



I N A P - INSTITUTO DE ASUNTOS PÚBLICOS
DEPARTAMENTO DE GOBIERNO Y GESTIÓN PÚBLICA



Universidad de Chile

DOCUMENTOS APOYO DOCENTE Nº 17 – Mayo 2007

Nociones de Evaluación de Programas

Mauricio Olavarría Gambi, Ph. D.

Los Documentos de Apoyo Docente (DAD) son una Publicación del Departamento de Gobierno y Gestión Pública del Instituto de Asuntos Públicos de la Universidad de Chile.

Los DAD tienen como objetivo poner a disposición de la comunidad académica la experiencia docente de los/as profesores/as del Instituto de Asuntos Públicos.

Serie de Documentos de Apoyo Docente:

1. Discusión teórica conceptual sobre la disciplina
2. Revisión bibliográfica y exposición de autores
3. Desarrollo de contenidos de los Programas de las Asignaturas
4. Propuesta de Ejercicios, Análisis de Casos e Instrumentos Metodológicos

Editora

Karina Doña Molina, Académica Instructora
Departamento de Gobierno y Gestión Pública
INAP – UNIVERSIDAD DE CHILE

Comité Editorial

Prof. Verónica Figueroa
Prof. Thomas Griggs
Prof. Cristián Pliscoff
Prof. Ariel Ramírez

Asistente de Publicaciones

Diego Barría Traverso

Se autoriza la reproducción total o parcial del material publicado, previa cita de la fuente.

DOCUMENTOS DE APOYO DOCENTE

Nociones de Evaluación de Programas*

Mauricio Olavarria Gambi, Ph. D.**

PALABRAS CLAVES

PROGRAMAS PÚBLICOS –ADMINISTRACIÓN PÚBLICA – METODOLOGÍAS DE
EVALUACIÓN DE PROGRAMAS

* Este documento ha sido preparado en el contexto del programa de formación del Centro de Estudios en Seguridad Ciudadana de la Universidad de Chile, el que es patrocinado por el Open Society Institute de Estados Unidos.

** Profesor Asociado, Departamento de Gobierno y Gestión Pública, Instituto de Asuntos Públicos, Universidad de Chile.

PRESENTACIÓN

El presente documento ha sido preparado para apoyar el proceso de aprendizaje de los alumnos del Instituto de Asuntos Públicos de la Universidad de Chile que se interesan en adquirir nociones acerca de la conceptualización, métodos y prácticas de la evaluación de proyectos y programas públicos. La finalidad del documento, por tanto, es utilidad antes que originalidad. Así, el documento presenta el estado del arte y, para ello, recoge las contribuciones más significativas en el campo de la evaluación de programas.

La evaluación de programas ha llegado a ser una herramienta insustituible en el debate, diseño, implementación y juicio respecto de si las intervenciones del Estado alcanzan las finalidades previstas, de si su implementación es efectiva o si sus resultados e impactos satisfacen las aspiraciones ciudadanas. Diversos programas docentes del Instituto de Asuntos Públicos abordan esta temática, por lo que este documento busca aportar las nociones básicas a quienes desean dar los primeros pasos en la evaluación de las intervenciones del Estado y a quienes desean participar informadamente en el debate sobre los asuntos públicos.

Este documento constituye una sistematización de las notas de clases que he preparado para los cursos sobre políticas públicas, política social y gestión pública de la Escuela de Gobierno y Gestión Pública, Magíster en Ciencia Política, Magíster en Gobierno y Gerencia Pública, y diversos Diplomas impartidos en el Instituto de Asuntos Públicos. El elemento

desencadenante de esta sistematización ha sido el requerimiento de preparación de material sobre este tema para cursos del Centro de Estudios en Seguridad Ciudadana del Instituto de Asuntos Públicos. Por ello debo expresar mi agradecimiento a mi amigo y colega Hugo Frühling y a Ximena Tocornal, una muy talentosa colaboradora suya, por generarme el requerimiento mencionado y que ha sido el estímulo necesario para formalizar estas notas de clases.

MAURICIO OLAVARRIA GAMBI

Abril 2007

ÍNDICE

1. INTRODUCCIÓN: Contexto, Consideraciones y Utilidad de la Evaluación de Programas	8
1.1. Evolución de la Evaluación de Programas	8
1.2. ¿Qué es la Evaluación de Programas?	10
1.3. ¿Para Qué se Evalúan Programas de Intervención Social?	16
1.4. Tipos de Evaluación	21
1.5. Instrumentos de Recolección de Información	27
1.6. Problemas a Considerar en la Evaluación	29
1.7. Los Destinatarios de la Evaluación	31
1.8. El Contexto de la Evaluación	35
2. EVALUACIÓN EX - ANTE	37
2.1. Concepto de Evaluación Ex – Ante	37
2.2. Evaluación de Diseño	38
2.3. Evaluación de Eficiencia de los Programas	39
2.4. Conceptos básicos	41
2.5. El Análisis Costo – Beneficio	42
2.6. El Análisis Costo - Efectividad	46
3. EVALUACIÓN INTERMEDIA	53
3.1. La Evaluación Intermedia	53
3.2. Seguimiento y Monitoreo	56
3.3. Monitoreo de la cobertura	61

3.4. Monitoreo de la Implementación del Programa	68
3.5. Monitoreo de los Avances del Programa	70
3.6. Datos para el Monitoreo	72
4. EVALUACIÓN EX - POST	74
4.1. Concepto de Evaluación Ex – Post	74
4.2. Tipos de Evaluación Ex – Post	75
4.3. Evaluación de Resultados	76
4.4. Evaluación de Impacto	79
4.5. Evaluación de Impacto basada en un Diseño Experimental	84
4.6. Evaluación de Impacto basada en un Diseño Cuasi – Experimental	92
4.7. Evaluación de Programas de Cobertura Total	101
5. IDEAS CENTRALES DEL TEXTO	108
GLOSARIO	110
REFERENCIAS	114

1. INTRODUCCIÓN: Contexto, Consideraciones y Utilidad de la Evaluación de Programas

1.1. Evolución de la Evaluación de Programas

La evaluación de programas sociales ha ganado crecientemente importancia como actividad profesional. La demanda por evaluaciones ha crecido junto con el incremento del número de programas sociales que se han ido implementando en diversos países para mejorar las condiciones de vida de la población. Gobiernos, agencias internacionales y contribuyentes crecientemente desean saber si los programas en que se han invertido grandes cantidades de recursos, tiempo y esperanzas producen los resultados previstos, si es necesario mantenerlos, abandonarlos, introducir rectificaciones o si han alcanzado los objetivos o han producido efectos adicionales no previstos inicialmente.

Los inicios de la evaluación de programas, como actividad profesional, pueden ubicarse en la década de 1930 en Estados Unidos, cuando científicos sociales planteaban la necesidad de aplicar rigurosos métodos de investigación social al análisis de efectos y resultados de los programas sociales y de acción comunitario. Luego, al término de la Segunda Guerra Mundial, en Estados Unidos y Europa, junto con la reconstrucción de post - guerra se dio inicio a grandes programas sociales en áreas de desarrollo urbano y vivienda, educación, capacitación laboral y salud preventiva, planificación familiar y nutrición, y desarrollo rural, las que, por un lado, demandaron inversiones cuantiosas y, por otro, generaron una insistente y creciente demanda por conocer los resultados

de esos programas. Así, a partir de la década de 1950, en Estados Unidos y Europa, la evaluación de programas de gran escala comenzó a ser más frecuente y hoy es una práctica habitual en los países desarrollados.

En los países subdesarrollados, la introducción de la práctica de la evaluación de programas sociales ha ido aparejada con la ampliación de la cooperación de organismos y agencias internacionales y de gobiernos de países desarrollados. Ellos han ido destinando crecientes recursos para desarrollar programas sociales y humanitarios en países subdesarrollados, pero, al mismo tiempo, ha ido creciendo la exigencia de evaluar sus resultados, implementación e impacto producido. El tema de fondo es saber si los recursos fueron bien invertidos y produjeron los resultados que los promotores y financistas de los programas deseaban. Si los resultados no se lograron, también es necesario saber cual habría sido la razón que no permitió lograr los resultados previstos y qué otros efectos se produjeron.

Hacia fines de la década de 1980 la Organización para la Cooperación Económica y Desarrollo (OECD) estimó que un donante promedio desarrollaba entre 10 y 30 evaluaciones al año en países subdesarrollados, mientras que la USAID y el Banco Mundial implementaron unas 250 evaluaciones cada año, cada uno. Según el Banco Mundial (Valadez y Bamberger, 1994), hacia mediados de la década de 1990, alrededor de 100 países subdesarrollados realizaban regularmente alguna forma de monitoreo y actividades de evaluación sobre los programas sociales que implementaban. Muchas de estas evaluaciones son conducidas por agencias internacionales.

Según el Banco Mundial (Valadez y Bamberger 1994), hacia mediados de la década de 1990 ya había un pequeño – pero creciente – número de países subdesarrollados (que incluían a Chile, Colombia, Brasil, India y Malasia) que desarrollaba actividades de evaluación de programas para su propio control de proyectos, planificación financiera y análisis de política pública. En Chile el establecimiento de una Unidad de Evaluación de Programas Gubernamentales en el Ministerio de Hacienda, en la década de 1990, hubo extendido e intensificado la evaluación de programas como práctica regular de la gestión del Estado.

1.2. ¿Qué es la Evaluación de Programas?

a) Definición de la Evaluación de Programas

Rossi y Freeman (1993) se refieren a la evaluación de programas como “evaluación” o “investigación de evaluación” indistintamente. Para ellos “la evaluación de programas es la aplicación sistemática de procedimientos y métodos de investigación social para evaluar y sopesar la conceptualización, diseño, implementación y utilidad de programas de intervención social”. En otras palabras, los evaluadores usan metodología de investigación social para juzgar, analizar, evaluar y mejorar las formas en que programas y políticas sociales son conducidos, desde las etapas más tempranas de definición y diseño, hasta las etapas de desarrollo e implementación.

Los altos oficiales públicos, autoridades de gobierno, organizaciones donantes, administradores de programas necesitan distinguir aquellos programas útiles, que están produciendo efectos positivos, de aquellos

inefectivos y/o ineficientes. De este modo, ellos pueden planear, diseñar e implementar nuevos esfuerzos que alcancen una efectiva y eficiente implementación, como asimismo generar los resultados y el impacto deseado sobre los miembros de la comunidad y su medioambiente.

Las preguntas que normalmente se hacen los evaluadores, impulsores, donantes, sostenedores y administradores son del siguiente tipo:

- ¿Qué problemas requieren de nuevas intervenciones, o expandir o modificar los programas sociales actualmente existentes?, ¿cuál es la naturaleza y ámbito de esos problemas sociales?, ¿dónde se localizan y a quién afectan?
- ¿Qué intervenciones sociales posibles, probablemente aminorarían significativamente el problema social?
- ¿Cuál es la población objetivo apropiada para una particular intervención?, ¿llega la intervención a su población objetivo?
- ¿Está la intervención siendo implementada de la manera como se pensó?, ¿es la intervención efectiva?
- ¿Cuánto cuesta?, ¿cómo son los costos en relación a la efectividad del programa y a sus beneficios?
- ¿Qué resultados e impactos ha producido la intervención?

Las respuestas a estas preguntas no sólo son necesarias para grandes intervenciones, diseñadas para ser implementadas en todo el país, como los programas de salud, educación o capacitación, sino, que también son preguntas aplicables a pequeños y especializados programas que pueden llevarse a cabo en una localidad, comunidad o servicio público.

Dar respuesta a estas preguntas es el trabajo que desarrollan las personas que llevan a cabo evaluación de programas.

b) Criterios de evaluación

Los criterios de evaluación son principios que permiten distinguir, calificar, juzgar, estimar, afirmar que se ubican en el plano conceptual del estudio y actúan como referente para organizar la apreciación y lectura que el evaluador realiza de la información reunida. Entre los criterios más utilizados el Centro de Estudios en Seguridad Ciudadana (CESC) de la Universidad de Chile (2005) distingue a los siguientes:

- *Pertinencia*: La explicitación de este criterio alude a la pregunta ¿hay consistencia entre el problema identificado y el proyecto o programa que se está implementando o implementará?
- *Relevancia*: Alude al grado de significación e impacto que las acciones del proyecto o programa tienen para las personas a quienes está dirigida la intervención.
- *Coherencia*: Hace referencia al grado de integración y consistencia de los diferentes elementos del proyecto.

- *Efectividad*: Expresa el nivel de logro de los objetivos, resultados e impacto esperado, en relación a los recursos a disposición de la intervención.

Cohen y Franco (2000: 107) expresan que la efectividad es una relación entre resultados alcanzados y objetivos esperados. Algebraicamente lo expresan de la siguiente forma:

$$\text{Efectividad} = \frac{\text{Resultados}}{\text{Objetivos}}$$

Los resultados son los productos concretos entregados por el programa. Éstos, a su vez, están supeditados o en relación a los recursos asignados. La identificación de los resultados expresa, entonces, una medida de eficiencia, esto es, una relación entre los recursos que se disponen y los productos que se obtienen.

Los objetivos, por otro lado, representan los estados que se pretende alcanzar a través del desarrollo de las actividades y la aplicación – o consumo – de los recursos considerados en la intervención. La medición de los objetivos hace referencia a una medida de eficacia, toda vez que este concepto expresa el grado de cumplimiento del objetivo

El concepto de efectividad expresa una relación recíproca y dinámica de los conceptos de eficiencia y eficacia en la implementación de un programa o proyecto. Es, por tanto, un concepto relativo y no absoluto. Así, los programas o proyectos implementados podrán ser clasificados en relación a la efectividad demostrada.

Los criterios de evaluación son desglosados en categorías más concretas que, a su vez, puedan ser observables y medibles, de tal manera de generar **indicadores** para su tratamiento analítico. La formulación de indicadores está presente en todo tipo de evaluación. Se utilizan para saber si se han alcanzado logrado los objetivos del proyecto, los productos y resultados esperados, así como para evaluar la gestión del mismo.

d) Concepto de Indicadores

Un indicador es la expresión cuantitativa de un cierto fenómeno o realidad que busca ser influido por la acción del proyecto o programa. Es una unidad de medida que expresa normalmente una relación entre dos dimensiones. Ejemplo de indicadores en seguridad ciudadana son la tasa de homicidios por cada 100.000 habitantes, tasa anual de victimización, el índice mensual de denuncias de delitos, dotación de policías por cada 100.000 habitantes, dotación de jueces por cada 100.000 habitantes y similares.

La ventaja que ofrece trabajar con indicadores es que permiten la comparación. Por ejemplo, el cuadro 1 muestra una comparación entre

indicadores de victimización y homicidios de países latinoamericanos en la década de 1990.

