es falible

La evaluación es solo una de las fuentes de evidencia

La evidencia es solo uno de los insumos de las políticas

Las políticas son solo una de las influencias sobre las prácticas

Las prácticas son solo una de las influencias sobre los resultados

Carol Weiss et al. (2005)

La cita anterior de Carol Weiss invita a reflexionar sobre una serie de cuestiones fundamentales en la evaluación de políticas, y que frecuentemente son desdeñadas:

- ¿Cómo se genera conocimiento en el marco de la evaluación de políticas?
- ¿Cómo se asegura validez interna y externa de los hallazgos?
- ¿Cómo es la relación entre el conocimiento generado en la evaluación (o en cualquier otro proceso) y su aplicación



al campo de las políticas?

- ¿Cuál es el objetivo de este conocimiento: más inmediato y dirigido a las políticas, o más general y de contexto?
- ¿Cuál es la relación entre conocimiento y política?
- ¿Cómo se comunican los hallazgos de una evaluación?
- ¿Cómo se institucionaliza la evaluación y cómo se establecen relaciones sistemáticas entre el ámbito de la evaluación y el ámbito de la toma de decisiones?

Estas preguntas permiten estimar de manera más acertada los alcances y las limitaciones de la evaluación, especialmente en todo lo relacionado con la complejidad conceptual y valorativa de la generación de conocimiento, así también con la factibilidad y pertinencia de la aplicación del conocimiento a la toma de decisiones. En los próximos puntos se expondrán elementos de juicio para una reflexión en este sentido.

3.2 El contexto evaluativo

El contexto evaluativo comprende el marco más amplio de procedimientos, instituciones y expectativas que rodean la realización de actividades evaluativas.

Estas actividades están inmersas en una matriz de relaciones sociales de distintas dimensiones:

- El marco de procedimientos y reglas de juego de la producción de políticas públicas, dentro del cual existe (o no) un espacio para la realización de las actividades evaluativas, porque representan un requisito (o no) en todo el proceso de una política.
- La cultura evaluativa presente en las organizaciones políticas y públicas.
- Las expectativas que existen de manera más general o de manera más específica y coyuntural alrededor de la realización de una evaluación (momento político, relevancia de un programa, entre otras cuestiones).

El que una evaluación se realice en el marco de procedimientos institucionalizados, donde la evaluación es ponderada favorablemente y goza de legitimidad, y las expectativas se centran en un proceso de reconocimiento y aprendizaje, entonces esta evaluación verá facilitada su implementación, la comunicación de sus hallazgos y la traducción de sus recomendaciones.

La generación de conocimiento sobre el desempeño de una política o de un programa no es una cuestión exenta de discrepancias y disputas, principalmente por el proceso a partir del cual se construye un juicio sobre la realidad empírica. Al igual que en cualquier investigación social (y no olvidemos que la evaluación de políticas públicas tiene por objeto las intenciones y acciones de individuos y organizaciones sobre relaciones sociales existentes que producen situaciones problemáticas), hay enfoques (epistemológicos) que tienen perspectivas distintas acerca de qué conocimiento y cómo se construye.

Para el enfoque *positivista*, la realidad social es fácilmente apprehendida a través de instrumentos principalmente cuantitativos, que aprovechan el hecho de que hay un mayor consenso acerca de la rigurosidad de cómo ha sido definido el problema y cómo ha sido formulado el programa o el proyecto. En la medida en que haya este consenso sobre el mundo empírico, los indicadores seleccionados son considerados inequívocos para juzgar los avances del programa y las posibilidades de comparación no presentan limitaciones. Para este enfoque, los datos que se obtienen no ameritan juicios diferentes por parte de los evaluadores.

Para el enfoque *interpretativo*, la realidad se nos presenta como mucho más compleja e incierta, y las discrepancias sobre lo que observamos provienen de las diversas posiciones valorativas de los observadores, de los criterios a partir de los cuales es posible realizar un juicio sobre un programa y de la imprecisión de los métodos de observación en relación con la complejidad de la realidad social. Este enfoque admite que hay distintas *miradas* y distintas evaluaciones posibles, así como confía en métodos cualitativos para la generación de conocimientos. La experiencia es también una importante fuente de aprendizaje y de conocimiento (Feinstein, 2007).

Más allá de que para el desarrollo de evaluaciones son imprescindibles tanto métodos y técnicas cuantitativas como cualitativas, la investigación asociada a las políticas públicas requiere de un abordaje más propiamente interpretativo, en la medida en que el objeto de la indagación son relaciones sociales y sus efectos, los cuales pueden ser valorados desde distintas dimensiones y desde distintas perspectivas, todas ellas admisibles en cuanto a la generación de un conocimiento sobre las políticas.

Algunos ejes de análisis son importantes en este punto:



la identificación de los criterios de evaluación, la cuestión de la validez y la cuestión de la atribución.

La evaluación que se hace de un programa puede ser realizada desde distintos criterios. Estos criterios son complementarios, aunque pueden ser contradictorios entre sí. A continuación, se describen algunos de los más comunes:

- **Calidad:** este atributo alude a distintas facetas pero la principal es la satisfacción de los sujetos destinatarios de política en cuanto a la naturaleza del servicio o del bien ofrecido, así como a las condiciones ambientales y relacionales de su provisión.

- **Pertinencia:** este criterio evalúa la adecuación de la respuesta pública a las características del problema que le da origen.

- **Eficacia:** este criterio es quizás uno de los preferidos; alude al cumplimiento de las metas (cantidad y tiempo) convenidas.

- **Eficiencia:** este criterio es también uno de gran demanda y refiere al costo y agilidad en la provisión de los bienes y servicios ofrecidos por el programa.

- **Legitimidad:** este criterio, más abstracto que los demás, alude al apoyo y la confianza que la política o el programa recibe por parte de distintos públicos, entre los cuales se encuentran los destinatarios, los ejecutores y las organizaciones en general.

- **Relevancia:** por relevancia se entiende la significación y la incidencia que el programa tiene dentro del conjunto de las intervenciones en un sector determinado, o dentro de las propias condiciones de vida de la población destinataria (Feinstein, 2007).

- **Transparencia:** alude a los procedimientos de ejecución y su grado de adecuación a los estándares de publicidad y rendición de cuentas del sistema administrativo.

- **Innovación y aprendizaje:** se refiere a la capacidad que tiene un programa para generar su propio conocimiento y aplicarlo a la mejora de su ejecución.

Frecuentemente es posible encontrar políticas o programas eficaces pero poco eficientes, o muy pertinentes y relevantes, pero poco transparentes en su ejecución.

La cuestión de la *validez* es crucial en la generación de un conocimiento. La pregunta puede formularse de la siguiente manera: ¿cómo se entiende el concepto de la validez y cómo se aborda el mismo? El conocimiento que puede considerarse

válido, al menos en un tiempo determinado, es aquel que ofrece una interpretación lógica y comprobada de las relaciones sociales y sus efectos, al menos, desde los criterios y valoraciones desde los cuales partió el o la observadora. Esta interpretación es sujeta de un análisis de validez interno y externo. El interno alude a la consistencia lógica de la secuencia argumental, en especial, a las relaciones entre causa y efecto y sus interdependencias. El externo remite a la adecuación de la argumentación respecto del mundo empírico, a partir de los métodos y fuentes de observación utilizados.

Una aplicación diferente de la cuestión de la validez a las evaluaciones de impacto es la que ofrece Xavier Ballart (1992), cuando dice que las dos cuestiones esenciales en la evaluación del impacto de un programa consisten en determinar:

1. Hasta qué punto los efectos del programa son realmente debidos al programa y no a otros factores.

2. Hasta qué punto los resultados de un programa concreto pueden ser generalizados a otras situaciones.

La primera es una cuestión de validez interna, mientras que la segunda es una cuestión de validez externa. En general se afirma que las cuestiones de validez interna son más importantes que las de validez externa, dado que hasta que no se puede afirmar que los efectos observados se pueden atribuir al programa, no tiene sentido generalizarlo a otras situaciones (1992: 113).

El primer punto planteado por Xavier Ballart nos remite al problema de la *atribución*. Esta cuestión es una de las más significativas en el campo de la evaluación y ha dado pie a distintos modelos, algunos de los cuales (sobre todos los experimentales y semiexperimentales) se esgrimen cómo los únicos que en efecto puede dar cuenta de este problema (se analizarán sintéticamente estos modelos en las unidades 4 y 5). La cuestión de la atribución puede resumirse de la siguiente: *¿hasta qué punto los cambios producidos en una situación problema determinada que fue el disparador de una intervención pública se han debido a esta intervención y no a otro factor?* Dicho de otra manera, *¿es posible atribuir los cambios observados en la realidad a la intervención de este programa?*

Para Feinstein (2007: 6), el problema de la atribución se formula a partir de una serie de preguntas:

a) ¿cómo identificar el efecto de la intervención X en las variables clave?



b) ¿cómo aislar o separar dicho efecto del correspondiente a otros factores?

c) ¿qué hubiera sucedido en ausencia de la intervención X?

d) ¿qué se suponía que hubiera sucedido si no se realizaba la intervención X?

Las actividades evaluativas no siempre producen un conocimiento que permite despejar las dudas sobre la atribución de un cambio determinado. En la mayoría de los casos, los cambios se producen tanto por la intervención de un programa como por otros factores. Algunas evaluaciones generan un conocimiento que permite vincular cambios en la realidad a un programa, concluyendo sobre su necesidad aunque no sobre su suficiencia. Este conocimiento es de todas maneras muy valioso, ya que permite valorar la calidad de las prácticas y sus efectos concretos. Otro tipo de evaluación como las de impacto con diseños muestrales semiexperimentales, ofrece evidencia más concluyente que permite discernir los efectos de un programa en relación con otros factores.

3.4 Conocimiento y toma de decisiones

La relación entre la evaluación y el conocimiento que eventualmente genera, y la práctica de las políticas públicas no es directa. La anterior frase de Carol Weiss ilustra cierto escepticismo acerca del uso de las evaluaciones por parte de hacedores y políticos en las políticas públicas. Por lo pronto, ¿cuál es el uso que puede hacerse de las evaluaciones? ¿para qué se hacen? Carol Weiss (1999) tiene algunas orientaciones. Para la notable experta, la evaluación, como investigación, puede estar destinada a proveer:

- Datos: evidencia empírica.
- Ideas: nuevas aproximaciones posibles a un problema.
- Argumentos: especificaciones de causa-efecto o de relaciones sociales alrededor de un problema y su resolución.

Sobre los usos en las políticas públicas, Weiss entiende que las evaluaciones pueden aportar a distintas dimensiones o momentos:

- Expansión del conocimiento y puesta a disposición para su aplicación (oferta).
- Resolución de problemas (respuesta a demandas).
- Aporte a la comprensión general (*enlightment*).
- Uso en la lógica política (racionalización de decisiones).

- Uso táctico (situacional con respecto a la acción).
- Interactivo (contrastación de distintas fuentes).
- Ejercicio intelectual (investigación de políticas).

La autora estadounidense es escéptica respecto del uso inmediato de las conclusiones y recomendaciones de las evaluaciones en la reformulación de políticas, inclinándose más por el rol de *ilustración* (*enlightment*) de las mismas: generando evidencias y conocimientos sobre un campo más general de relaciones sociales alrededor de problemas y modos de resolución, cuyo uso es más mediato y asociado a otro de tipo de fuentes para la toma de decisiones.

En la discusión sobre el uso de las evaluaciones en las políticas públicas se entretiene el necesario debate sobre la relación entre conocimiento y política. ¿Cuáles son las fuentes a partir de las políticas, políticas o hacedores de política pública toman decisiones? La pregunta lleva en primer término a la diferencia que existe entre los hechos y los valores, como fuentes primordiales de decisión en la conducta humana.

Los hechos implican enunciados sobre el mundo empírico, sobre los cuales es posible emitir un juicio sobre su veracidad. Los valores en cambio, no habilitan una argumentación técnica. Sobre los valores (como el del sentido de la justicia o la responsabilidad), se puede estar de acuerdo o en desacuerdo. Los primeros representan una fuente de argumentación por parte de los profesionales y técnicos de las políticas públicas (aunque no exclusivamente). Los segundos, son una fuente a la que acuden frecuentemente los políticos (aunque no exclusivamente).

Hechos y valores se encuentran entrelazados como fuentes de decisión en las políticas públicas. Existen elementos de ambos en la fundamentación de cualquier programa, y es importante distinguirlos, aunque no deben ser erradicados como fuente de decisión. De lo contrario, pueden construirse espacios gobernados o por tecnócratas sin valoraciones explícitas, o gobernados políticos despreciativos de las evidencias.

Vale en este punto un ejemplo para ilustrar este punto. En los programas de transferencias condicionadas de ingreso, la condicionalidad (el hecho de que los destinatarios deban cumplir ciertas condiciones como la de acreditar la escolaridad de los hijos o su atención de salud) es un



componente necesario para la obtención del beneficio. La condicionalidad está asociada a una concepción de la responsabilidad de los destinatarios: no debe recibir un subsidio público quien no acredite *responsabilidad* en cuanto a su uso. Esta concepción no es aplicable a todos los subsidios públicos, sino a aquellos donde los destinatarios son familias vulnerables. La condicionalidad es producto de una *valoración*, no de un enunciado empírico. Si bien es posible generar un conocimiento empírico acerca del efecto de la *condicionalidad* sobre otras dimensiones (escolaridad, salud), su adopción o no puede estar asociada a la valoración que se ha descripto.

Como ha sido mencionado en la Introducción, en los últimos años y gracias al avance de los sistemas de información y del conocimiento empírico acumulado, ha ganado fuerza la reivindicación de unas decisiones de política pública *basadas en evidencia* (*evidence-based policy analysis*). Este avance ha sido muy importante para el desarrollo de la disciplina y para el despliegue de las actividades evaluativas. No debe sin embargo, sobreestimarse el uso de la *evidencia* en la toma de decisiones. En primer lugar, no todo lo que es presentado como evidencia (científica) debe ser considerado como tal. En la producción de evidencia puede haberse desarrollado una serie de sesgos de distinto tipo: en la selección de los datos, por el objeto de las investigaciones o por la finalidad de las mismas (Justin Parkhurst, 2015).