Cuadro 1: Victimización y tasas de homicidios en América Latina en la década de 1990*

País	Victimización (% de hogares)	Tasas de homicidios (por 100,000 habitantes)
Argentina	38	4.8
Bolivia	34	Nd
Brasil	38	23.3
Chile	31	3.0
Colombia	37	89.5
Costa Rica	37	5.6
Ecuador	42	15.3
El Salvador	47	41.3
Guatemala	55	75.3
Honduras	37	45.0
México	43	17.8
Nicaragua	37	18.3
Panamá	28	10.9
Paraguay	33	4.0
Perú	38	11.5
Uruguay	25	4.4
Venezuela	45	16.0
<p>* Las tasas de victimización son de Gaviria y Pagés. Las tasas de homicidios han sido tomadas de las fuentes más recientes y confiables posibles, especialmente en el caso de los países de Centroamérica. Fuentes: Gaviria y Pagés (1999); Buvinic et al. (1999); De Mesquita Neto (2002); CIEN (2002); PNUD (2001); Krug et al. (2002).</p>		

Nota: Cuadro tomado de Cruz (2003: 28)

1.3. ¿Para Qué se Evalúan Programas de Intervención Social?

Según Rossi y Freeman (1993), siendo la evaluación de programas la aplicación de métodos de investigación social al análisis de intervenciones sociales, es también una actividad política y gerencial. Ello porque la evaluación de programas genera información relevante a los estamentos políticos desde los que emergen decisiones de política pública y destinación de recursos para la planificación, diseño, implementación y continuidad de programas orientados a mejorar las condiciones de vida de las personas.

a) Razones para implementar la evaluación de programas

Las evaluaciones de programas son ejecutadas por una variedad de razones:

- Para juzgar el valor de programas en ejecución y estimar la utilidad de los intentos por mejorarlos. Verifica la efectividad de la estrategia de intervención, el grado de cumplimiento de los objetivos y las razones de éxito y/o fracaso del programa.
- Para evaluar y sopesar la utilidad de nuevos programas e iniciativas. Así aumenta la calidad de los diseños de intervención.
- Para incrementar la efectividad de la administración de programas. Permite retroalimentar a autoridades y administradores sobre el funcionamiento de los programas y, de esa manera, aumenta la calidad de las decisiones gerenciales sobre el programa.

- Para la rendición de cuentas ante los patrocinadores del programa: autoridades de gobierno, líderes políticos, agencias donantes, ciudadanos.

b) Utilidad de la evaluación de programas

Aunque la variedad de usos que pueda tener la información que provee la evaluación de programas es amplia, desde un punto de vista de su utilidad para los procesos de toma de decisiones sus usos pueden ser los siguientes (Rossi y Freeman, 1993).

- *Decisiones de Autoridades Públicas*

Los resultados de una evaluación de programas es información valiosa para autoridades de gobierno que tienen la autoridad y responsabilidad de destinar recursos y dar forma a los programas sociales. Por ejemplo, una evaluación sobre la implementación del Plan Cuadrante¹, los resultados, el impacto, el beneficio neto o efectividad alcanzados por esta intervención arrojará información relevante al mando de Carabineros y a las autoridades gubernamentales a cargo de la seguridad ciudadana. Ello les permitirá tomar decisiones, basadas en información cierta, acerca de si se debe mantener tal como fue concebido el programa, introducirle rectificaciones o discontinuarlo.

¹ El Plan Cuadrante es un plan de vigilancia policial por cuadrantes territoriales implementado por Carabineros de Chile, con la finalidad de incrementar los niveles de seguridad ciudadana en el país. Mayor información sobre este Plan puede encontrarse en el siguiente sitio web <http://www.carabineros.cl/cuadrante/cuadrante.htm>

- *Aprobar/apoyar o rechazar/oponerse al programa*

Muchas decisiones políticas son de este tipo. Una variedad de actores influyen la manera en que programas sociales son diseñados e implementados. Para ellos los resultados de una evaluación de programas les dará información valiosa para fijar su posición en el debate de política pública que se dará en torno al programa o como votarán, si la decisión se toma en una votación en el Congreso Nacional, en el Concejo Comunal u otro organismo colegiado de decisión.

Por ejemplo, una evaluación ex - ante del proyecto "Cámaras de Vigilancia para el Centro de San Andrés" permitirá al Concejo Municipal y a las autoridades a cargo del Fondo Nacional de Seguridad poder estimar si, de acuerdo al conocimiento disponible y a las lecciones que han dejado experiencias similares, es posible esperar - razonablemente - una disminución del temor y victimización en el centro de San Andrés y cual es el beneficio neto que generaría, dado el monto de la inversión social. Es decir, en primer término, la evaluación del proyecto permitirá estimar si el diseño del proyecto podrá - razonablemente - disminuir la incidencia delictual en el área en que se aplicará. Por otro lado, la evaluación ex - ante permitirá establecer cual es el beneficio neto esperado del proyecto y si hay otros proyectos que arrojen beneficios netos similares o mayores a costos menores o similares que las cámaras de vigilancia.

Los antecedentes que aporte la evaluación del proyecto sobre "Cámaras de Vigilancia para el Centro de San Andrés" permitirá, tanto al Concejo Municipal como a los responsables del Fondo Nacional de Seguridad, tomar una posición respecto de si deben o no otorgar

financiamiento a ese proyecto. Del mismo modo, la evaluación que se haga de la propuesta permitirá fundamentar la decisión que se adopte respecto de la aprobación o rechazo de los fondos para esta iniciativa.

- *Uso directo de la evaluación*

Dado que las evaluaciones entregan documentadamente sus hallazgos y, a menudo, identifican las causas de los problemas que estarían afectando el rendimiento de los programas, los resultados de los estudios de evaluación son usados por los formuladores de política pública para introducir rectificaciones o reformas en los programas. Por ejemplo, evaluaciones sobre proyectos similares a las “Cámaras de Vigilancia para el Centro de San Andrés”— y que ya se hubieran implementado – permitirán establecer si este tipo de intervenciones es efectiva en la reducción de la percepción de temor y la victimización y que rectificaciones serían recomendables para incrementar sus grados de efectividad.

- *Legitimación y Rendición de Cuentas*

La información provista por la evaluación de programas permitirá a las autoridades, ciudadanos y representantes poder conocer como han sido implementadas las intervenciones, si llegan a su población objetivo, su impacto, sus resultados, beneficios y efectividad alcanzados. De este modo, las autoridades y administradores de programas podrán informar a los ciudadanos acerca de los logros alcanzados con las intervenciones implementadas, el destino y utilización de los recursos públicos. De este modo, por un lado, se favorece la participación de los ciudadanos en los

asuntos públicos, se expande el debate, lo que – a su vez – permite un mayor flujo de ideas, perspectivas y traer a la vista aspectos que inicialmente podrían no haber sido previstos. Con ello, el programa y su implementación ganan en legitimidad técnica, política y ciudadana.

c) Variables que afectan la utilización de la evaluación de programas

- Relevancia del programa evaluado
- Comunicación entre investigadores y quienes usan los resultados
- Procesamiento de la información por quienes la usarán
- Aplicabilidad de los resultados de la investigación
- Involucramiento de los usuarios de los resultados de la investigación.

d) ¿Cómo mejorar el aprovechamiento de la evaluación de programas?

- Los evaluadores deben entender el marco cognitivo de los formuladores de políticas públicas. Los reportes deben ser apropiados para una audiencia de políticos. Si son muy sofisticados probablemente no serán consumidos.
- Los resultados de la evaluación deben ser oportunos y accesible. Deben estar disponibles cuando la oportunidad lo requiere.

- La evaluación debe respetar la diversidad de intereses que hay sobre el programa. El evaluador debe explicar las diferencias en valores y perspectivas entre los usuarios de los servicios y el evaluador. Es particularmente importante cuando una investigación es desarrollada por un equipo específico.
- Los planes de utilización y evaluación de los resultados deben ser parte del diseño de evaluación. Deben incluir potenciales usos en enseñanza, comunicación de los resultados a formuladores de políticas públicas y a la comunidad de especialistas temáticos y evaluadores.
- La evaluación debe incluir una evaluación de la utilización. El diseño de evaluación debe incluir el establecimiento de criterios que permitan verificar si ha habido una exitosa utilización de los resultados de la investigación.

1.4. Tipos de Evaluación

Los esfuerzos de evaluación pueden ser clasificados según la etapa del proyecto en que se implementan, la finalidad perseguida en el estudio de evaluación, los instrumentos utilizados y la ubicación del evaluador.

a) Según la etapa del proyecto

- *Evaluación ex - ante.* Corresponde a las evaluaciones y análisis efectuados antes del inicio de la intervención. Luego de identificado, conceptualizado y dimensionado el problema, este tipo

de evaluación tiene por finalidad identificar la consistencia entre el problema identificado y el diseño de la intervención, por un lado, y, por otro, seleccionar aquella alternativa de intervención (programa o proyecto) que más eficientemente permita resolver o aminorar el problema que se abordará. En el primer caso se hará una evaluación del diseño del proyecto o programa y, en el segundo caso, se podrá implementar una evaluación costo – beneficio o evaluación costo – efectividad.

- *Evaluación intermedia.* Hace referencia a las evaluaciones efectuadas durante la implementación de la intervención, con la finalidad de constatar el estado de avance, identificar áreas y situaciones críticas, corregir problemas y, si fuera el caso, enmendar rumbos. Este tipo de evaluación entrega valiosa información a las autoridades y gerentes de programas sobre la marcha de la intervención, lo que le permite tomar decisiones orientadas a implementar rectificaciones oportunas y necesarias. Ejemplo de este tipo de evaluación son las evaluaciones intermedias de resultados y monitoreo (o seguimiento).
- *Evaluación ex – post.* Aluden a las evaluaciones que se hacen luego de concluida la fase de implementación de la intervención. Su propósito es determinar si los objetivos establecidos a la intervención se han cumplido – o no – y los cambios que se pueden verificar en la realidad intervenida como consecuencia del proyecto o programa. La finalidad última de este tipo de evaluación es poder identificar lecciones y aprendizajes que se deducen del proyecto o programa que se ha puesto en práctica. Ejemplo de evaluaciones ex – post son

las evaluaciones de resultados y las evaluaciones de impacto. Con frecuencia también se hacen evaluaciones costo – beneficio y/o costo – efectividad ex – post con la finalidad de estimar la relación entre los resultados alcanzados por una intervención y los recursos invertidos en ella. En todo caso, para poder capturar adecuadamente los resultados e impactos que ha provocado la intervención es necesario que medie un tiempo entre la finalización de la intervención y el inicio del estudio de evaluación. Este plazo variará dependiendo del periodo de maduración del proyecto o programa en cuestión.

b) Según la finalidad de la evaluación

- *Evaluación de diseño.* Corresponde al juicio que se hace sobre la coherencia entre el problema identificado, el contexto en que se dará la intervención y la organización interna del proyecto o programa. Entre estos últimos elementos se cuentan el marco conceptual, los objetivos, métodos, cronogramas de trabajo, recursos asignados, acciones emprendidas o a emprender, y jerarquía, secuencia y vinculación entre esas actividades. La evaluación del diseño busca responder la pregunta de si hay razones para pensar que el programa, tal como ha sido diseñado, podrá razonablemente resolver el problema, o al menos aminorarlo.
- *Evaluación de eficiencia.* Este tipo de evaluación se orienta a establecer la relación entre los resultados alcanzados o previstos y los recursos invertidos en el programa. Los ejemplos de este tipo de evaluación son las evaluaciones costo – beneficio y evaluación

costo – efectividad. Cuando estas evaluaciones son realizadas ex – ante su finalidad es poder identificar la alternativa de intervención, entre todas las disponibles, la que maximiza los resultados minimizando los costos. Cuando se realizan ex – post su finalidad es identificar la relación de costos y resultados y si ella diverge de las expectativas iniciales que se tenían, cuando se decidió la intervención.

- *Seguimiento o Monitoreo.* Cohen y Franco (2000) hacen sinónimo seguimiento y monitoreo. La definen como una actividad gerencial interna, que se lleva a cabo durante el periodo de ejecución del proyecto o programa, y que tiene por finalidad que los procesos de trabajo que se han implementado permitan entregar los productos comprometidos y alcanzar los objetivos establecidos, de acuerdo al plan trazado.
- *Evaluación de resultados.* Busca identificar el grado de cumplimiento de los objetivos que se le establecieron al proyecto o programa. A su vez, el grado de cumplimiento de los objetivos se expresa en los indicadores que presenta la ejecución del proyecto. Así, la evaluación de resultados busca comparar las metas, expresada en los indicadores que se esperaba que el proyecto o programa alcanzara, con los indicadores que expresan el grado de cumplimiento efectivo o resultados realmente alcanzados por la intervención.
- *Evaluación de impacto.* Se orienta a identificar los cambios producidos que puedan ser razonablemente asociados a la

intervención. Los impactos que identifica este tipo de evaluación pueden ser impactos previstos e impactos no previstos, deseados o no deseados. Alternativamente, la evaluación de impacto también responde a la pregunta acerca de qué hubiera ocurrido si no se hubiera implementado el proyecto o programa. Los métodos normalmente usados en la evaluación de impacto son los diseños experimentales y cuasi - experimentales. En las evaluaciones de impacto de programas de cobertura completa se usan métodos basados en los llamados controles genéricos y controles sombras.

c) Según los instrumentos utilizados

El Centro de Estudios en Seguridad Ciudadana (CESC) de la Universidad de Chile (2005) distingue a los siguientes tipos de evaluación de acuerdo al tipo de instrumentos utilizados:

- *Evaluación cuantitativa.* Consiste en evaluaciones que utilizan *metodologías cuantitativas de recolección de información.* Los instrumentos, tales como, la encuesta o cualquier otro sistema de registro, tipifican la información en términos numéricos o de frecuencia y, luego, las bases de datos generadas son analizadas estadísticamente. Entre las encuestas más utilizadas en seguridad ciudadana se encuentran las de victimización y, entre los sistemas de registros, aquellos, tales como, estadísticas de denuncias, sentencias, juicios, entre otros.
- *Evaluación cualitativa.* Consiste en evaluaciones que utilizan *metodologías cualitativas de recolección de información.* Los

instrumentos de recolección de información cualitativa permiten acceder a significados, opiniones, actitudes y vivencias, a través del registro escrito de discursos que luego son analizados de acuerdo a diversos modelos. Los instrumentos más utilizados son las entrevistas (tanto individuales como grupales), el análisis de prensa y la observación participante.

- *Evaluación mixta.* Consiste en evaluaciones que utilizan de manera complementaria *metodologías cuantitativas y cualitativas de recolección de información.*

d) Según la ubicación del evaluador

Desde el punto de vista de quien realiza la evaluación, Cohen y Franco (2000) distinguen los siguientes tipos de evaluación.

- *Evaluación interna.* Es una evaluación que se lleva a cabo dentro de la unidad o equipo a cargo de la gestión del programa. Aunque la evaluación efectuada por los mismos equipos que implementan el proyecto tiene la ventaja que los evaluadores conocen los aspectos sustantivos del problema y del programa, y con ella se reducen las resistencias a la evaluación, la gran dificultad que presenta es que es muy difícil superar el conflicto que se produce entre el interés de hacer una evaluación objetiva y el interés de los responsables del programa de mostrarse efectivos en su gestión.
- *Evaluación externa.* Es implementada por evaluadores ajenos a la unidad o equipo humano encargado de implementar el programa y

que han sido contratados específicamente para ese fin. Estos evaluadores son personas con reconocimientos en las áreas que evaluarán y, en ello, aplican métodos estandarizados de evaluación. Aunque la ventaja con la que se presenta la evaluación externa es su mayor objetividad, dado que el evaluador no está comprometido con el diseño o gestión del programa, también es necesario reconocer que frecuentemente se producen desacuerdos entre los expertos respecto de la interpretación de los hechos sociales, que no existe un método único e infalible para realizar evaluaciones y que los expertos cargan con sus propios prejuicios a la hora de analizar los hechos.

- *Evaluación conjunta.* Busca combinar los esfuerzos de equipos externos e internos en un proceso de evaluación de programas, de modo de aprovechar las ventajas de ambos tipos de evaluaciones.
- *Evaluación participativa.* Tiene por finalidad acercar a evaluadores y población beneficiaria del programa. Ello permite capturar información relevante sobre la satisfacción de los usuarios, aspectos sustantivos de la distribución de los productos, de la cobertura, resaltar las dimensiones del programa que resultan más relevantes para los usuarios e información del contexto en que se desenvuelven los usuarios y que resulta relevante para la gestión del programa.

1.5. Instrumentos de Recolección de Información

- *Uso de indicadores sociales.* Permiten recoger información para estimar la evolución que ha tenido el problema sobre el que se está

implementando la intervención que se evaluará y responder preliminarmente a las siguientes preguntas: ¿está creciendo el problema?, ¿está disminuyendo? o ¿está estancado? El análisis de indicadores sociales también permite estimar el tamaño, localización y distribución de la población afectada por el problema. El uso de indicadores también nos permitirá hacer una primera estimación del efecto de programas sociales actualmente en operación.

- *Análisis de registros de agencias y/o servicios.* Puede ser útil si los servicios /agencias mantienen excelentes registros de sus usuarios y actividades, pero usualmente las situaciones sociales que llegan a la categoría de problema son aquellas que han escapado al control de la agencia a cargo.
- *Encuestas y Censos.* Estos instrumentos proveen valiosa información cuantitativa que nos permite responder a la pregunta de ¿Cuán grande es el problema? Aunque estos instrumentos son técnicamente complejos y de alto costo proveen información que permiten estimar, de mejor manera, la incidencia de un problema social y su evolución.
- *Instrumentos cualitativos,* como grupos Focales, estudios etnográficos (antropológicos), entrevistas en profundidad, observación participante permiten entregar valiosa información sobre las características básicas de un hecho o problema social. Este tipo de instrumento permite responder a la siguiente pregunta: ¿de qué se trata el problema?

1.6. Problemas a Considerar en la Evaluación

a) Problemas políticos y organizacionales

Un tema central es el poder que ejercen las agencias que desarrollan evaluaciones o los evaluadores sobre el destino del programa, sobre los recursos que se le asignarán o sobre la evolución y continuidad de él. Así, estas agencias son percibidas como poderosas fuerzas o centros de poder que necesitan ser controlados. Por ejemplo, ¿cuál es el poder de DIPRES del Ministerio de Hacienda?

Asimismo, es conveniente hacer notar que recurrentemente algunas evaluaciones son conducidas por organizaciones no gubernamentales, financiados por donantes que tienen un interés particular en evaluar ciertos aspectos de los programas.

El tipo de información que las evaluaciones de programas requieren es, con cierta frecuencia, sensible y puede generar situaciones políticamente delicadas y crear complejidades organizacionales, en cuanto a que puede ser información que no está rutinariamente disponible (hay que prepararla). El modo como la información que arroja la evaluación es diseminada y usada también puede plantear problemas políticos.

b) Problemas de dirección de la evaluación

A menudo muchos esfuerzos de programas no identifican claramente los usuarios del estudio. Eso lleva a que los consumidores del estudio sientan que no se les da la información que ellos requieren, o que

la información se presenta de una forma que no se entiende fácilmente o que llega tarde. Por ejemplo, una evaluación de impacto de programas del FOSIS hecha en 1993 por el Departamento de Economía de la Universidad de Chile, que incluía profusa información estadística y sofisticada metodología, fue vista como un excelente trabajo, aunque muy académico y no entendida por personas a cargo de programas, lo que derivó en “falta de recomendaciones concretas para mejorar o reformular” las intervenciones(ver: FOSIS 1995).

A menudo los gerentes de programas se resisten a la evaluación porque ven que ella puede debilitar la justificación del programa, restarle recursos y muy raramente podrían esperar beneficios – como incremento de recursos o similares – de una evaluación. Por otro lado, la obtención de una aprobación es un largo proceso burocrático y desarrollar la evaluación y presentar sus resultados requiere también mucho tiempo.

c) Problema con la focalización del estudio

Usualmente los reportes que se presentan como evaluación de un programa más bien presentan reportes administrativos de actividad (actividades por unidad de tiempo, gasto del programa, recursos invertidos y similares), pero pocos estudios abordan los temas acerca de cómo ellos operan, si llegan a la población definida como beneficiaria, cómo se sostienen, si logran los resultados previstos o si son capaces de producir los impactos que se esperan. De este modo, poco se sabe si grandes programas sociales logran sus objetivos y producen los cambios o efectos para los que fueron diseñados. Por ejemplo, evaluaciones sistemáticas de los programas en seguridad ciudadana – como Comuna Segura, Plan

Cuadrante y similares – o programas de intervención social – como los Programas Puente, Chile Solidario, Chile Barrio, entre otros – permitirían arrojar información valiosa sobre su efectividad, eficiencia, resultados e impacto.

d) Problemas metodológicos

Normalmente se producen agudos reclamos acerca del rigor metodológico de estudios que se presentan como evaluación de programas. Los principales temas apuntan a la calidad de los datos, a lo adecuado del diseño de evaluación y al sesgo que pudiera haber en diferentes opciones metodológicas.

1.7. Los Destinatarios de la Evaluación

En el diseño de una evaluación es esencial considerar quienes serán los principales “consumidores” de la investigación de evaluación de impacto.

Los consumidores o destinatarios de estudio de investigación es un grupo múltiple. También debe considerarse que en el desarrollo de su trabajo los evaluadores se ubican en el centro de una red en que diversos intereses se confrontan. Habrá personas o grupos que se sienten “propietarios” de la idea que originó el programa o del programa mismo. Habrá otros que tienen expectativas sobre la evaluación, bien porque la propusieron, impulsaron, defendieron, aprobaron o se opusieron. Puede que estos grupos sean incluso contendores. Los intereses de estos grupos y de otros como los funcionarios que trabajan en el programa, los

beneficiarios o los que aspiran a ser beneficiarios, los rechazados por el programa, las autoridades y sus detractores, los científicos sociales que se dedican a la evaluación de programas se encontrarán entre los "consumidores" del estudio de evaluación.

Los evaluadores deben tener claro que hay una multiplicidad de "consumidores" de los estudios de evaluación. Según Rossi y Freeman (1993) esto genera dificultades de tres principales maneras:

(i) Los evaluadores a menudo están inseguros acerca de qué perspectiva tomar: ¿la de la sociedad como un todo?, ¿la del servicio público involucrado?, ¿la de los funcionarios del programa?, ¿los beneficiarios?, ¿la de los "críticos" del programa?, ¿cuál?

(ii) Debe tenerse presente también que el hecho que exista una multiplicidad de "consumidores" generará dificultades al desarrollo de la evaluación y a los resultados que ella produzca.

(iii) Debe tenerse claro también que quienes impulsan la evaluación de un cierto programa, usualmente plantean la necesidad de una evaluación cuando perciben que los resultados del programa pueden estar contradiciendo el tipo de políticas e intervenciones que ellos promueven.

En este ambiente, el evaluador debe tener presente que hay una variedad de intereses sobre los resultados del estudio de evaluación y que esos intereses pueden estar en conflicto.

Conducir una evaluación es aportar información que puede ser usada para sostener un juicio, opinión o decisión. Esta precisión es clara y útil en abstracto, pero es a menudo difícil de establecer en la práctica. Suele ocurrir que los “consumidores” de las evaluaciones están más orientados a percibir los resultados de una evaluación como juicios críticos al programa en cuestión y en consecuencia a reaccionar defensivamente, cuando son apoyadores del programa, o a usar la evaluación para intensificar las críticas al programa, cuando son opositores.

¿Quiénes son los “típicos consumidores” de un estudio de evaluación? Quienes tienen algún interés o están involucrados en el programa o los resultados de la evaluación típicamente son:

- *Autoridades públicas y altos funcionarios públicos.* Estos son personas que están en posición de decidir si un programa será implementado, seguirá, se discontinuará, se expandirá o reducirá.
- *Patrocinadores del programa o de la evaluación del programa.* Estos pueden ser líderes políticos o sociales que apoyaron o que emitieron opiniones durante el debate que se dio previo a la decisión sobre la implementación del programa, organizaciones nacionales o internacionales que promueven y/o financian el programa o su evaluación.
- *Población beneficiada.* Personas, hogares, organizaciones en otras unidades que participan en el programa o reciben los beneficios o servicios de la intervención que está en evaluación.

- *Gerentes del programa.* El grupo responsable por supervisar y coordinar la intervención que está siendo evaluada.
- *Personal del programa.* Funcionarios responsable por la aplicación de la intervención o programa
- *Evaluadores.* Personas encargadas de diseñar y conducir la evaluación.
- *Competidores del programa.* Organizaciones o individuos que compiten con el programa por capturar los recursos que el programa capturó, o por otros recursos disponibles.
- *Consumidores en el contexto.* Organizaciones, grupos o individuos y otras unidades ubicadas en el contexto inmediato del programa. Por ejemplo: autoridades municipales, regionales, parlamentarios o personas con influencia política que se ubican en el espacio o cerca del espacio en que se implementa el programa.
- *Comunidad de evaluadores.* Otros evaluadores que leen el estudio y comentan sobre la calidad técnica de la evaluación.

Aunque estos grupos no son todos los que en cierto momento pueden estar interesados en el programa o en su evaluación, son éstos los que más a menudo ponen atención al desarrollo de un programa y los resultados de una evaluación sobre él.

1.8. El Contexto de la Evaluación

Los evaluadores no realizan su trabajo en el vacío. Hay un contexto en el que ellos trabajan sobre el que tendrán una mutua influencia. El valor del trabajo de los evaluadores depende de su utilización, por lo que es necesario entender el contexto en el cual ellos desarrollan su trabajo.

La evaluación está en el centro de diversos intereses, necesidades, visiones contrapuestas y enfoques diversos. Los evaluadores deben determinar la perspectiva con la que abordarán una evaluación, reconocer la existencia de otras perspectivas, estar preparados para críticas, incluso de los patrocinadores, y estar atentos a mantener una fluida comunicación con las diversas partes que tienen interés en la evaluación.

Las evaluaciones son investigaciones llevadas a cabo para afectar el desarrollo de políticas públicas, para influir y dar forma al diseño e implementación de intervenciones sociales y para mejorar la administración de los programas sociales.

La evaluación es un ingrediente dentro del proceso político. La evaluación puede aportar información muy útil acerca de un programa en operación, pero eso no significa que automáticamente sus recomendaciones serán entendidas, aceptadas y utilizadas. En cualquier sistema político hay una cierta sensibilidad a ponderar, evaluar y balancear demandas que pueden ser, incluso, contradictorias. En este contexto, el rol del evaluador es el de un testigo experto que testifica acerca del grado de efectividad de un programa. Quienes toman la

decisión (autoridades, patrocinadores) pueden darle al testimonio del evaluador gran relevancia, pero son ellos, y no el evaluador, quienes finalmente decidirán. Los resultados de la evaluación es una entre varias consideraciones que se tomarán en cuenta para la decisión final.

Si no fuera de esta manera, el evaluador tendría poder de veto en el proceso de decisiones políticas, lo que significaría que son los evaluadores, y no las autoridades, los llamados a decidir.

2. EVALUACIÓN EX - ANTE

2.1. Concepto de Evaluación Ex – Ante

Las evaluaciones ex – ante son evaluaciones que se llevan a cabo antes del inicio de la implementación y puesta en práctica de una intervención. Este tipo de evaluaciones cumplen con dos finalidades:

- Poder anticipar dificultades que podrían aparecer en la ejecución del proyecto o programa y generar las correcciones necesarias con anterioridad a la puesta en práctica de la iniciativa,
- Poder establecer si la intervención generará un beneficio cierto (neto) en la realidad en que se implementará. Para ello es necesario identificar tanto los beneficios que reportará, como los costos en que se incurrirá. También permite identificar, entre varias, aquella iniciativa o alternativa que reporta los mayores beneficios netos, con lo que se transforma en una herramienta valiosa para la toma de decisiones sobre selección de programas y proyectos que pasarán a la etapa de implementación

Dos son los tipos de evaluación ex – ante que recurrentemente se ejecutan: evaluación de diseño y evaluación de eficiencia. Esta última, a su vez, puede implementarse a través de evaluaciones costo – beneficio y/o evaluaciones costo – efectividad.

2.2. Evaluación de Diseño

La evaluación de diseño es una valiosa herramienta, crecientemente usada por gobiernos y organismos internacionales, para estimar la utilidad de iniciativas a implementar. Por ejemplo, el FMI ha usado esta herramienta para conceder asistencia financiera a diversos países en sus procesos de superación de la pobreza y mejoramiento de las condiciones sociales de la población (ver FMI, 2006).

La evaluación del diseño, cuando es ex - ante, tiene como propósito estimar, a priori, si un proyecto o programa podrá abordar el problema social al que está dirigido con una razonable probabilidad de éxito. Ello implica analizar la consistencia entre:

- La identificación del problema, que surge de los diagnósticos que se efectúan acerca de la realidad social que se intervendrá, los objetivos de la intervención, el marco conceptual de la intervención y las hipótesis de intervención.
- Los objetivos establecidos al proyecto o programa, las actividades consideradas y la organización de esas actividades.
- Las actividades consideradas, los tiempos establecidos y los recursos asignados.