En segundo lugar, la *evidencia* obtenida a través de investigaciones o evaluaciones no es ni puede ser considerada la única fuente explicativa de las decisiones públicas. En la toma de decisiones operan también en mayor grado criterios vinculados a la conveniencia y la oportunidad política, a la negociación o a la misma intuición. No solo la *evidencia* no resuelve toda la incertidumbre circundante, sino que las decisiones tienen que necesariamente incorporar criterios de factibilidad y de oportunidad, que la mayor parte de las veces depende del análisis político de los actores participantes.

Existe entonces un sano equilibrio entre el conocimiento desarrollado de acuerdo a los cánones de la investigación social (interpretativa, en este caso de las evaluaciones), y el conocimiento o la motivación que abreva en el análisis de la oportunidad y de la factibilidad política. Las políticas públicas deben desplegarse en un contexto de gobernabilidad necesario, cuyas fuentes se nutren de evidencia y negociaciones.

El mundo del conocimiento académico investigativo difiere bastante del mundo de la política. No existe necesariamente armonía, sincronía, o simplemente puntos de contacto. La distancia cultural es significativa: varían muchos los ámbitos institucionales de referencia de los dos sectores, las orientaciones temporales con sus plazos de realización, los sistemas de incentivos, y referido a la evaluación de políticas, la apreciación que cada uno de los sectores haga de los resultados.

Con estos elementos de juicio es muy probable que las evaluaciones no surtan el efecto inmediato buscado por los evaluadores. Si bien esta constatación podría llevar a cierto escepticismo respecto del rol de las evaluaciones, lo importante, por el contrario, es reconocer los alcances y limitaciones de las actividades evaluativas y reconocer la necesidad de integrar las lógicas tanto académicas como políticas en un proceso de mejora gradual de las intervenciones.

4. Modelos de seguimiento de políticas y programas

En la apartado 2 se distinguieron de manera general, las actividades de seguimiento de las de evaluación de políticas y programas. En algunas oportunidades, la diferencia en las actividades no es tan nítida, y los nombres de las mismas pueden generar alguna confusión.

Las actividades evaluativas asociadas al seguimiento tienen que ver con lo que en la Introducción se definió como “evaluaciones *durante* la ejecución” de las políticas y programas, y que en términos de las distinciones realizadas por Nirenberg et al. (2003), están orientadas principalmente a los *procesos*.

Aunque no todas las actividades de seguimiento se corresponden con esta característica, uno de los rasgos más importantes es que el seguimiento tiene un fuerte foco en los *procesos* productivos de la organización y en los *productos* de las políticas. Las actividades de seguimiento (o monitoreo) de políticas y programas ofrecen información durante la ejecución de las mismas y con el objeto de generar acciones adaptativas y correctivas.

A continuación, se ofrecen variedades de actividades evaluativas de seguimiento o de evaluación *durante* la ejecución de políticas y programas, identificando en cada



caso su foco principal y algunas aplicaciones.

4.1 Seguimiento (monitoreo) de políticas y programas

Esta modalidad responde a la definición más común del seguimiento. A partir de la cadena de valor productivo (matriz lógica de vinculación entre insumos, productos, resultados e impactos) y de la programación operativa, se identifican indicadores que permiten auscultar con cierta periodicidad los avances en la producción de bienes y servicios (productos-entregables) y la realización de los distintos procesos de abastecimiento y de producción de bienes intermedios (administrativos).

El relevamiento de la información operativa de procesos y productos puede combinarse elaborado indicadores de eficiencia, eficacia y productividad.

La periodicidad del relevamiento de la información depende de múltiples factores. Un primer factor es la necesidad de información. Las autoridades de un programa pueden fijar como estándar la supervisión de los avances de un programa con carácter semanal, mensual o trimestral, para tomar algunos ejemplos. Esto puede llevar a solicitar información periódica sobre el estado de los procesos de compras, sobre la cantidad de raciones diarias o mensuales suministradas en establecimientos de desarrollo infantil, o sobre la cantidad de microcréditos otorgados por mes o por trimestre, para tomar algunos de los innumerables ejemplos. Esta información debe ser cotejada con la programación establecida para el período de referencia. Los indicadores de seguimiento se complementan con un análisis más cualitativo sobre los factores que incidieron en el cumplimiento o no de las metas esperadas para este período de referencia.

Un segundo factor tiene que ver con las características de la política o el programa. Por ejemplo, los programas agropecuarios se ajustan a la realidad biológica de los cultivos o de la crianza de animales, por lo que pueden ofrecer información en determinados períodos (y no en otros). Finalmente, otro factor que defina la periodicidad del seguimiento tiene que ver con la factibilidad de obtener información. En muchos casos, esto se relaciona con operativos estandarizados de relevamiento de información, que solo suceden en períodos determinados. Por ejemplo, hay información escolar que es relevada masivamente una o dos

veces en el año, por lo que el proceso de seguimiento debe ajustarse a estos tiempos.

4.2 Evaluación del desempeño

La evaluación del desempeño organizacional ha sido una herramienta que ha encontrado un fuerte impulso con la difusión de la Nueva Gestión Pública y su énfasis en la medición de resultados en el sector público (*performance measurement*). El término resultados, en este caso, alude a distintas dimensiones del logro de las acciones públicas, sean estos logros lo que hemos llamado productos o resultados e impactos, con referencias también a indicadores que miden eficiencia, eficacia y productividad. Más allá del uso del término evaluación, esta herramienta está más asociada al seguimiento del funcionamiento de las organizaciones públicas, a partir del análisis continuo de sus logros (y especialmente, sus productos).

De acuerdo a quienes utilizan esta metodología, el sector público debe explicitar de manera permanente los logros que deben alcanzar (productos y resultados, así como medidas relativas a la eficiencia, productividad u otro criterio). Todas las organizaciones públicas están sujetas a la definición de las medidas para la evaluación de su desempeño, no solo aquellas que realizan sus actividades bajo la forma de proyectos, donde deben estipularse *ex ante* las metas y los recursos necesarios. Hay organizaciones públicas (una gran cantidad de ellas) que producen servicios permanentes, sin término fijo y que, aun así, deben especificar sus medidas de logro o desempeño. Algunos ejemplos: las cárceles, los hospitales, las escuelas, las bibliotecas públicas, las unidades policiales, entre tantas otras.

Las medidas de desempeño más usuales son las siguientes:

- Insumos (inputs), productos (outputs) y resultados (outcomes).
- Medidas asociadas a distintos criterios generales de valor en los resultados (equidad, calidad, pertinencia, coherencia, innovación).
- Medidas de carga de trabajo, eficacia, eficiencia, productividad (Tabla 2).