Las preguntas que normalmente aborda la evaluación de diseño incluyen a las siguientes: ¿están objetivos y metas del proyecto o programa

articulados?, ¿están expresados con claridad?, ¿están articuladas las actividades consideradas en la intervención?, ¿está el equipo de trabajo que implementará la intervención en conocimientos de las actividades que se implementarán y lo que se espera de ellos?, ¿hay algún plan de contingencia?, ¿está considerado desarrollar un monitoreo o seguimiento de la implementación de la iniciativa?

Cuando la evaluación del diseño se realiza con posterioridad a la implementación de la intervención tiene como propósito identificar cuales fueron las fortalezas o debilidades en el diseño de la intervención que llevaron a los resultados obtenidos. En este caso, la finalidad de una evaluación de diseño ex – post es identificar las lecciones aprendidas, para efectuar las correcciones en el mismo programa – si aún es posible – o para tenerlas en consideración para el diseño de futuras intervenciones.

Las evaluaciones de diseño podrán hacerse de manera interna, externa o ambas modalidades conjuntamente, con las ventajas y desventajas que cada una de estas opciones tiene y que han sido señaladas en la sección precedente de este documento.

2.3. Evaluación de Eficiencia de los Programas

El análisis costo – beneficio y costo – efectividad provee importante información a los gerentes de los programas, autoridades y partes interesadas sobre la relación de costos y beneficios alcanzados por una intervención.

Tanto el análisis costo-beneficio como el de costo – efectividad son medios para juzgar la eficiencia de un programa. A su vez, la evaluación de eficiencia provee un marco de referencia para relacionar los costos a los resultados alcanzados por un programa. Al proporcionar un elemento de juicio con el cual se indica la aceptabilidad de programas, es posible estar en condiciones de relacionarlo con otros para establecer prioridades y jerarquías (Cohen y Franco, 2000).

El análisis costo – beneficio y costo – efectividad tiene su equivalente en el área de la administración de negocios, donde costos son comparados con ingresos. Asimismo, el análisis de eficiencia es normalmente efectuado en el área de negocios privados. En el ámbito de los programas sociales los análisis de eficiencia han ganado creciente aceptación. Aún así, los métodos empleados para valorar costos y beneficios no han dejado de tener controversia, tanto desde un punto de vista de los tipos de beneficios y costos que deben ser incluidos, la certeza de las estimaciones producidas, las técnicas empleadas para cuantificarlos y dimensiones éticas que surgen en aquellas intervenciones que hacen referencia a valoraciones sobre la vida humana, por ejemplo, ¿cuánto vale la vida humana?.

La diferencia entre el análisis costo – beneficio y el de costo – efectividad está en la manera en que los resultados del análisis son expresados. En el análisis costo – beneficio los resultados del programa son expresados en términos monetarios. En el análisis costo – efectividad los resultados se expresan en términos sustantivos: cuánto producto por unidad de costo. Mientras en el análisis costo – beneficio reduciremos a dinero costos y beneficios, en el análisis costo – efectividad establecemos

una relación entre los productos expresados en unidades sustantivas con los costos monetarios (Cohen y Franco, 2000: 199).

El análisis costo - efectividad permite comparar el grado de eficiencia relativa que tienen proyectos alternativos para obtener los mismos productos. Así, entre distintas opciones podemos identificar cuál es la que maximiza los resultados con los mismos costos o cuál es la que minimiza costos para similares resultados.

En el análisis costo - beneficio identificamos el beneficio neto (o costo neto) que distintas alternativas presentan. Este beneficio o costo neto es la diferencia entre beneficios y costos valorados en términos monetarios. De este modo el análisis costo beneficio aporta un antecedente racional para jerarquizar programas según el beneficio neto que presenten.

2.4. Conceptos básicos

- *Beneficios:* Son los resultados o productos del programa. Los beneficios pueden incluir los efectos directos o indirectos. En el análisis costo beneficio son expresados en términos monetarios. En el análisis costo efectividad son expresados en términos sustantivos.
- *Costos:* son los recursos, directos o indirectos, requeridos para producir una intervención. Normalmente son expresados en términos monetarios.
- *Beneficios Netos:* El total de los beneficios menos el total de costos, ambos valores actualizados (incluyen una tasa de descuento).

- *Tasa de beneficios a costo*: Beneficios totales actualizados, divididos por costos totales actualizados.
- *Beneficios y Costos Directos*: Se refieren a la valoración de productos – bienes y servicios – entregados a los ciudadanos y a los insumos requeridos para generar estos bienes y servicios.
- *Beneficios y Costos Indirectos o secundarios*: Serie de consecuencias y efectos – negativos o positivos – generados por un programa fuera del ámbito de implementación o que afectan a terceros.
- *Precio de cuenta, social o sombra*: Son aquellos valores que son corregidos de las imperfecciones de mercado – información incompleta, control de precios y cambios, monopolios, monopsonios – a una situación de competencia perfecta, mediante la cual se refleja la verdadera escasez del resultado del programa y otorga una medida real de evaluación.

2.5. El Análisis Costo – Beneficio

El análisis costo – beneficio requiere identificar la totalidad de beneficios que reporta o reportará el proyecto o programa, sean éstos directos o indirectos, tangibles o intangibles. Del mismo modo, debe hacerse una identificación completa de los costos en que incurre la intervención. A su vez, costos y beneficios deben traducirse en una unidad de medida común, que es normalmente el dinero.

a) Perspectivas en costos y beneficios

El modo de estimar los costos y beneficios no está exento de controversia, sobre todo en programas sociales. El primer tema que debe resolverse es qué perspectiva tomar para el análisis. ¿Desde qué perspectiva estimamos costos y beneficios?, es decir ¿costos y beneficios para quién?.

Este punto es crítico, pues el análisis debe tener una sola perspectiva. Mezclar puntos de vista puede llevar a contabilizar dos o más veces costos y beneficios.

Los análisis costo – beneficio son efectuados normalmente desde tres tipos de puntos de vista.

- Beneficiarios individuales o agregados (población objetivo)
- Patrocinadores del programa
- Comunidad

Los análisis efectuados desde la perspectiva del beneficiario estiman costos y beneficios que asumen las unidades que reciben la intervención. Según Rossi y Freeman (1993), este enfoque normalmente produce estimaciones de beneficios más altas. La razón sería que si el patrocinador del programa (el Estado) o la sociedad asumen los costos (o buena cantidad de ellos) o subsidios, entonces para los beneficiarios del programa hay mayores beneficios netos.

Los análisis basados en la perspectiva del patrocinador toman la perspectiva de quién financia la intervención. Desde este punto de vista, los costos en que incurre este patrocinador serán costos relevantes para el programa. Esta perspectiva se usa normalmente cuando hay programas alternativos que pueden ser implementados y sólo hay disponibilidades para uno de ellos.

Los análisis basados en la perspectiva de la comunidad toman el punto de vista de la sociedad o grupo de referencia como un todo. Esta es la perspectiva más comprehensiva, pero también es la más difícil de aplicar. Tomar esta perspectiva implica hacer esfuerzos adicionales por incorporar los efectos indirectos y externalidades producidos por el programa.

b) La medición de costos y beneficios

La medición de costos y beneficios enfrenta tres desafíos, a saber:

- Identificar y medir todos los costos y beneficios de un programa.
- Expresar todos los costos y beneficios en un denominador común, normalmente unidades monetarias. En el análisis costo – efectividad los beneficios se expresan en unidades sustantivas.
- Expresar costos y beneficios en unidades de medida – normalmente monetarias – equivalentes. Ello implica considerar el valor tiempo del dinero y expresarlos – normalmente – en valor presente, lo que

implica aplicar una tasa de descuento para traer beneficios y costos, desde el pasado o futuro, a valor presente.

El primer desafío está relacionado con la disponibilidad de fuentes de información que permitan al analista apreciar tipos de costos y tipos de beneficios. En los análisis *ex-ante* puede haber cierta identificación previa, una idea de cuáles serían, pero eso puede ser controversial. Aquí la controversia se centrará en discutir porque se identificaron los costos y beneficios estimados en el estudio y no otros. En los análisis *ex-post* la dificultad radica en poder identificar la totalidad de costos que se incurrió y beneficios que se obtuvieron en los hechos.

Respecto del segundo desafío, la cuestión es cómo medir y en qué unidad de medida especifican costos y beneficios del programa.

Respecto al tercer desafío la cuestión es como hacer comparables los costos y beneficios, especialmente cuando ellos surgen en momentos distintos. La identificación de una tasa de descuento permite llevar – o traer – valores que surgen en tiempos distintos a un mismo momento y hacerlos comparables².

En este punto surgen diversas controversias, sobre todo cuando es necesario estimar costos y beneficios en unidades monetarias. Por ejemplo, ¿cuál es el valor monetario de una vida?, ¿cuál es el valor monetario de los daños psicológicos producidos por una agresión? En análisis de programas

² Para mayor detalle acerca de la comparación temporal de flujos de un proyecto y la tasa de descuento ver Jansson (1994) y Budnick (1993).

médicos o evaluaciones orientadas hacia intervenciones en seguridad ciudadana estas son preguntas recurrentes y que no tienen fácil respuesta.

Diversos métodos existen para valorar costos y beneficios. Entre los más usados:

- Contabilización
 - Datos del presupuesto, gasto
 - Costo de oportunidad

- Valorización contingente

- Precios hedónicos

- Precios sombra

2.6. El Análisis Costo - Efectividad

El análisis costo efectividad es utilizado cuando los beneficios de programas sociales no son susceptibles de ser transformados a unidades monetarias. Rossi, Freeman y Wright (1979; citados en Cohen y Franco, 2000) definen al análisis costo beneficio como una "técnica analítica que compara los costos de un programa con los beneficios resultantes, no expresados en la misma medida". Así, los costos son llevados a medidas monetarias y relacionados con beneficios resultantes. Por ejemplo, en un programa de rehabilitación de jóvenes delincuentes, la medida podría ser número de rehabilitados por unidad de costo.

a) Utilidad del análisis costo – efectividad

El análisis costo – efectividad permite evaluar alternativas de decisión, comparando los resultados, medidos en unidades no monetarias, con los costos incurridos, medidos en unidades monetarias. Esta relación costo efectividad jerarquiza las opciones según la magnitud de recursos disponibles tomando en cuenta sólo resultados o fines últimos del programa.

Teniendo en cuenta lo anterior, la utilidad del análisis costo – efectividad está en que permite obtener información sobre si el programa se encuentra en el óptimo de resultados/productos con un monto limitado de recursos o si el programa minimiza el costo con resultados predeterminados. La consecuencia de aquello es que el análisis de costo efectividad puede jerarquizar alternativas de programas, lo cual constituye una herramienta para la toma de decisiones efectiva sobre las prioridades de quién asume dicho rol.

La forma en la cual se establecen las prioridades teniendo un conjunto de alternativas está en la comparación del grado de eficiencia relativa que tienen programas distintos para los mismos productos y la misma población objetivo. Así, entre un conjunto de alternativas puede descubrirse la óptima, minimizando los costos para cierto nivel de logros, o bien, maximizado el alcance de los fines para un nivel total predeterminado de costos.

Cohen y Franco (2000) afirman que la principal orientación de la evaluación costo – efectividad es prospectiva, es decir, que se evalúa las

capacidades del programa antes de su implementación. Por tanto, el problema de la evaluación en el análisis costo efectividad se reduce a encontrar la opción más eficiente para obtener los productos –bienes y servicios- que el programa va a generar.

Hemos dicho que para realizar el análisis costo – efectividad se deben comparar los resultados en términos sustantivos por unidad de costo. Pero ¿qué debe incluir la estimación de costos? Se establece que deben incluirse todos los costos que el programa demanda en su vida útil. Para determinarlos es necesario identificar, primero, los costos de inversión – construcción de inmobiliario, compra de bienes raíces, entre otros – prorrateados en los años de duración del programa con el fin de contabilizarlos anualmente; segundo, los costos operacionales derivados del funcionamiento a lo largo del desarrollo del programa, como también otros costos menores y, tercero, actualizarlos o ponerlos en el mismo momento del tiempo, para los cual se aplicará una tasa de descuento. También es necesario determinar la demanda que va a satisfacer el programa, quiénes y cuántos son los beneficiarios y qué beneficios se utilizarán para satisfacer dicha demanda con el objeto de establecer el valor sustantivo que se relacionará con la unidad de costo. En ese sentido, sólo si se determinan los costos y se relacionan con la unidad de producto resultante es posible escoger la alternativa óptima. Así, el análisis costo – efectividad permite realizar una elección racional.

b) Etapas específicas del análisis costo – efectividad en la evaluación *ex ante*.

De acuerdo a Cohen y Franco (2000) existen dos opciones de evaluar un programa de forma prospectiva. Una es la metodología de llevar los valores futuros de rentabilidad socioeconómica a valores presentes llamada Valor Actual Neto (VAN) y Tasa Interna de Retorno (TIR). Sin embargo, esta metodología se encuentra con la dificultad de transformar los beneficios a unidades monetarias en proyectos o programas sociales. La segunda metodología es la del análisis de costo efectividad cuyas etapas son, a saber:

- Definición del problema: consiste en la comparación entre la situación vigente y el estándar políticamente deseable. Esta etapa responde a qué necesidades y qué servicios va a otorgar el programa; quiénes integran la población objetivo; dónde están localizados; cuántos recursos se disponen para satisfacer sus necesidades y cómo serán provistos dichos recursos.
- El diagnóstico de la situación: consiste en describir la oferta del servicio y la demanda que va a satisfacer, además de indicar la cobertura del programa; explicar las causas del problema que se va a intervenir; y predecir la proyección de la situación con programa y sin programa.
- Planteamientos de alternativas de solución: consiste en la profundización de aquellas alternativas consideradas en las etapas anteriores para analizar caminos viables, implementación y recursos

necesarios con el objeto de alcanzar el nivel de atención establecido en cada una de ellas. La relación costo eficiencia se establece conociendo los costos de puesta en marcha e implementación de cada opción del programa, lo cual permite escoger la que tiene el mínimo costo.

- La selección de la mejor alternativa es lograda cuando se comparan todas aquellas que son planteadas. Si el resultado de todas las opciones es homogéneo, esto es, si prestan el mismo servicio con igual calidad, se escoge el que represente el mínimo costo. Cuando se relacionan costos con producto se obtiene el costo unitario con lo cual se realiza el análisis costo – eficiencia; pero cuando se incluye los resultados que genera la intervención, pasamos al análisis costo – efectividad que permite seleccionar la mejor opción que minimiza los costos por unidad de impacto.
- En la etapa de la elaboración del proyecto definitivo, Cohen y Franco (2000) establecen que ésta se realiza cuando el programa social implica la construcción de obra física. Si no fuera tal el caso, debe obviarse este paso y pasar directo a la implementación.
- La incertidumbre y el análisis de sensibilidad están asociadas primero al análisis del riesgo – de fracaso – del programa, y segundo a todas aquellas variables críticas para el éxito – que ejercen un efecto significativo en la aceptabilidad del programa – que pueden variar durante el desarrollo de la implementación.

En el análisis de la evaluación *ex ante* es necesario mirar de una forma prospectiva las alternativas que tiene el programa social, es decir, establecer una proyección hacia el futuro, con el objeto de establecer *a priori* los efectos sobre la población objetivo.

c) Análisis costo – efectividad en la evaluación *ex post*.

Según Cohen y Franco (2000), este tipo de evaluación está más desarrollada metodológicamente y es la que ha tenido mayor aplicación. De acuerdo a estos autores, ella puede ser de procesos o de impacto. Ambas se distinguen en su finalidad. La primera tiene como objetivo mejorar la eficiencia operacional, mientras que la segunda su fin radica en determinar los cambios que son generados en la población objetivo después de la implementación de un programa.

Así como en la evaluación *ex ante*, este tipo de evaluación en el análisis costo efectividad consta de diversas etapas, a saber:

- Identificar con precisión los objetivos del programa. Aquí se consideran sólo los objetivos que son traducidos en actividades, es decir, que implican la asignación de recursos y que reflejan el verdadero fin del programa.
- Traducir los objetivos en dimensiones operacionales e indicadores. Es útil transformar los objetivos cualitativos en metas para poder cuantificarlos y así construir indicadores.

- Especificar las alternativas que serán evaluadas y medir los recursos afectados en valores monetarios. Las alternativas surgen de la ejecución del programa o de la comparación con un modelo de optimización que supere las restricciones identificadas en la evaluación. Es necesario que las alternativas sean medidas homogeneizando las unidades de producto para que sean comparables.
- Valorar comparativamente las alternativas. En esta etapa deben identificarse tanto los efectos buscados y previstos como aquellos no buscados que generalmente son costos o beneficios secundarios o indirectos, no susceptibles de ser medidos. Con esto se podrá determinar el saldo neto de un proyecto (Cohen y Franco, 2000).

Ahora bien, las dificultades del análisis costo efectividad dicen relación con que no pueden hacerse generalizaciones de los resultados y no permite comparar proyectos con objetivos distintos, ya que no contienen el denominador común esencial para hacerlos comparables. Entonces, la elección entre dos programas con distintos resultados u objetivos esperados debe hacerse priorizando la satisfacción de necesidades. En ello, la satisfacción de necesidades básicas se ubica en primer lugar.

3. EVALUACIÓN INTERMEDIA

3.1. La Evaluación Intermedia

La evaluación intermedia ocurre durante la ejecución del programa o proyecto. Su propósito es poder identificar si hay circunstancias que estén impidiendo que la intervención marche como se preveía o cumpla las metas que se habían presupuestado para los distintos momentos de la ejecución.

Puede ser que hayan habido cambios en la realidad que no estaban previstos, lo que hace que el diseño de la intervención pueda haber quedado desfasado del problema que aborda y, en ese caso, la evaluación intermedia podrá identificar esos cambios y generar la información necesaria a los tomadores de decisiones para las correcciones que haya lugar. Puede ocurrir también que los usuarios del programa se comporten de manera distinta a como lo preveía el diseño de la intervención, ya sean porque estén haciendo un sobre consumo, sub-consumo del servicio o utilizando el servicio para otras finalidades, distintas de aquellas por las cuales se estableció. Por ejemplo, el 1.000.000 de llamadas "bromistas" que recibe el número 133, de emergencias policiales, muestra que los usuarios no se comportan con la racionalidad que esperaban los diseñadores del servicio, a la vez que generan una seria dificultad para la efectividad de este servicio, por cuanto, la sobre carga de llamadas hace que quienes están en una real emergencia policial tengan serias dificultades para recibir la protección policial, si es que la

reciben. Adicionalmente, puede ocurrir que la ejecución no respete el diseño que se tenía del programa y esa ser la causa de los incumplimientos de metas que se observan.

En cualquiera de estos casos, la evaluación intermedia podrá identificar la causa de los problemas que está afectando a la implementación del programa, informar a los responsables de la gestión, para que éstos implementen las acciones correctivas. Éstas pueden apuntar a correcciones del diseño o correcciones en la gestión del programa.

Así, la evaluación intermedia se conduce a través de estudios de monitoreo o seguimiento y es una actividad esencial en la gestión misma del programa. El monitoreo puede ser visto como una actividad orientada a ajustar la implementación del programa con su diseño. Cuando las evaluaciones de resultados e impacto revelan que ellos no se lograron, se lograron sólo parcialmente o que hubo escaso o ningún impacto es necesario averiguar por qué. Sin un adecuado monitoreo o seguimiento es imposible saber si la intervención se está encaminando a los resultados que se esperan.

La literatura sobre evaluación es marcadamente prolífica en terminología. Es común encontrar referencias a tipos de evaluaciones que en otros textos son cubiertos con otras denominaciones. Uno de esos casos es la llamada evaluación de procesos. Cohen y Franco (2000: 110) señalan que ella “determina la medida en que los componentes de un proyecto contribuyen o son disfuncionales a los fines perseguidos, (que) se realiza durante la implementación y, por tanto, afecta a la organización y las

operaciones, (que) busca detectar las dificultades que se dan en la programación, administración, control, para corregirlas oportunamente, disminuyendo los costos derivados de la ineficiencia, y (que) no es un balance final, sino una evaluación periódica". Así definida, la primera parte de la definición de Cohen y Franco se aproxima bastante al concepto entregado en la sección anterior sobre evaluación de diseño. El resto de la definición que nos ofrecen los autores citados se aproxima al concepto de monitoreo – cuestión central en esta unidad. Éste, como se verá en esta unidad, concentra su análisis en la cobertura del programa (si llega a la población especificada), en la entrega o distribución de los bienes y servicios y avances logrados en la ejecución. El monitoreo, por tanto, centra su análisis en la programación, administración y control de las actividades establecidas para entregar el bien o servicio a la población especificada en el diseño. Así, en este documento entenderemos que el concepto de monitoreo también cubre a aquel tipo de análisis que en otra literatura es señalada como evaluación de procesos. El análisis de la cobertura y la distribución es, en último término, el análisis de los procesos que permiten distribuir el servicio (o bien) y llegar con él a la población especificada en el diseño. Este tipo de monitoreo es presentado en la sección 3.5 de esta Unidad.

Asimismo, la evaluación intermedia supone la existencia de un plan, que establezca metas e indicadores con el cual comparar la ejecución. En este contexto, una herramienta útil para los procesos de planificación y evaluación que ha alcanzado gran difusión es el llamado "marco lógico".

El método del marco lógico fue elaborado inicialmente por la Agencia de Cooperación de Estados Unidos (USAID) hacia fines de la

década de 1960. Buscaba incrementar la efectividad de la ejecución de proyectos por la vía de definir con precisión los objetivos, establecer con claridad la responsabilidad de los ejecutores y tener una base cierta y nítida con la que comparar la ejecución y sus resultados y, a partir de eso, identificar y explicar las desviaciones. Las ventajas que ofrece la metodología del marco lógico son que aporta una terminología uniforme que sirve para eliminar ambigüedades, lleva a generar una estructura lógica de las actividades y permite establecer criterios de verificación del cumplimiento de metas y logros.

3.2. Seguimiento y Monitoreo

El monitoreo ha sido siempre una actividad esencial de evaluación. Cuando las evaluaciones de resultados revelan que hubo escaso o ningún impacto es necesario estimar por qué. Sin un adecuado monitoreo es imposible determinar si la intervención ha sido eficaz. El monitoreo puede ser visto como una actividad orientada a ajustar la implementación del programa con su diseño.

El monitoreo de programas con propósitos de evaluación es la evaluación sistemática de la cobertura e implementación de los programas.

a) Conceptos

Valadez y Bamberger (1994) definen al monitoreo como una actividad gerencial interna y continua cuyo propósito es asegurar que el programa alcance sus objetivos, en el tiempo previamente establecido y

conforme al presupuesto asignado. De este modo, el monitoreo involucra la provisión de información de retroalimentación sobre los avances y problemas detectados durante la implementación. Las actividades en las que se centra el monitoreo, para estos autores, son el análisis de la obtención y asignación de recursos, producción y entrega del servicio, y el registro de los costos.

Cohen y Franco (2000) hacen sinónimo evaluación y seguimiento. Así, el monitoreo o seguimiento sería el examen continuo o periódico que efectúa la administración, en todos sus niveles jerárquicos, acerca de la manera en que se está ejecutando una actividad. Agregan – basados en los manuales de la ONU - que en ello se busca asegurar que la entrega de insumos, los calendarios de trabajo, los productos esperados se conformen a metas establecidas y que otras acciones que son necesarias progresen de acuerdo con el plan trazado.

Para Rossi y Freeman (1993) el monitoreo es el examen sistemático de la cobertura del programa y su implementación, distribución y entrega. La evaluación de la cobertura consiste en estimar en qué medida el programa llega a su población objetivo. El análisis de la implementación consiste en la determinación de la congruencia entre lo establecido en el plan respecto del otorgamiento de los servicios del programa y la que realmente se está entregando.

El análisis de estas conceptualizaciones muestra que Valadez y Bamberger y Cohen y Franco centran su atención en las cuestiones más operativas del monitoreo mientras que Rossi y Freeman se centran, más bien, en lo sustantivo del programa.

¿Están ambas visiones aludiendo a instrumentos distintos? No. Mientras los primeros puntualizan el tipo de actividades que se llevan a cabo en el monitoreo, Rossi y Freeman se centran en los resultados de la aplicación del programa y, desde esa perspectiva, analizan las desviaciones de la ejecución con respecto al plan. Según Rossi y Freeman el monitoreo provee la información necesaria que permite responder la pregunta de ¿por qué un determinado programa no alcanzó el impacto esperado?

b) Perspectivas en el Monitoreo

Aunque el tipo de información requerida y los procedimientos de recolección de datos usados son generalmente los mismos, o se superponen considerablemente, para Rossi y Freeman (1993) el monitoreo tiene diferentes propósitos cuando es implementado desde la perspectiva de una evaluación, de la efectividad y responsabilidad pública (accountability), y de la gestión del programa.

- *Desde la perspectiva del evaluador*

Muy a menudo el impacto de un programa es bajo o inexistente, ya sea porque el tratamiento apropiado no es aplicado o no se aplica completamente, porque la intervención no es aplicada a la población objetivo señalada en el diseño, porque hay variaciones sustantivas en el modo en que se aplica la intervención en diversos ámbitos, localidades o grupos poblacionales, o porque pueden darse todas estas situaciones problemáticas conjuntamente.

Los estudios de monitoreo son esenciales para entender e interpretar los resultados de la evaluación de impacto. Saber que ocurrió es necesario para explicar o hipotetizar por qué un programa funcionó o no funcionó (logró o no sus resultados e impacto). Para el evaluador el monitoreo es esencial para comprender por qué un programa no alcanzó sus objetivos o impacto deseado. De esta manera, el monitoreo entregará información valiosa y complementaria a la evaluación de impacto.

Los evaluadores necesitan información sobre la medida o formas en que los elementos individuales de cada programa son implementados para establecer la utilidad de aspectos particulares de la intervención o de la frecuencia y duración del tratamiento requerido.

Adicionalmente, el monitoreo provee información necesaria para la difusión del programa. Las características esenciales de una intervención pueden ser reproducidas sólo si el programa puede ser descrito en sus detalles operacionales.

- *Desde una perspectiva del Control, Efectividad y Responsabilidad Pública*

La información que provee el monitoreo es también crítica para los formuladores del programa y, también, para las autoridades que lo han respaldado políticamente y financiado, para los políticos que lo han apoyado y para los grupos que tienen intereses en él (a favor o en contra).

Aunque parte de esta información es rutinariamente provista por los gerentes o administradores de los programas, información de monitoreo provista por fuentes independientes es altamente valorada y necesaria para dar juicios no comprometidos con intereses particulares. Por eso, muy a menudo las autoridades, grupos de interés, patrocinadores y financistas de los programas perciben a los evaluadores que hacen monitoreo como sus “ojos y oídos” y como una segunda línea de información independiente. Así se podrá saber si el programa debe mantenerse, expandirse, reducirse o discontinuarse.

Monitoreo es diferente de la auditoría. Mientras ésta es una actividad normativa que compara “lo que es” con “lo que debería ser”; el monitoreo es el análisis de lo que ha ocurrido realmente en los hechos y por qué está pasando. Así, el monitoreo se centra en la identificación empírica y medición de lo que realmente ha ocurrido o está ocurriendo.

- *Desde la perspectiva del Administrador*

El monitoreo con una orientación gerencial a menudo plantea los mismos temas que los estudios de evaluación de resultados, impacto o de efectividad; la diferencia radica en el propósito para el cual la información que arroja el monitoreo es utilizada. El monitoreo con orientación gerencial está menos orientado a hacer juicios y más inclinado a incorporar medidas correctivas como parte regular de las operaciones del programa. Utilizando la misma información, el monitoreo se centra en cuáles deben ser las medidas correctivas que deben introducirse durante la operación del programa.

Consecuentemente, ésta debe ser una actividad regular de la Gerencia de Programas Públicos. El monitoreo desde la perspectiva gerencial es particularmente importante durante la etapa piloto y la fase de implementación de nuevos programas.

c) Objeto del monitoreo

El monitoreo busca responder tres preguntas claves:

- En qué medida un determinado programa está llegando a su población objetivo,
- Si la implementación de los servicios es consistente con las especificaciones establecidas en el diseño del programa, y
- Si las acciones implementadas y/o los recursos consumidos por el programa son los previstos.

3.3. Monitoreo de la cobertura

¿Está llegando el programa a la población objetivo?, ¿qué cobertura ha alcanzado el programa?, ¿hay sesgos en la incorporación de beneficiarios del programa? Estas son preguntas que abordan uno de los objetivos del monitoreo: el análisis de la cobertura.

Es, por tanto, esencial definir la población objetivo a la que se orientará el programa. Ello es determinante para las etapas de

planificación, diseño e implementación de los programas y, por cierto, para el monitoreo.

a) Focalización y Cobertura

Cualquier intervención social puede encontrar dificultades para atraer a su población objetivo. Por ello es necesario verificar si el programa está llegando a la población identificada como beneficiaria en el programa.

La idea de verificar la efectividad de la cobertura del programa se asocia con el concepto de focalización. Según Margaret Grosh (1994), la focalización se refiere específicamente a la identificación de aquellas personas que serán o no serán beneficiarias de un programa social. La práctica de la focalización se asocia con la idea que algunas personas deben ser excluidas de recibir los beneficios de un determinado programa social. La focalización supone superar el esquema homogéneo, propio de los programas universales, donde los servicios son ofrecidos a toda la población sin considerar sus diferencias y peculiaridades.