Tabla 2: Medidas de desempeño en el nivel local

Función Municipal	Medida de carga de trabajo	Medida de eficiencia	Medida de eficacia	Medida de productividad
Secretaria	Números de grupos de minutas de las reuniones del Concejo Municipal preparadas	Horas empleadas por minutas preparadas del Concejo Municipal	Porcentaje de minutas del Concejo Municipal aprobadas y sin enmiendas.	Porcentaje de minutas del Concejo Municipal preparadas dentro de los 7 días posteriores a las reuniones y aprobadas sin necesidad de enmiendas
Biblioteca	Circulación Total	Circulación por empleado	Circulación per cápita	Circulación por cada \$1.000
Reparación de medidores	Número de medidores reparados	Costo por medidor reparado	Porcentaje de medidores reparados que aún funcionan apropiadamente seis meses después	Costo por medidor reparado apropiadamente (p.e. costo total de todas las reparaciones de medidores dividido por el número de medidores que no necesitan reparación después de seis meses)
Personal	Solicitudes de trabajo recibidas	Costo por solicitud de trabajo procesada, costo por vacante ocupada	Porcentaje de nuevas contrataciones, ascensos que han completado el período de prueba y se desempeñan satisfactoriamente seis meses	Costo vacante ocupada satisfactoriamente o sea, el empleado que se desempeña satisfactoriamente seis meses

Fuente: Ammons (1999).

Las medidas de desempeño organizacional requieren de una actualización periódica y permiten una revisión continua de los avances de una organización.

4.3 Seguimiento estratégico

El denominado seguimiento estratégico de la acción pública es un concepto que abarca a una gran variedad de instrumentos y prácticas que pueden llevar el nombre de Tablero de Control o Tablero de Comando de la acción de gobierno, y que en inglés lleva el nombre de *Balanced Score Card*.

Estos instrumentos son sistemas de información gerencial que seleccionan áreas prioritarias de política

o programación para enfocar el seguimiento sobre las variaciones que allí se producen. Pueden incluir información (y su consiguiente relevamiento y actualización periódica) estructural de una colectividad (mortalidad), así como información de productos, procesos y resultados.

Como sistemas de información estratégica y gerencial, están contruidos sobre la base de las preferencias (personales o institucionales) de las y los gerentes públicos, habilitando salidas específicas de información agregada y sistematizada. No obstante ello, representan sistemas de una gran complejidad, que involucran a una multiplicidad de agentes públicos debidamente identificados, que tienen la responsabilidad de actualizar fehacientemente la información que luego será procesada en el nivel gerencial. Los sistemas de estas características necesitan adicionalmente de una instancia externa de validación aleatoria de los datos, para que ejerza la presión requerida para lograr una información precisa.

Si bien los datos más agregados y como totalidad de la muestra son *seguidos* por las máximas autoridades de la organización, es necesario desarrollar aperturas para cada una de los niveles subalternos de la organización, para que el seguimiento sea también una función de todos, con un impacto en el aprendizaje y en la mejora.

4.4 Evaluación desde la perspectiva de los beneficiarios

La evaluación social participativa desde la perspectiva de los beneficiarios busca conocer el funcionamiento y el impacto de un programa, incorporando activamente la participación de sus beneficiarios para mejorar el proceso de toma de decisiones.

Las actividades desarrolladas contribuyen a relevar la perspectiva de los beneficiarios para impulsar y establecer mecanismos participativos en el seguimiento presente del programa, en la generación de mecanismos de consulta permanente y en la propuesta de recomendaciones para reformular los objetivos presentes.

El instrumento recaba información de procesos y productos proveniente de las apreciaciones que hacen los propios destinatarios, tanto desde criterios de eficacia y efectividad, como de calidad y transparencia, con el objeto de producir mejoras en el funcionamiento actual de los programas.



Las técnicas de relevamiento de información son las entrevistas a informantes calificados (destinatarios) y los talleres con beneficiarios a partir de los cuales se extraen valoraciones producidas al calor de los intercambios de pareceres, orientados por preguntas prioritarias. La sistematización de esta información permite generar un conocimiento aplicable a la mejora del programa.

5. Modelos de evaluación de políticas y programas

Ya se explicitaron algunos modelos de evaluación. En este apartado se abocará a distinguir distintos tipos de evaluación siguiendo la caracterización planteada originalmente. Por lo pronto, se distinguirán algunos modelos de evaluación según el foco que tienen (tomando el criterio establecido en Nirenberg y otras, 2000, pero agregando modalidades).

Por un lado, es posible identificar evaluaciones que tienen un foco puesto en el diseño de los programas (o en el perfeccionamiento del diseño): la evaluación de evaluabilidad y la evaluación *ex ante*. Por el otro, evaluaciones que tienen el foco puesto en la valoración de los resultados y de la estructura analítica y operativa que lleva a los mismos: evaluación de la teoría del programa, evaluación de resultados propiamente tal y evaluación de impacto. A continuación se explicitan algunas de estas modalidades.

5.1 Evaluación de evaluabilidad

Por evaluabilidad puede entenderse, en un sentido más restringido, el grado en el que las características particulares del diseño de un programa o plan afectan a la habilidad de proveer una evaluación efectiva (Youtie, Bozeman y Shapira, 1998). De manera más general, la evaluabilidad es definida como las posibilidades que tiene una intervención pública para ser evaluada (Gobierno de Navarra, 2008).

Si bien toda política o programa es evaluable en su sentido más estricto, es necesario saber previamente si el proceso de evaluación será beneficioso en términos de la generación de conocimiento y su aprovechamiento, y es necesario saber también si están dadas las condiciones económicas, organizacionales y políticas para realizar la evaluación. En el primero de los casos, y como lo advierte la definición de Youtie et al. (1998) anteriormente mencionada, la calidad del

diseño de una política y programa incide fuertemente en las posibilidades de obtener un conocimiento más acertado y consecuentemente, en generar las condiciones para un mejoramiento de esta política.

La evaluación de evaluabilidad es el ejercicio que busca concluir acerca de la factibilidad de proseguir con una evaluación más amplia de un programa o política (Marisa Merino, 2007). Las preguntas que guían una evaluación de evaluabilidad son diversas:

- ¿Está estructurado conceptual y analíticamente el programa para facilitar su evaluación? ¿Está claramente explicitado el modelo de intervención del programa y sus alcances?
- ¿Están dadas las condiciones organizacionales y políticas para realizar la evaluación y aprovechar sus conocimientos y recomendaciones?
- ¿Secuenta con los recursos, la logística y la organización necesaria para realizar la evaluación posteriormente?

Desde el punto de vista de esta clasificación de modelos de evaluación, la evaluación de evaluabilidad se justifica en términos de aporte a la revisión de la calidad del diseño de un programa, en vista a aprovechar un ejercicio evaluativo más exhaustivo. El foco principal tiene que ver con la calidad del diseño, y por ende, su objeto es el perfeccionamiento del mismo. Los componentes evaluables (en términos de evaluabilidad) son los siguientes:

- Las características del diseño en relación con la teoría del programa, y con los propósitos y los medios establecidos.
- La especificación técnica de todas variables y dimensiones evaluables, considerando la calidad de los indicadores.
- Las características de la especificación del entramado organizacional de implementación.