El objetivo de la Focalización es concentrar recursos en quienes más lo necesitan. La idea básica es concentrar los recursos de un determinado programa en un cierto grupo de personas que necesitan recibir los beneficios de ese programa, con la finalidad de aumentar el impacto (resultados esperados) del programa. Luego de identificar el grupo al que está dirigido el programa se produce la aplicación del programa, lo que es conceptualmente distinto.

Por ello, es necesario verificar si el programa ha logrado la cobertura esperada y si hay posibles sesgos en la cobertura del programa.

- Cobertura: es el grado (tasa o proporción) en que la población objetivo es usuaria del programa.
- Sesgo: ocurre cuando un grupo es cubierto más densamente que otro. Por ejemplo, más mujeres que hombres que reciban los beneficios de un programa indicaría que hay un sesgo a favor de las mujeres. Los sesgos pueden darse también por autoselección, es decir, voluntariamente unos participan o usan más que otros los servicios (o bienes) que entrega el programa.

Por ejemplo, la Junta de Vecinos América, con fondos aportados por el Fondo de Prevención del Consumo de Drogas, implementó el proyecto "Prevención la Mejor Inversión frente a la Drogadicción". El proyecto puso en práctica talleres de desarrollo personal, de prevención de drogas, deportivos y recreativos para jóvenes. Se esperaba que asistieran, al menos, 30 jóvenes del barrio que presentaban diversos grados de consumo de drogas. En realidad asistieron, en promedio, 17 jóvenes, la mayoría de ellos hijos de socios de la organización. Aunque fueron invitados, jóvenes que se juntan en las esquinas – y que muestran involucramiento en drogas – no participaron de los talleres.

Aunque los dirigentes perciben que el proyecto ha sido positivo para la comunidad, porque se cumplieron las actividades programadas y las cuentas están en orden, el monitoreo de la cobertura nos podría dar luces sobre aquellos aspectos en los que debería haber mayor atención.

Quienes participaron en el taller fueron los hijos de los dirigentes, es decir, aquellas personas más motivadas por prevenir el consumo de drogas en la población joven, por lo que, presumiblemente, sus hijos presentan un bajo o inexistente consumo de drogas. Aunque los talleres puedan haber afianzado la conciencia acerca del daño que causa el consumo de drogas en este grupo, ellos – dada la formación del hogar otorgada por sus padres – ya habían desarrollado esa conciencia. Sin embargo, el otro grupo, el que presenta el problema, estuvo ausente de los talleres. Dado que el grupo consumidor de drogas no recibió el tratamiento, es probable que su consumo de drogas no se altere. Así, los beneficios, resultados e impactos en la comunidad agregada probablemente serán bajos.

El ejemplo permite mostrar cómo problemas en la cobertura pueden hacer que el programa o proyecto no logre las finalidades para las que fue diseñado.

b) Problemas de cobertura

Otras dos situaciones que pueden ocurrir es que el programa presente sub o sobre cobertura.

- La *sub-cobertura* se da cuando el programa no logra alcanzar la cobertura prevista. Esto puede ocurrir por sesgos en la manera en que los beneficiarios son seleccionados o porque una parte de la población objetivo no se interesa en los servicios (o bienes) que provee el programa. La sub-cobertura es medida por la proporción de la población objetivo de un programa que realmente es usuaria de él.

La proporción de la población objetivo que usa el programa se expresa del siguiente modo:

$$POP = 100 * \left(\frac{PUN}{POB} \right) ; \text{ donde}$$

POP : Proporción de Población Objetivo que Usa el Programa

PUN : Población Usuaría que Necesita el Programa

POB : Población Objetivo (del Programa)

La proporción de la población objetivo que no usa el programa se expresará como:

$$PONU = 100 - POP ; \text{ donde}$$

PONU : Proporción de Población Objetivo No Usuaría

POP : Proporción de Población Objetivo que Usa el Programa

- La *sobre-cobertura* ocurre cuando hay personas que son usuarias del programa, pero que no cumplen con la definición del perfil de

población objetivo. En este caso, el costo del programa es superior a lo que se sería si se concentra en la población objetivo.

La sobre-cobertura se expresa como el número de participantes usuarios del programa, pero que no cumplen con el perfil de la población objetivo, comparado con el número total de participantes en el programa (Rossi y Freeman 1993). Se expresa de la siguiente forma:

$$PUNO = 100 * \left(\frac{PFP}{PUP} \right) ; \text{ donde}$$

PUNO : Proporción de Población Usuaría que no corresponde a Población Objetivo

PFP : Población (usuaría) Fuera del Perfil de la Población Objetivo

PUP : Población Usuaría del Programa

b) Eficiencia de la cobertura

En general, lo que se busca es hacer un uso eficiente de los recursos. Ello significa maximizar el número de beneficiarios que realmente necesitan el servicio (o bien) que entrega el programa y minimizar el número de beneficiarios que podrían no necesitarlo u obtenerlo por sus propios medios.

La eficiencia en la cobertura se expresa como:

$$\text{Eficiencia de Cobertura} = 100 * \left\{ \frac{\text{Número de usuarios}}{\text{Nº total población objetivo}} - \frac{\text{Número de usuarios no necesitados}}{\text{Nº total usuarios}} \right\}$$

Valores que van de -100 a +100 indican el grado de eficiencia de la cobertura. El valor +100 indicará que el programa cubre a la totalidad de la población que necesita el servicio y a nadie que no lo necesite, o lo pueda obtener por sus propios medios. EL valor -100 indica que sólo personas fuera del perfil de la población objetivo son usuarias del programa.

Las fuentes de información para desarrollar el monitoreo de la cobertura pueden ser: archivos del servicio, encuestas a los beneficiarios del programa y encuestas a la comunidad.

c) Sesgos de la cobertura

El análisis de los sesgos en la cobertura del programa puede hacerse a través del examen de las diferencias existentes entre los beneficiarios del programa, de aquellos que lo abandonan y de aquellos que siendo elegibles no se inscriben. Estas dos últimas situaciones pueden ser signos de insatisfacción con el programa. También pueden estar indicando la

existencia de barreras a la incorporación al programa de personas que cumplen con el perfil de la población objetivo.

El análisis de los sesgos de la cobertura será esencial para estimar la efectividad de los esfuerzos de focalización y gestión del programa y, en caso de ser necesario, reorientarlo hacia quienes cumplen el perfil de la población objetivo, de modo de mejorar los resultados e impacto del programa.

3.4. Monitoreo de la Implementación del Programa

El monitoreo de la implementación busca establecer si en la ejecución real del programa han habido desviaciones respecto de las especificaciones establecidas en el diseño que estén teniendo incidencia en los resultados y/o impactos observados. Este tipo de monitoreo es importante para tomar decisiones acerca de si es adecuado mantener, expandir, reducir o terminar un cierto programa.

La información que provee el monitoreo de la implementación permite a los gerentes de los programas poder identificar problemas que ya se estén presentando o situaciones que, de no atenderse oportunamente, pueden devenir en serios problemas que comprometan la viabilidad del programa. Los informes que provee este tipo de monitoreo pueden permitir identificar problemas como retrasos en la entrega de los beneficios del programa, gastos por sobre lo estimado, insatisfacción de los usuarios, entre otros (Valadez y Bamberger 1994).

Según Rossi y Freeman (1993) hay tres tipos de problemas que se presentan en la implementación y que inciden en que un determinado programa pueda no estar alcanzando los resultados esperados o produciendo el impacto deseado.

1) Otorgamiento Incompleto de los Beneficios

Ocurre cuando, en el hecho, los beneficios no llegan a la población objetivo o el modo de implementación del programa diluye de tal modo los beneficios que sólo una pequeña parte de ellos llega a los beneficiarios. Puede ocurrir que la falta de entendimiento o compromiso de los funcionarios operativos haga que los usuarios reciban servicios o bienes por debajo del nivel señalado en el diseño del programa. Por ejemplo, es frecuente escuchar reclamos respecto del trato de los funcionarios a quienes acuden a solicitar atención de salud en los consultorios o de personas pobres que acuden a solicitar beneficios en las Municipalidades. Dado ello, un buen número de potenciales beneficiarios se abstiene de solicitar nuevamente los beneficios. Otros, aparentemente, estarían recibiendo los beneficios por debajo de lo definido. Estos podrían ser casos de "tratamiento incompleto".

2) Implementación Errónea

Puede ocurrir de dos maneras. En el primer caso, puede ocurrir que las actividades emprendidas no tengan relación con el diseño del programa y que, por tanto, los beneficios que el programa establecía para la población objetivo identificada no estén llegando a esta. En este caso, el modo en que se implementa el programa, en el hecho, lo anula. La otra

manera en que se puede dar este problema es que se requieran procedimientos de implementación tan sofisticados que el personal de que se dispone no es capaz de implementarlo y, finalmente, ejecutan el programa según como ellos lo entendieron. Es decir, puede ocurrir que los resultados de la implementación de la fase piloto fueron adecuados, pero que se hicieron con un personal altamente capacitado y motivado, y que no será aquel que trabaje en la etapa de implementación normal del programa.

3) Implementación No Estandarizada

Ocurre cuando la implementación deja mucha discreción a los ejecutores, lo que puede resultar en que el modo como se ejecute el programa y los resultados que se obtengan varíen muy sensiblemente de un lugar a otro.

3.5. Monitoreo de los Avances del Programa

La sección 3.2 de esta Unidad ha señalado que una de las preguntas que busca abordar la práctica del Monitoreo es si las actividades emprendidas en nombre de la intervención y los recursos que se han consumido en su implementación corresponden a lo que se planeó. Por extensión, también busca estimar los progresos en la ejecución del programa o proyecto y si éste logra cumplir las metas previstas para las etapas de implementación transcurridas.

Este tipo de monitoreo también es denominado en variada literatura como “evaluación de procesos”, porque su finalidad es estimar si el

programa está siendo implementado conforme al plan y porque puede hacerse una o varias veces durante la vida del programa o proyecto (Frechtling, 2002). Otra literatura se refiere a esta modalidad como “monitoreo insumo – producto” o “monitoreo del rendimiento” (ver: Valadez y Bamberger 1994: 121).

Este tipo de monitoreo busca comparar el diseño del programa, que incluye su plan de implementación, con la ejecución real. Esta comparación se centra en establecer si las metas alcanzadas por la ejecución real corresponden a las previstas para la etapa de implementación en análisis; si las actividades desarrolladas son aquellas previstas en el diseño y si los recursos consumidos son los que se preveían.

Las fuentes de información para desarrollar un monitoreo de este tipo provienen de los documentos que describen el diseño del programa y su plan de implementación, los reportes de implementación del programa y los flujos de fondos (presupuestados y ejecutados). La comparación de esta información permitirá establecer si se están cumpliendo las metas previstas y, en caso de haber desviaciones, si éstas corresponden a problemas o aciertos en la implementación, en el diseño y/o en el uso de los recursos.

La información que provee este tipo de monitoreo permitirá a los administradores del programa efectuar las acciones correctivas necesarias, en las áreas señaladas, para poder lograr las metas asignadas al programa.

Por ejemplo, el proyecto "Construcción de Multicanchas en la Villa Los Copihues" establecía aportes monetarios del Ministerio del Interior y de mano de obra por parte de la comunidad, representada por la Junta de Vecinos. A cinco meses de iniciado el proyecto, los avances son escasos. El monitoreo de los avances del proyecto permitiría, en este caso, establecer si el cronograma del proyecto se ha cumplido, cual sería la causa de retraso de las obras y que acciones correctivas sería necesario implementar para cumplir las metas comprometidas.

3.6. Datos para el Monitoreo

Pueden recolectarse de las siguientes fuentes

- *Observación directa.* Permite apreciar como está implementándose efectivamente el programa, cual es su funcionamiento real. Este modo de recolección de datos es posible cuando no es intrusiva, no es distractiva para el personal o no crea efectos adversos en la gestión del programa.
- *Registros y archivos de la institución que implementa el programa.* Entre otros aspectos, el análisis de esta documentación permite apreciar las características de los usuarios reales y compararlos con la definición de población objetivo señalada en el diseño; las decisiones rutinarias que se han ido tomando en la ejecución del programa; la distribución de funciones, competencias y responsabilidades entre el personal del programa; el consumo de los recursos asignados y costo en que se ha ido incurriendo en cada una de las actividades.

- *Personal del Programa.* La información recolectada del personal del programa puede hacerse a través de entrevistas en profundidad, reportes acerca de las actividades que ellos han emprendido en la ejecución del programa, encuestas basadas en cuestionarios estructurados o semi - estructurados. El personal aportará información acerca de cómo ellos entienden que el programa debe implementarse y como lo están ejecutando en la práctica.
- *Usuarios del programa.* La recolección de información de los usuarios puede tomar formas similares a como se obtiene la información del personal. Los usuarios aportarán información acerca de la satisfacción que les genera el programa, los servicios (o bienes) que reciben, su periodicidad, su percepción del programa, entre otros aspectos.

4. EVALUACIÓN EX - POST

4.1. Concepto de Evaluación Ex – Post

Como se ha señalado en la Unidad 1 de este módulo, las evaluaciones ex – post son aquellas que se implementan una vez concluida la fase de implementación del programa o, al menos, concluida una etapa del programa. Las evaluaciones ex – post son las más desarrolladas (Cohen y Franco 2000: 218) y sofisticadas, desde un punto de vista metodológico, y son también las que han ido concitando el mayor interés en desde la década de 1990.

El atractivo de las evaluaciones ex – post es que permiten identificar los resultados, impactos, efectividad y beneficios netos alcanzados. Con ello, este tipo de evaluaciones permiten lograr un conocimiento relevante para el análisis social, toda vez que identifican los tipos de programas y modos de intervenir que más probablemente generarán los resultados e impactos deseados, y que tipo de intervenciones o en que contextos ellas no serán efectivas en lograr los efectos, beneficios y resultados previstos.

En el mismo sentido, al lograr generar conocimiento relevante, este tipo de evaluaciones orienta el diseño de nuevos programas, que presumiblemente podrán tener mejores resultados, mayores beneficios netos e impacto – que si no existiera este conocimiento –, dado que están basados en similares experiencias anteriores.

El conocimiento producidos por estas evaluaciones ex - post, asimismo, podrá apoyar la toma de decisiones sobre programas y proyectos de interés público, con información cierta obtenida de la aplicación de rigurosos métodos de investigación de ciencias sociales.

4.2. Tipos de Evaluación Ex - Post

El siguiente esquema presenta una sistematización de tipos de evaluación ex - post.