Para uno de los más conocidos modelos para la evaluación de la evaluabilidad de una intervención pública (Gobierno de Navarra, 2008), las dimensiones de esta evaluación son las siguientes:

- a. La calidad de la planificación.
 - Calidad del diagnóstico
 - Definición de objetivos
 - Coherencia interna
- b. La calidad del sistema de información.
- c. La claridad en los propósitos de la evaluación.



d. Los recursos destinados a la evaluación: costo, tiempo y recursos humanos.

e. La implicación del personal relacionado con la intervención en el seguimiento y evaluación (Gobierno de Navarra, 2008).

Este modelo, como se observa, incluye además de los componentes técnicos y programáticos de la intervención, aspectos relacionados con los recursos y su dedicación.

Una evaluación de evaluabilidad aporta elementos de juicio sobre la claridad y calidad del diseño del programa y sus componentes (variables, dimensiones, indicadores), y aporta también elementos sobre la factibilidad y mérito de emprender una evaluación más costosa de resultados e impactos a la luz de los recursos existentes y del uso potencial de sus recomendaciones en un contexto político y organizacional determinado. Un interesante caso de aplicación de este tipo de evaluación a un conjunto de programas sociales argentinos puede verse en Aquilino y otros (2013).

5.2 Evaluación ex ante

Otra modalidad de evaluación con foco en el diseño es la evaluación *ex ante*. En realidad, este ejercicio evaluativo puede generar un conocimiento que no solo perfeccione el diseño de la política o del programa, sino que lleve a las autoridades a desistir de su ejecución.

La evaluación *ex ante* puede aplicarse a dos situaciones. Por un lado, es un tipo de evaluación que se practica para ajustar un proceso de formulación de *una* alternativa única, para corroborar si cumple con las condiciones requeridas para su implementación o para ajustar su diseño para facilitar esta ejecución. Por otro lado, es un ejercicio que también puede aplicarse para establecer valoraciones entre *más de una* opción de intervención.

En la primera de estas modalidades, la más usual, la evaluación *ex ante* se propone establecer:

- *La pertinencia de la propuesta* (adecuación de la intervención respecto del problema).
- *La coherencia interna y completitud* (consistencia entre los medios y los fines y alcances de estos últimos).
- *La factibilidad interna y externa* (capacidad organizacional interna para lograr los objetivos y viabilidad en el contexto político y social).

Esta modalidad de evaluación *ex ante* ofrece información a los decisores para distintos propósitos: decidir sobre la aprobación y/o selección de propuestas; perfeccionar la formulación del programa; facilitar el seguimiento del programa durante la ejecución.

A partir de lo mencionado sobre los ejes que integran esta evaluación (pertinencia, coherencia y factibilidad), los elementos de juicio versarán sobre la claridad de la justificación del programa, la coherencia de la respuesta respecto del diagnóstico, la pertinencia de las estrategias adoptadas en función del problema, la claridad y verosimilitud de las metas establecidas, la consistencia entre los medios previstos y las metas establecidas, la consistencia de la secuencia del plan de trabajo, la adecuación presupuestaria, la capacidad institucional del organismo o de los organismos responsables de la ejecución, la claridad de los arreglos interorganizacionales, la viabilidad política de la implementación y la sostenibilidad de las acciones en el futuro cercano. En algunas oportunidades, la evaluación *ex ante* va acompañada de una estimación global del riesgo del programa o proyecto, contemplando algunas dimensiones propias de los ejercicios de auditoría.

La otra modalidad de evaluación *ex ante* se practica cuando es necesario elegir entre más de una alternativa de programa para enfrentar un problema. En estos casos, las opciones de programa son evaluadas a partir de herramientas como el Análisis Costo Beneficio (ACB) o el Análisis Costo Efectividad (ACE). La primera de estas herramientas es utilizada cuando las dimensiones evaluables (productos y sus costos, resultados esperados) pueden monetizarse, y se aplican entre otros, a programas de promoción económica o de infraestructura. El ACE es un instrumento que compara costos monetizados con resultados cuya medición no es monetaria, como por ejemplo, años de vida, años de escolaridad o estado nutricional (Ernesto Cohen y Rolando Franco, 1992).

5.3 Evaluación de Teoría de un Programa

La evaluación de la *teoría de un programa* puede incluirse en las evaluaciones de resultado, pero es pertinente describirla de manera separada.

Por lo pronto, lo primero es distinguir entre lo que



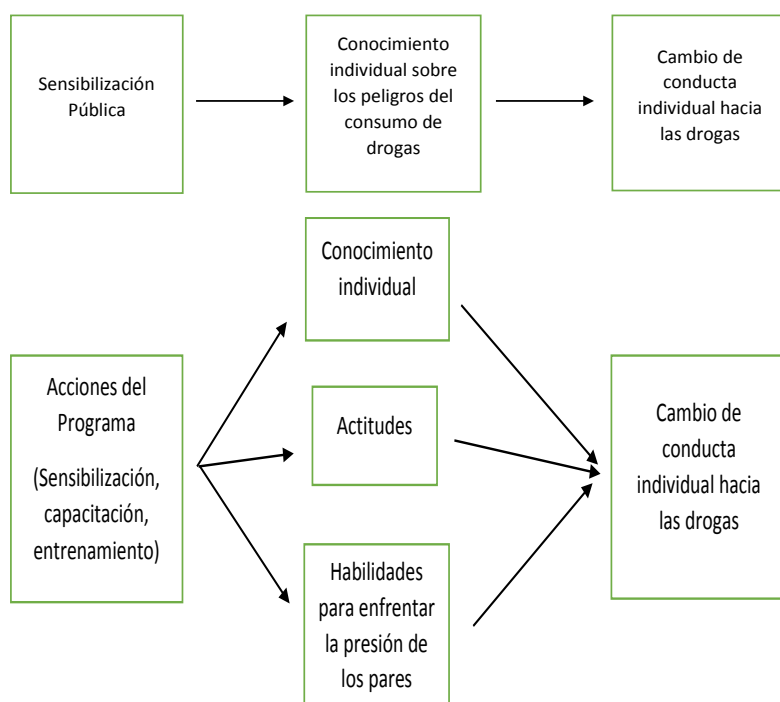
llamamos la teoría del programa y lo que conocemos como la *lógica del programa* (asociada a distintos instrumentos de registro programático como es el marco lógico). La *teoría del programa* es una teoría o modelo explícito acerca de cómo un programa causa los resultados intencionados u observados. En cambio, la *lógica del programa* especifica los insumos, su transformación en productos y su vinculación con resultados e impactos. No especifica los mecanismos que operan detrás de esta lógica, sino que detalla la cadena de valor entre insumos, productos, resultados e impactos (ver Figura 1).

La evaluación de la teoría del programa explicita, en primer lugar, los *mecanismos* relacionales que vinculan sus distintos componentes (muchos de los cuales no están previamente comprobados y por lo tanto son hipotéticos), para facilitar luego la recolección de información y la posterior evaluación de resultados del programa. Pero adicionalmente, y esta es la contribución más importante, aporta elementos de juicio para evaluar la pertinencia de esta teoría, es decir, su comprobación empírica. Ambas funciones son desde luego, complementarias (Ligero Lasa, 2016).

Como toda evaluación, la evaluación de la teoría de un programa ofrece elementos de juicio para conocer cómo se relacionan eventos o fenómenos sociales y qué mecanismos inciden en esta relación. Pero como toda evaluación, y a partir de los aprendizajes, estos elementos de juicio permiten introducir *mejores teorías* en los programas para optimizar sus resultados.

En resumidas cuentas, la evaluación de teoría aporta elementos de juicio para: comprender por qué los programas funcionan (o no) de acuerdo a sus previsiones, esclareciendo los mecanismos que vinculan los distintos componentes de la estructura lógica; identificar componentes o relaciones innecesarias en la secuencia lógica total del programa; identificar posibles nuevos mecanismos; atribuir resultados a la secuencia de acciones y relaciones establecidas en el programa; mejorar, a la postre, el desempeño general del programa.

A continuación, y a los efectos de ilustrar mejor lo que implica una teoría en un programa y los mecanismos implícitos o explícitos que exhibe, en la Figura 2 se ofrecen algunos ejemplos.

Figura 2: Teoría de programas

Fuente: Elaboración propia.

En la figura anterior pueden observarse dos teorías, una más simple y otra más compleja, para un programa para la lucha contra el consumo juvenil de drogas. En el primer caso, la teoría asume implícita o explícitamente que la sensibilización pública altera (positivamente) el conocimiento que los jóvenes tienen sobre los peligros que conlleva el consumo de drogas (mecanismo 1), y que este conocimiento activa un cambio en su conducta respecto del consumo (mecanismo 2).

La explicitación de esta teoría y sus mecanismos, y la constatación de que el consumo de drogas no se ve alterado por esta intervención, revela que el mecanismo 2 no se verifica. El cambio de conducta depende de algo más que el simple conocimiento (en el caso en que se haya generado): depende también de otras actitudes y aptitudes. Por lo tanto, la matriz del programa debe incluir acciones no solo de sensibilización, sino de capacitación individual o de entrenamiento para producir los cambios requeridos en las competencias psicosociales de los jóvenes. Más allá de ello, también es necesario evaluar si los supuestos del mecanismo 1 y sus nuevos agregados también se observan empíricamente.



5.4 Evaluación de resultados

La evaluación de resultados suele ser presentada como el prototipo de la evaluación, y a partir de ella se formulan las definiciones más frecuentes.

La evaluación de resultados tiene como *objeto principal* estimar el grado de cumplimiento del objetivo general del programa, cuyo logro ha sido denominado *resultado del programa*, y que contiene el cambio observado en la *situación problema* que dio origen a la intervención pública (Tabla 3).

Tabla 3: Enunciados programáticos

Enunciados Problemáticos	Enunciados de Intencionalidad	Logro
Efectos	Finalidad	Impacto
Problema principal	Objetivo central/Propósito	Resultado
Causas o Factores directos	Objetivos específicos/ Actividades	Productos
Causas o Factores fuera del control directo de la organización	Objetivos específicos/ Actividades en el marco de acuerdos con otras organizaciones	Productos

Fuente: Elaboración propia.

Además de aportar elementos sobre el grado de cumplimiento del objetivo general del programa, la evaluación de resultados ofrece elementos de juicio sobre una variedad de otros componentes o efectos de la intervención:

- La verosimilitud de las relaciones causales que conectan las intervenciones públicas (los productos) con el resultado observado.
- La verosimilitud de los mecanismos explícitos e implícitos que explican las conexiones entre los distintos componentes de la cadena de valor.
- La existencia de otros efectos no previstos de la intervención del programa.
- La pertinencia de los arreglos institucionales e interorganizacionales de la intervención.
 - La calidad de los productos.
 - La calidad de los procesos.
 - La disponibilidad y calidad de los insumos requeridos en la intervención.

La evaluación de resultados requiere del *diseño de un modelo evaluativo* (Olga Nirenberg y otras, 2000). Este modelo evaluativo debería derivarse sin más de la información suministrada por los instrumentos de planificación/programación del programa, como es por ejemplo, el marco lógico. El problema es que algunos programas o proyectos no han explicitado su modelo de intervención, identificando claramente la cadena de valor que une los insumos con los productos y los resultados e impactos. En ese caso, la primera tarea de los evaluadores es construir el marco conceptual del programa, en una interacción con sus ejecutores, partiendo de la explicitación del objetivo general y el resultado esperados, y luego identificando las dimensiones de intervención, las variables específicas de cada dimensión y los indicadores de cada variable y su fuentes de información. Esta estructura permitirá elaborar un plan de recolección y procesamiento de datos que permitan relacionar las intervenciones (con sus metas esperadas) con su resultado.

Además de los componentes conceptuales, el modelo evaluativo contiene las previsiones organizacionales, logísticas y tecnológicas necesarias para la realización de la evaluación. De acuerdo con las autoras citadas (Olga Nirenberg y otras, 2000), la estructura principal de un modelo evaluativo es la siguiente (Tabla 4):

Tabla 4: Modelo evaluativo

Componente	Contenidos
I. Precisiones conceptuales	Marco teórico general Dimensiones y variables
II. Aspectos metodológicos	Abordaje metodológico Técnicas sugeridas/Instrumentos Fuentes Actores participantes Cronograma de aplicación

Fuente: Nirenberg y otras (2000).

La evaluación se nutre de abordajes tanto cualitativos como cuantitativos, así como hace uso de distintas técnicas de recolección (Tabla 5).

Si bien la evaluación de resultados busca discernir la contribución que el programa realiza al mejoramiento de una situación problema determinada, intentando dar una respuesta a la cuestión de la *atribución* que ha sido



mencionada anteriormente, la estrategia metodológica que adopta no permite extraer conclusiones definitivas sobre esta atribución. Una conclusión más definitiva sobre esta atribución se logra a partir de la realización de evaluaciones de impacto, cuyo costo o requerimiento metodológico no siempre está a disposición de los evaluadores.

Tabla 5: Métodos

Métodos y técnicas cuantitativas	Métodos y técnicas cualitativas
Paradigma positivista empírico	Paradigma interpretativo constructivista
Objetividad	Subjetividad (sentidos, actitudes, motivaciones)
Obtención de datos	Obtención de sentidos de acción
Importancia de la deducción	Preferencia por la inducción
Técnica estadística	Técnicas de caso y de contenidos
Impronta generalizadora	Impronta de particularidades
Diseño planificado preestablecido	Diseño adaptativo y en construcción
Tendencia a programas de gran alcance	Tendencia a programas de alcance más local

Fuente: Elaboración propia.

5.5 Evaluación de impacto

La evaluación de impacto es la metodología evaluativa más rigurosa, porque logra distinguir los efectos de una intervención de los efectos producidos por otros factores externos. Las conclusiones de estas evaluaciones ofrecen evidencia importante para *atribuir* (o no) los resultados observados a una intervención específica (programa o proyecto).

Si bien el concepto tiene otras acepciones, la evaluación de impacto tiene el objeto de determinar en forma más general si el programa produjo los efectos deseados en las personas, hogares e instituciones y si esos efectos son *atribuibles* a la intervención del programa (Judy Baker, 2000). También permite examinar consecuencias no previstas en los beneficiarios, ya sean positivas o negativas.