- Evaluación de Eficiencia
 - Análisis Costo - Beneficio
 - Análisis Costo - Efectividad

- Evaluación de Resultados

- Evaluación de Impacto
 - Programas de Cobertura Parcial
 - ✓ Diseño Experimental
 - ✓ Diseño Cuasi - Experimental

 - Programas de Cobertura Total
 - ✓ Controles Reflexivos
 - ✓ Controles Sombras

Los análisis Costo - Beneficio y Costo - Efectividad fueron presentados en la sección 2, sobre evaluaciones ex - ante, como evaluaciones que se efectúan con anterioridad a la implementación del

programa. En ese caso, este tipo de evaluaciones provee antecedentes e información valiosa que permite decidir si el programa será puesto en práctica o seleccionar la alternativa de intervención que a - priori se presenta como más eficiente.

Adicionalmente, las evaluaciones de eficiencia también pueden hacerse una vez concluido el programa. En ese caso, entregarán información acerca de los beneficios (o costos) netos que el programa realmente produjo.

Dado que la evaluación de eficiencia ha sido presentada en la sección 2 de este documento, las siguientes secciones de esta unidad se concentrarán en la descripción de las evaluaciones de resultados e impacto.

4.3. Evaluación de Resultados

El Sistema Nacional de Inversiones (SNI) del MIDEPLAN (2007) señala que “el objetivo de la evaluación de resultados es determinar la pertinencia y el grado de logro de los objetivos planteados originalmente” al programa.

Conforme a ello, este es un tipo de evaluación que se implementa una vez concluido el proyecto o programa o, al menos, cuando ha terminado la implementación de una etapa significativa.

La evaluación de resultados necesita que el programa que está siendo evaluado haya sido formalizado, tanto en su diseño, su cronograma

de implementación, como también en las metas que debía lograr. Los resultados que alcance la implementación serán comparados a lo expresado en el diseño, cronograma y metas previstas del programa para “determinar la pertinencia y el logro de los objetivos planteados originalmente”.

a) Objetivos de la Evaluación de Resultados

Basado en las formulaciones del Sistema Nacional de Inversiones (SIN) del MIDEPLAN (2007), la evaluación de resultados se orienta a los siguientes objetivos:

- Determinar el grado de cumplimiento de los objetivos establecidos a las iniciativas de inversión (proyectos y programas) y la validez de las proyecciones ex - ante;
- Generar información para retroalimentar el diseño de nuevos programas;
- Identificar correcciones y rectificaciones necesarias para mejorar e incrementar los resultados que logren los programas y proyectos;
- Identificar mejoras que deben introducirse en los métodos y criterios de preparación, formulación y evaluación ex - ante de programas.

b) Información para la Evaluación de Resultados

La indicación básica acerca de qué información recoger para preparar el informe de evaluación de resultados está dada en el propio documento que formaliza el programa (ver, Unidad 2 de este Módulo). En él se identifican los objetivos buscados, los recursos asignados (dinero, instalaciones, equipos, personal, tiempo), las actividades consideradas y su cronograma de implementación. A su vez, los objetivos serán descompuestos en metas y éstas, en indicadores.

Así, la información que se deberá recoger de la ejecución real será aquella relativa a las actividades desarrolladas, al consumo de recursos y a los indicadores de logro alcanzados. La comparación entre lo logrado y lo planeado será la esencia de la evaluación de resultados.

La evaluación de resultados, además, deberá identificar las lecciones y aprendizajes que deja la experiencia. Para ello, es necesario recoger las impresiones de usuarios, comunidad y personal del programa.

c) El Informe de la Evaluación de Resultados

La CEPAL (2007) sugiere la siguiente estructura para un informe de evaluación de resultados.

- Resumen del proyecto o programa
- Objetivos del proyecto o programa
- Resultados
- Lecciones aprendidas

- Recomendaciones para futuras experiencias
- Comentarios adicionales

4.4. Evaluación de Impacto

a) Finalidad de la Evaluación de Impacto

La evaluación de impacto busca identificar efectos – deseados y no deseados – que se producen como consecuencia de la implementación de un programa o proyecto. Así, la pregunta que una evaluación de impacto intenta responder es ¿qué cambios ocurridos en la realidad pueden ser razonablemente atribuidos a la intervención analizada (programa o proyecto)?.

El objetivo básico de una evaluación de impacto es producir una estimación de los efectos netos de una intervención; esto es, una estimación del impacto de la intervención que no esté contaminado por la influencia de otros procesos y eventos que también pueden haber afectado la conducta o condiciones en las cuales la intervención (programa o proyecto) que está siendo evaluada es implementada.

Los efectos netos de la intervención son aquellos cambios en la realidad que pueden – razonablemente – ser identificados como consecuencias de la intervención analizada. Ello implica poder separar el efecto que produce el programa o proyecto en análisis de los efectos producidos por otras intervenciones, estímulos o variables que están presentes en la misma realidad en que actúa la intervención (programa o proyecto) cuyo impacto está siendo evaluado.

El mayor desafío de la evaluación de impacto es separar el efecto neto de un programa del efecto bruto observado. El efecto o resultado bruto consiste en el cambio total o medida final que se obtiene. En otras palabras, es la diferencia total que se obtiene de la comparación de los cambios observados en dos grupos en un cierto espacio de tiempo, o de la comparación antes – después. Los resultados netos son aquellos que pueden razonablemente atribuirse a la intervención, descontados los efectos de otras causas que pudieran haber estado influyendo.

Desde el punto de vista de la evaluación de impacto la comprensión de este tema es de vital importancia, pues lo que se observa a simple vista en la realidad es el resultado bruto, es decir, la totalidad de cambios que ocurren en la realidad en un cierto periodo. Éstos incorporan las variaciones producidas por la intervención y aquellas producidas por otros factores. Así, siendo el resultado bruto los cambios totales que es posible observar en la realidad analizada, la evaluación de impacto apunta a poder identificar aquella parte de los resultados brutos que pueden ser, razonablemente, atribuidos a la intervención en análisis. El siguiente esquema presenta la relación entre resultado bruto y efecto neto del programa en evaluación.

$$\text{Resultados Brutos} = \left\{ \begin{array}{c} \text{Efectos} \\ \text{Netos de la} \\ \text{Intervención} \end{array} \right\} + \left\{ \begin{array}{c} \text{Efecto de} \\ \text{otros} \\ \text{procesos} \\ \text{(factores} \\ \text{extraños)} \end{array} \right\} + \left\{ \begin{array}{c} \text{Efectos} \\ \text{del} \\ \text{Diseño} \end{array} \right\}$$

Fuente: Rossi y Freeman (1993)

La evaluación de impacto apunta primariamente a la determinación del efecto neto.

b) Consideraciones para la Evaluación de Impacto

El diseño de una evaluación de impacto debe tomar en cuenta dos tipos de presiones competitivas:

- las evaluaciones deben ser implementadas con suficientes rigor para que puedan obtenerse conclusiones relativamente firmes.
- Deben tenerse en cuenta consideraciones prácticas de tiempo, costo, cooperación y protección de la persona humana. Ellas limitan las opciones de diseño y metodología que pueden ser implementadas.

Todas las evaluaciones de impacto se hacen a través de estudios comparativos. Determinar el impacto de un programa requiere comparar, con el mayor rigor, las condiciones de aquellos que experimentaron la

intervención con aquellos grupos “equivalentes” que experimentaron un tratamiento alternativo o simplemente no recibieron tratamiento alguno.

La identificación del grupo de comparación es una cuestión crítica en la práctica de la evaluación de impacto, pues permite responder a la pregunta de “que podría haber ocurrido en la realidad en ausencia de la intervención en análisis”. A esto se le llama el “contrafactual”.

La identificación de grupos que recibieron la intervención (grupo de tratamiento) y de grupos que no la recibieron o recibieron alguna intervención alternativa, indica que una cuestión metodológica básica para efectuar la evaluación de impacto de una intervención es que ésta tenga cobertura parcial. Siendo el programa o proyecto de cobertura parcial, habrá un segmento significativo de la población que presenta similares características relevantes para los propósitos de la intervención, que no ha recibido los efectos de la intervención y desde el cual se podrá obtener suficientes casos para construir el grupo de comparación.

Aunque conclusiones más firmes de las evaluaciones de impacto es posible obtenerlas de evaluaciones de programas de cobertura parcial, también es posible evaluar el impacto de programas de cobertura total. En este caso, habrá que estar atento a identificar – y documentar – que otras variables pudieran estar influyendo en la realidad observada, aparte de la intervención que está siendo evaluada, como también identificar sesgos y problemas que pudiera presentar el método elegido. Las secciones siguientes de estas notas abordan este tema.

Los evaluadores evalúan el resultado de programas sociales a través de:

- la comparación de información acerca de participantes y no participantes en el programa.
- Hacer mediciones sobre los participantes a través del tiempo. La práctica más común en estas mediciones es comparar la situación antes y después de la intervención.
- Otros métodos que pueden producir resultados equivalentes a esas comparaciones. Esta consideración es especialmente cuando se quiere evaluar el impacto de programas de cobertura total.

c) Tipología de los Estudios de Evaluación de Impacto

A continuación se presenta una clasificación de los estudios de evaluación de impacto basada tanto en la cobertura del programa en evaluación, como en los métodos empleados. Las siguientes secciones de esta unidad abordarán en mayor detalle la descripción de cada uno de estos tipos de evaluación.

- Programas de Cobertura Parcial
 - Diseño Experimental
 - Diseño Cuasi – Experimental

- Programas de Cobertura Total
 - Controles Reflexivos
 - Controles Sombras
 - ✓ Panel de Expertos
 - ✓ Evaluación de Administradores del Programa
 - ✓ Evaluación de Usuarios

4.5. Evaluación de Impacto basada en un Diseño Experimental

El modelo experimental es el más robusto diseño de investigación para evaluar el impacto neto de una intervención. La lógica del diseño experimental es la base para el diseño de todos los tipos de evaluación de impacto.

Un diseño experimental de evaluación de impacto sólo puede ser empleado para evaluar el impacto de programas de cobertura parcial. Los programas de cobertura parcial son aquellos que llegan sólo a una parte de las personas que conforman el grupo objetivo al que está orientado el programa. La evaluación de impacto basada en un diseño experimental no es aplicable, en consecuencia, a programas que alcanzan cobertura total de la población.

a) Conceptos

El diseño experimental se basa en una comparación entre dos grupos de población objetivo: uno que recibe el tratamiento y otro que no

lo recibe. La asignación de las personas a uno y otro grupo se hace al azar. Según Baker (2000: 2) "al distribuir aleatoriamente la intervención entre los beneficiarios calificados, el proceso de asignación mismo crea grupos de tratamiento y de comparación comparables que son estadísticamente equivalentes entre sí, a condición de que las muestras sean de tamaño adecuado". De este modo, el resultado neto de una intervención puede ser entendido como la diferencia entre personas que han participado en el programa y personas comparables que no han recibido el tratamiento.

La expresión tratamiento hace referencia al servicio o bien que es distribuido por el programa que está siendo evaluado. De este modo, quienes reciben ese servicio o bien son considerados en este método, como el grupo de tratamiento. Por el contrario, quienes no reciben ese "tratamiento" se constituyen en el grupo de comparación. Así, asignando a personas que provienen de la población objetivo a que se orienta el programa a alguno de los dos grupos – de tratamiento y de comparación – es posible hacer apropiadas comparaciones entre quienes recibieron las intervenciones y personas comparables que no lo recibieron.

Si se obtiene perfecta comparabilidad, los mismos eventos o situaciones que pueden tener incidencia en los resultados observados se encontrarán presentes en ambos grupos. Siendo ambos grupos comparables, la única diferencia entre ellos estará causada por la intervención y por los efectos del diseño de investigación.

De este modo, los efectos netos de una intervención, derivados de la aplicación de un diseño experimental, se pueden presentar esquemáticamente como sigue:

$$\text{Efectos Netos} = \left\{ \begin{array}{c} \text{Resultado} \\ \text{Bruto del} \\ \text{Grupo de} \\ \text{Tratamiento} \end{array} \right\} - \left\{ \begin{array}{c} \text{Resultado} \\ \text{Bruto del} \\ \text{Grupo de} \\ \text{Comparación} \end{array} \right\} + \left\{ \begin{array}{c} \text{Efectos} \\ \text{del} \\ \text{Diseño} \end{array} \right\}$$

Fuente: Rossi y Freeman (1993)

En términos ideales, comparabilidad significa que ambos grupos son idénticos y que su única diferencia radica en el hecho de haber participado o no participado en el programa de evaluación. De este modo, al haber comparabilidad es posible establecer causalidad, es decir, que los efectos identificados pueden ser atribuidos a la intervención.

La principal ventaja del modelo experimental es que sus resultados se pueden interpretar fácilmente, "puesto que el efecto del programa sobre el resultado que se evalúa se puede medir a través de la diferencia entre las medias de las muestras del grupo de tratamiento y del grupo de comparación" (Baker, 2000: 2).

Para que dos grupos sean comparables se requiere:

- *Idéntica composición*: los grupos de tratamiento y de comparación contengan la misma mezcla de personas y características;
- *Idénticas experiencias*: Durante el tiempo de observación, ambos grupos – de tratamiento y de comparación – experimentan los mismos procesos. Por ejemplo: ambos son afectados por las mismas

variables económicas, la composición etaria se mantiene equivalente en los dos grupos, o procesos similares que afecten a los dos grupos igualmente. La única diferencia es que el grupo de tratamiento recibe la intervención y el grupo de comparación no la recibe;

- *Idéntica predisposición:* ambos grupos - de tratamiento y de comparación - están igualmente predispuesto hacia el programa. Es decir, las tendencias de autoselección deben ser idénticas en ambos grupos.

La comparabilidad entre ambos grupos es muy importante para poder obtener conclusiones firmes acerca del impacto neto de un programa en evaluación. El mejor modo de lograr esta comparabilidad es asignar al azar miembros de la población objetivo al grupo de tratamiento y al grupo de comparación.

Asignar al azar - a personas comparables - a ambos grupos requiere de asegurarse que todas las personas de la población objetivo tienen la misma oportunidad de ser seleccionados para alguno de los dos grupos (de tratamiento y de comparación) sobre los cuales se basa la evaluación de impacto.

b) Estimación del Impacto

La estimación del impacto normalmente se obtiene de efectuar comparaciones, entre el grupo de tratamiento y el grupo de comparación, en los momentos antes y después de la intervención. A este tipo de

comparación se le conoce como el método de la doble diferencia. Algebraicamente se presenta de la siguiente forma:

$$\text{Impacto} = (GT_{t1} - GT_{t0}) - (GC_{t1} - GC_{t0}) ; \text{ donde}$$

GT es el grupo de tratamiento,

GC es el grupo de comparación,

t1 representa al momento después de la intervención, y

t0 representa al momento antes de la intervención.

c) Limitaciones del Diseño Experimental

Inicialmente el diseño experimental se construyó para uso en laboratorios y luego su uso se ha extendido en evaluación de programas sociales, aunque hay limitaciones en su aplicación en este último campo. Entre las dificultades que encuentra la aplicación del diseño experimental en las áreas sociales se encuentran:

❖ Problemas Prácticos

- *Consideraciones éticas.* Hay argumentos que señalan que no sería ético dar un tratamiento, que implicaría un beneficio, a un grupo y no dárselo a otro grupo que está en las mismas condiciones.