Para distinguir los efectos de las intervenciones y atribuir incidencia efectiva (o no) a la intervención que hace el programa, la evaluación de impacto despliega una estrategia metodológica compleja con varios componentes:

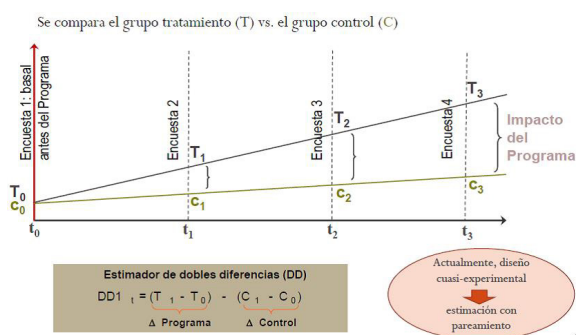
- Creación de un escenario contrafactual o simulado alternativo.
- Conformación de uno o varios grupos de control para compararlos con grupos de tratamiento (quienes están bajo la intervención del programa).
- Aplicación de diseños muestrales (grupos de control y de tratamiento) de dos tipos: diseños experimentales (aleatorios) y diseños cuasiexperimentales (no aleatorios). Posteriormente se describirán con más detalle estos diseños.

En las evaluaciones de impacto son comunes los diseños semiexperimentales. Se conforman al menos dos grupos homogéneos, de tratamiento y de control. Los grupos (como muestras representativas de dos universos también homogéneos) deben tener características similares en una gran variedad de dimensiones sociales, demográficas, económicas y culturales. La diferencia fundamental entre los dos grupos es su condición de destinatarios (o no) de la política o programa bajo evaluación.

Los dos grupos son evaluados en su trayectoria a partir de los componentes evaluables del programa, como por ejemplo: desempleo, estado nutricional, asistencia o rendimiento escolar, mortalidad infantil, entre otros muchos, dependiendo de cada intervención pública.

El efecto del programa sobre la población de tratamiento no es el cambio que experimenta esta población a lo largo de la intervención, sino la diferencia que se produce entre esta población y la población de control. Esta *diferencia*, si la conformación de los grupos ha sido rigurosa, es el valor de bienestar o de mejora que puede ser atribuido al programa, y ningún otro registro. De manera más apropiada, esta metodología utiliza el *estimador de dobles diferencias* entre la población beneficiaria y la población de control (Figura 3).

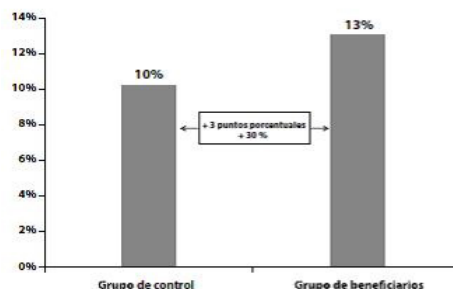
Figura 3: Método de evaluación de impacto cuantitativa



Fuente: Yaschine Arroyo, 2014



Gráfico 1: Efecto sobre la probabilidad de inserción laboral formal de los cursos sectoriales de formación profesional



Fuente: SSPTyEL-DGEyEL- OEDE, sobre la base del SIPA y registros administrativos de las Oficinas de Empleo.

La evaluación de impacto tiene requerimientos metodológicos muy estrictos y depende de mayores recursos para su diseño y para la recolección de la información. Como ejemplo, en Castillo y otros (2013) se evalúa el impacto de los cursos de formación profesional sectorial del Ministerio de Trabajo, Empleo y Seguridad Social de Argentina, donde la probabilidad de los destinatarios de los cursos de obtener empleo es 30% más grande que la de quienes que no han sido destinatarios (Gráfico 1).

6. Reflexiones finales sobre el fortalecimiento e institucionalización de prácticas y sistemas de seguimiento y evaluación de políticas públicas

El mejor desempeño de las políticas públicas, y en especial, de aquellas dirigidas a los sectores vulnerables, depende de muchos factores, entre los cuales se encuentra el mayor acceso a evidencia o conocimiento y una mejor aplicación de éstos en el proceso de toma de decisiones. Una de las principales fuentes de evidencia y conocimiento sobre las políticas públicas es la práctica de ejercicios de seguimiento y evaluación. En este trabajo, además de analizar los alcances y limitaciones de conocimiento generado, se han descrito distintos tipos de actividades evaluativas, tanto ligadas a lo que ha sido denominado el seguimiento, como a la evaluación propiamente tal. En la tabla 6 siguiente se resume la variedad de ejercicios descritos y sus características.

Tabla 6: Modalidades de seguimiento y evaluación

Tipo	Actividad	Alcance	Referencias
Seguimiento	Seguimiento o Monitoreo	Insumos - Procesos - Productos. Con menos frecuencia, resultados.	Bertranou (2018) BID (2011)
	Desempeño organizacional	Insumos - Procesos - Productos - Resultados. Indicadores relacionales: eficacia, efectividad, calidad, economía, productividad	Ammons (1999) BID (2011)
	Seguimiento estratégico	Problemas - Productos - Resultados Algunas veces, procesos	Banco Mundial (2004)
	Evaluación desde la perspectiva de los beneficiarios	Procesos - Productos, desde la perspectiva de los destinatarios.	Banco Mundial (2004)
Evaluación	Evaluación de evaluabilidad	Lógica de diseño. Factibilidad económica y política.	Gobierno de Navarra (2008) Aquilino y otros (2013)
	Evaluación <i>ex ante</i>	Lógica de diseño. Factibilidad política y organizacional.	Cohen y Franco (1992)
	Evaluación de la Teoría de un Programa	Mecanismos que vinculan componentes de cadena de valor	Ligero Lasa (2016), Birckmayer y Weiss (2000)
	Evaluación de Resultados	Resultados. También Insumos - Procesos - Productos.	Nirenberg y otras (2000, 2003)
	Evaluación de Impacto	Impacto (Resultado Neto). También Productos - procesos	Baker (2000), Yaschine (2014), Banco Mundial (2004)

Fuente: Elaboración propia.

Una condición significativa para que las actividades evaluativas se expandan y se fomente su uso en la mejora de las políticas públicas es su institucionalización (Acuña y otros, 2016). Institucionalizar el seguimiento y la evaluación de políticas públicas implica regularizar o normalizar su práctica, integrarla a otras actividades sustantivas de la gestión pública y asignar competencias generales o específicas a determinadas unidades de gobierno.

A continuación se exponen algunas características que debe incorporar un proceso de institucionalización de las actividades de seguimiento y evaluación:

a) Las metodologías de evaluación, si bien diversas, deben estandarizarse y consensuarse para facilitar la comparación o el uso interorganizacional de sus resultados.



b) La institucionalización puede tomar distintas formas en cuanto a normalización y cristalización organizacional, y debe ser gradual, incorporando paulatinamente estándares, aprendizajes, metodologías, comunidades de práctica y organismos.

c) La institucionalización requiere de organismos que asuman roles de rectoría y de organización de las actividades. El diseño institucional puede asumir distintas formas.

d) La evaluación y el seguimiento se fortalecen en el marco de un mayor desarrollo de los sistemas de información pública, que produzcan tanto datos estructurales como de gestión.

e) La institucionalización se fortalece con la integración de las actividades evaluativas al conjunto de actividades relacionadas con la producción de acciones públicas: la planificación y definición de metas, el desarrollo de sistemas de información, la administración financiera y la gestión de personas, entre otros subsistemas públicos. Mientras más esté integrada a estos subsistemas, mayor será su institucionalización.

f) La institucionalización va de la mano del fortalecimiento de las capacidades de gestión de evaluaciones, en el sentido de su producción y de su traslación a las instancias de toma de decisiones. Esta condición requiere de amplios procesos de formación de recursos humanos con habilidades para la programación y conducción de actividades de seguimiento y evaluación.

g) La evaluación y su institucionalización se basan también en la conformación de una comunidad de prácticas, sobre todo al interior del gobierno, que permita procesos de intercambio, aprendizaje, innovación y metaevaluaciones.

Referencias bibliográficas

- ACUÑA, Carlos; MARTÍNEZ NOGUEIRA, Roberto; RUBIO, Jimena, y POTENZA, Fernanda (2016). La evaluación de políticas públicas en Argentina. Sentido, actualidad y perspectivas. *Serie Documentos de Trabajo del IIEP* N°13, Buenos Aires.
- AMMONS, David (1999). Medidas de desempeño en los gobiernos estatales y locales. En Losada i Marrodán (ed.), *¿De burócratas a gerentes?* Washinton, BID.
- AQUILINO, Natalia; ARIAS, Eva; ESTÉVEZ, Sofía, y ECHT, Leandro

- (2013). Hacia un análisis de evaluabilidad de planes y programas sociales. Un estudio sobre 16 iniciativas nacionales. En *Documento de Políticas Públicas* N°124, Buenos Aires, CIPPEC.
- BALLART, Xavier (1992). *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudio de caso*. Madrid, Ministerio de las Administraciones Públicas.
- BAKER, Judy (2000). *Evaluación del impacto de los proyectos de desarrollo en la pobreza. Manual para profesionales*. Washington DC, Banco Mundial.
- BANCO MUNDIAL (2004). *Seguimiento y evaluación: instrumentos, métodos y enfoques*. Washington.
- BERTRANOU, Julián (2015). Planificación pública. Las exigencias de una adecuada problematización. En VV.AA., *Aportes teóricos a la formación de líderes políticos y comunitarios*, Neuquén: Casa de las Leyes, pp. 103-164.
- BERTRANOU, Julián (2018). Webinar Monitoreo de Políticas Públicas, Programa Nacional de Capacitación, INAP, 28 de agosto de 2018.
- BID (2011). Gestión para resultados en el desarrollo en el nivel subnacional, Unidad 6, PRODEV/INDES/BID, Washington.
- BIRCKMAYER, Johanna y WEISS, Carol Weiss (2000). Theory-based Evaluation in practice. What Do We Learn? En *Evaluation Review*, Vol. 24 N°4, pp. 407-431.
- CAMPBELL, Siobhan y otros (2007). *Analysis for policy: evidence-based policy in practice*. Government Social Research Unit, HM Treasury, Londres.
- CASTILLO, Victoria; OHACO, Moira, y SCHLESER, Diego (2013). Evaluación de impacto en la inserción laboral de los beneficiarios de los cursos sectoriales de formación profesional. En *Trabajo, ocupación y empleo*, Serie Estudios/12, Ministerio de Trabajo, Empleo y Seguridad Social, Buenos Aires.
- COHEN, Ernesto, y ROLANDO, Franco (1992). *Evaluación de proyectos sociales*. Siglo XXI Editores, Madrid, Buenos Aires.
- FEINSTEIN, Osvaldo (2007). La evaluación como instrumento de aprendizaje, ponencia presentada en XII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública, Santo Domingo, República Dominicana, 30 oct.- 2 nov.
- FLORES CRESPO, Pedro (2013). El enfoque de la política basado



- en la evidencia. Análisis de su utilidad para la educación de México. *Revista Mexicana de Investigación Educativa*, vol. 18, N° 56, enero-marzo, pp. 265-290.
- GOBIERNO DE NAVARRA (2008). Protocolo de Evaluabilidad de las políticas públicas en la Comunidad Foral de Navarra.
- GUERRERO AMPARÁN, Juan Pablo (1995). La evaluación de políticas públicas: enfoques teóricos y realidades en nueve países desarrollados. *Gestión y Política Pública*, México, Vol. IV, N° 1, México, primer semestre.
- LIGERO LASA, JUAN ANDRÉS (2016). Dos métodos de evaluación: criterios y teoría del programa. En MARTÍNEZ NOGUEIRA, Roberto (comp.) *La evaluación de políticas. Fundamentos conceptuales y analíticos*. Buenos Aires: Corporación Andina de Fomento.
- MARTÍNEZ NOGUEIRA, Roberto, (2016). *La evaluación de políticas. Fundamentos conceptuales y analíticos*. Buenos Aires: Corporación Andina de Fomento.
- MERINO, Marisa (2007), La evaluabilidad: de instrumento de gestión a herramienta estratégica de la evaluación de políticas públicas, Papeles 7/2007, Agencia de Evaluación y Calidad, España.
- NIRENBERG, Olga, BRAWERMAN, Josette, y RUIZ, Violeta (2000). *Evaluar para la transformación. Innovaciones en la evaluación de programas y proyectos sociales*. Paidós, Buenos Aires.
- NIRENBERG, Olga; BRAWERMAN, Josette y RUIZ, Violeta (2003). *Programación y evaluación de proyectos sociales*. Paidós, Buenos Aires.
- PARKHURST, Justin (2015). Appeals to evidence for the resolution of wicked problems: Complexity, contestation, and evidentiary bias. International Conference on Public Policy 2015, Milan.
- PNUD Argentina (2007). *La Evaluación Ex Ante. Estimación del Riesgo Global de Proyectos PNUD*. Anexo IV. Buenos Aires.
- SOTELO, Jorge (2012). Cadena de valor público y planteamiento estratégico, limitaciones y virtudes del modelo, trabajo presentado en el XVII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública, Cartagena, Colombia.
- WEISS, Carol (1999). La investigación de políticas: ¿datos, ideas o argumentos? En WAGNER, WEISS y WOLMAN (comp.) *Ciencias Sociales y estados modernos*, Fondo de Cultura Económica, México.
- WEISS, Carol; MURPHY-GRAHAM, Erin, y BIRKELAND, Sarah (2005).

An Alternate Route to Policy Influence: How Evaluations Affect DARE. En *American Journal of Evaluation*; 26; 12.

YOUTIE, J.; BOZEMAN, B., y SHAPIRA P. (1999). Using an evaluability assessment to select methods for evaluating state technology development programs: the case of the Georgia Research Alliance. En *Evaluation and Program Planning*, pp.55-64.

YASCHINE ARROYO, Iliana (2014). Evaluación de impacto del Programa Oportunidades de México. Presentación realizada en la Escuela de Política y Gobierno de la UNSAM.

Fecha de recepción: 30 de septiembre de 2018

Fecha de aceptación: 4 de febrero de 2019



Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