El contrargumento de ello es que sin la de evaluación de impacto no se sabe si, efectivamente, el programa será un beneficio o no, por lo que nadie estaría siendo excluido a priori de algo que no es un beneficio cierto. Si como consecuencia de la evaluación de

impacto se concluye que el programa es beneficioso se podrá extender su impacto a toda la población objetivo.

- *Tiempo y costos.* El diseño experimental consume tiempo y dinero. Un diseño experimental no debe implementarse cuando se requiere información del impacto del programa con cierta premura. Adicionalmente, los requerimientos técnicos y precisiones hacen que el diseño experimental no sea barato de implementar.

❖ Problemas de Validez Interna

Stock y Watson (2007) identifican las siguientes situaciones que podrían generar problemas de validez interna a los diseños experimentales.

- *Fallas en establecer una distribución aleatoria en los grupos de tratamiento y comparación.* Dado que el diseño experimental exige una distribución al azar de la muestra entre los grupos de tratamiento y de comparación, sus resultados presentarán problemas de validez interna si es que esa distribución entre ambos grupos no es aleatoria. Las fallas en aplicar un procedimiento de estas características hará que ambos grupos no sean idénticos, producto de lo cual los resultados del análisis estarán sesgados y, por tanto, serán inválidos.
- *Fallas en la aplicación del tratamiento.* El diseño experimental exige mucho control sobre la aplicación del tratamiento, por lo que si el personal a cargo no lo aplica correctamente o si los beneficiarios no usan el bien o servicio, o muestran un cumplimiento parcial del protocolo de implementación del tratamiento, habrá un problema

de validez interna, lo que implica que los resultados del experimento estarán sesgados.

- *Abandono o deterioro de la muestra (Atrición)*. Durante el periodo de evaluación personas incorporadas a la muestra pueden tener cambios en las características que las hicieron formar parte de este grupo o abandonar el experimento (por ejemplo: casarse, enfermarse gravemente, cambiar de residencia, morir y similares), lo que podría contaminar o sesgar los resultados del estudio, si este hecho no es advertido.
- *Efectos del experimento*. Puede ocurrir que el solo hecho de formar parte de la implementación de un experimento genere un estímulo o motivación que haga cambiar la conducta de las personas involucradas (sin siquiera recibir el tratamiento o mostrar una sobre-reacción), lo que puede llegar a afectar los resultados del experimento e invalidarlo. En este caso, las personas se comportan de manera diferente a como lo harían normalmente, por lo que los resultados que arroja el experimento no logran capturar efectivamente el efecto de la intervención.
- *Muestra pequeña o de tamaño insuficiente*. En este caso la muestra no presenta el número mínimo de casos que permiten interpretar "válidamente" el experimento.

❖ Problemas de Validez Externa

Los problemas de validez externa hacen referencia a las dificultades que enfrenta el experimento para que sus conclusiones sean generalizables al universo desde el cual se dedujo la muestra. Rossi y Freeman (1993) denominan a esta situación como problema de generalización, en tanto que Cohen y Franco lo denominan como problema de validez externa. En este ámbito, las dificultades que enfrenta un diseño experimental son las siguientes:

- *Muestra no representativa.* Hace referencia a que la muestra con la que se trabaja no recoge las características del universo de referencia o a la imposibilidad de generalizar las conclusiones del experimento hacia poblaciones que presentan características disímiles de la muestra a la que se ha aplicado el experimento. Por ejemplo, las conclusiones de la evaluación de impacto de un programa destinado al control de la delincuencia, aplicado en comunas con bajos índices delictuales, serán difícilmente generalizables a comunas con altos indicadores de delitos – y normalmente con diferentes realidades sociales.
- *Diferencias entre el experimento y el tratamiento real.* El tratamiento en condiciones experimentales puede ser diferente de las condiciones que el programa enfrente cuando se implemente en realidad. Los programas piloto normalmente disponen de recursos suficientes, instalaciones adecuadas, personal especialmente capacitado y motivado, cuestiones que difícilmente se tendrán en la etapa de implementación del programa. Las diferencias que se dan

entre el experimento – prueba piloto – y las condiciones reales de implementación que encontrará luego el programa plantean problemas a la generalización de los resultados de una evaluación de impacto.

4.6. Evaluación de Impacto basada en un Diseño Cuasi – Experimental

Al igual que el diseño experimental, el diseño cuasi-experimental puede ser usado en la evaluación de programas con cobertura parcial. Es decir, cuando hay población objetivo que no ha sido expuesta a la intervención.

Idealmente, cuasi – experimentos deben ser diseñados previo a la implementación del programa, aunque diseños cuasi – experimentales ex - post también son comunes.

a) Fundamentos

El término cuasi- experimento hace referencia a los diseños de evaluación en que la conformación de los grupos de tratamiento y comparación no es hecha al azar. Puede ocurrir que el evaluador no tenga incidencia en el desarrollo de la intervención, que la conformación de un grupo de comparación no pueda hacerse al azar, ya sea por razones éticas, políticas o de otra índole, o porque la decisión de evaluar el impacto del programa se haya tomado cuando éste ya se está implementando, por ejemplo. En estos casos la evaluación basada en un modelo cuasi – experimental es una alternativa aconsejable.

Ocurre con frecuencia que el propio diseño del programa identifica a los usuarios que recibirán el tratamiento. Para poder realizar una evaluación de impacto es necesario construir un grupo de comparación que se asemeje lo más posible al grupo que recibe el tratamiento. La cuestión es que el grupo de comparación se asemeje lo más posible al grupo de tratamiento en aquellas características que son relevantes para los resultados de la intervención.

El diseño cuasi - experimental es un diseño robusto y muy frecuentemente aplicado. El punto de cuidado es si el grupo de comparación que se ha elegido es suficientemente comparable al grupo que ha recibido el tratamiento en todas las características y experiencias relevantes para el resultado del programa o proyecto que está siendo evaluado.

Los efectos netos de una intervención, derivados de la aplicación de un diseño cuasi - experimental, se pueden presentar esquemáticamente como sigue:

$$\text{Efectos Netos}^3 = \left\{ \begin{array}{c} \text{Resultado} \\ \text{Bruto del} \\ \text{Grupo de} \\ \text{Trata-} \\ \text{miento} \end{array} \right\} - \left\{ \begin{array}{c} \text{Resulta-} \\ \text{do Bruto} \\ \text{del} \\ \text{Grupo de} \\ \text{Compara-} \\ \text{ción} \end{array} \right\} + \left\{ \begin{array}{c} \text{Diferen-} \\ \text{cias no} \\ \text{contro-} \\ \text{ladadas} \\ \text{entre GT} \\ \text{y GC} \end{array} \right\} + \left\{ \begin{array}{c} \text{Efectos} \\ \text{del} \\ \text{Diseño} \end{array} \right\}$$

Fuente: Rossi y Freeman (1993)

³ GT indica Grupo de Tratamiento y GC hace referencia al Grupo de Comparación.

La diferencia crítica entre el diseño experimental y el cuasi - experimento es que en el segundo hay que tener especial cuidado en la identificación de las diferencias entre ambos grupos de análisis - que pueden ser muy difíciles de observar y con el sesgo de selección. Ambos problemas son abordado más adelante en esta sección.

b) Tipos de Cuasi - Experimento

- *Ex - ante*. En esta modalidad el diseño de evaluación se establece antes que el programa se implemente, de modo similar a como ocurre en el diseño experimental. Ello permite a los evaluadores mayores y mejores oportunidades de planificar la recolección de datos acerca de la población objetivo y construir un grupo de comparación que replique las características - relevantes para los efectos de la evaluación - del grupo de tratamiento.
- *Ex - post*. Por las razones anotadas al inicio de la sección precedente, ocurre, con frecuencia, que el diseño de evaluación debe hacerse luego que la etapa de implementación del programa se ha iniciado o, también, una vez que ésta ha concluido. Dada esta limitación, el tema crítico en los cuasi - experimentos ex - post es la construcción de un grupo de comparación que se asemeje - lo más posible - al grupo de tratamiento en todas aquellas características relevantes para la evaluación.

c) Alternativas de Construcción del Grupo de Comparación

El diseño cuasi - experimental es posible aplicarlo en programas de cobertura parcial. De este modo, es posible construir un grupo de comparación de aquel segmento de la población objetivo del programa que no ha recibido el servicio o bienes distribuido por el programa. Rossi y Freeman (1993) identifican las siguientes opciones de construcción de grupos de comparación.

- *Pareo (Matching)*. La idea central de esta opción es constituir un grupo equivalente y paralelo, que se asemeja al grupo de tratamiento en aquellas características relevantes para la evaluación, pero que no ha sido expuesto al tratamiento. El grupo de tratamiento se identifica primero y luego se establece el par equivalente, es decir, el grupo de comparación.

La construcción del grupo de comparación, en esta opción, puede darse a través de un pareo individual o mediante un pareo grupal.

- o *Pareo Individual*. Aquí el evaluador busca un par para cada uno de los miembros del grupo que recibirá la intervención. Este par debe tener las mismas características relevantes que el respectivo miembro del grupo de tratamiento. La ventaja que ofrece el pareo individual es la posibilidad de disponer de un "gemelo" de cada uno de los miembros del grupo de tratamiento, lo que permite generar conclusiones firmes y robustas, dada la alta semejanza que adquieren los grupos de tratamiento y de comparación. Sin embargo, las dificultades

para encontrar un “gemelo” de cada miembro del grupo de tratamiento hace, con frecuencia, que se pierdan casos (miembros del grupo de tratamiento), lo que hace que se reduzca la muestra y con ello se pueda afectar la validez de los resultados de la evaluación. Esto se constituye en una importante desventaja del pareo individual.

- o *Pareo Grupal*. En este caso el evaluador construye un grupo de comparación cuidando que haya correspondencia en la composición, o distribución general, de ambos grupos. Es decir, la composición del grupo de comparación debe ser equivalente a las características – relevantes para la evaluación – que presenta el grupo de tratamiento.
- *Controles estadísticos*. En este caso, una vez establecido el grupo de tratamiento, el grupo de comparación es construido a partir de la población no expuesta a la intervención y que cumpla los criterios de elegibilidad del programa. A partir de ello, los controles estadísticos permiten identificar como difieren los grupos de tratamiento y comparación, de acuerdo a características seleccionadas, y, de esa manera, estimar el impacto.

Esta opción es algo superior al pareo, dado que pone una fuerte atención en identificar características que podrían estar correlacionadas con los resultados y efectos observados. Ello es importante, pues participantes y no participantes en el programa se diferencian en una serie de características que no logramos observar

a priori, las que, de no ponérsele atención, pueden influir en generar resultados sesgados de la evaluación.

- *Controles genéricos.* La lógica detrás de esta modalidad de comparación es que indicadores de – o derivados de – la población general son usados como si fueran el grupo de comparación. Así los indicadores que se obtienen de haber aplicado el programa a la población usuaria son comparados con los mismos indicadores que se disponen de la población general. Por ejemplo, una estimación del impacto producido por una intervención orientada a generar mayor seguridad ciudadana basada en una comparación que se hiciera sobre la variación, en un cierto año, de los índices de denuncia de delitos que se registran en una comuna respecto de los registrados en la región o país, sería una evaluación basada en controles genéricos.

Esta manera de estimar el impacto de un programa, aunque simple y fácil de realizar, produce, sin embargo, resultados de baja confiabilidad, pues la impresión que genera es que los cambios observados en la realidad se habrían producido a consecuencia de la intervención. Dado que otros estímulos podrían haber estado influyendo en los resultados observados, pero que no son identificados por el método de comparación, no es posible entregar conclusiones firmes cuando se compara por controles genéricos. Por ello, este tipo de análisis debe emplearse sólo cuando no es posible emplear otras alternativas de evaluación de impacto. Aún así, es necesario identificar qué otras variables, estímulos o condiciones ambientales estaban presentes al momento de la intervención y que

- de acuerdo al conocimiento que se disponga sobre el fenómeno en análisis - pudieran haber influido en los cambios observados.

La evaluación del "Programa de formación laboral para mujeres en situación de riesgo social", implementado por la Fundación de Promoción Familiar, puede ser visto como un caso de aplicación de controles genéricos. El programa otorga capacitación laboral en "corte y confección" y apoyo psicosocial a mujeres recién egresadas de la cárcel y a su entorno familiar más inmediato, especialmente hijos. Aunque hubo una deserción importante⁴, pues partieron 25 y terminaron 15 mujeres, el programa es visto como exitoso porque sólo dos mujeres reincidieron en conductas delictuales. El porcentaje de reincidencia de las participantes en el programa es 13%, si se considera al grupo que recibió el tratamiento completo, u 8%, si se considera al grupo con el que se inició el programa. Ambas cifras son vistas como muy positivas por los impulsores del programa, pues la comparan con el hecho que el 87% de las mujeres que están en la cárcel son reincidentes.

Más allá de si este es el indicador adecuado con el que se debe comparar los resultados alcanzados por el programa, este análisis ilustra un caso de aplicación de controles genéricos en la evaluación de un programa.

⁴ La deserción estuvo motivada por enfermedad y problemas económicos de las inscritas.

d) Limitaciones del Diseño Cuasi – Experimental

Al igual que el diseño experimental, un cuasi – experimento también enfrenta problemas prácticos y problemas metodológicos que pueden afectar su validez interna y/o externa. El debate ético que puede afectar al diseño experimental también se aplica al cuasi – experimento. Del mismo modo, las dificultades de tiempo y costo (señaladas en la sección 4.5) también afectan a las evaluaciones de impacto basadas en diseños cuasi – experimentales. Las situaciones que afectan la validez interna de un cuasi – experimento son similares a las que afectan a un experimento y que han sido tratadas en la sección 4.5 de este documento. Así, el diseño cuasi – experimental también enfrenta las amenazas de los problemas originados por fallas en la aplicación del tratamiento, por atrición o abandono o deterioro de la muestra, por efectos del experimento y por una muestra de tamaño insuficiente. Aparte de los señalados, el cuasi – experimento también enfrenta los siguientes problemas de validez interna:

❖ Otros Problemas de Validez Interna

- *Problemas de comparabilidad entre los grupos de tratamiento y control.* Siempre existe la posibilidad que una o más diferencias existan entre los miembros de los grupos de tratamiento y de comparación, y que estas diferencias – no la intervención – sean la causa de todo o parte de los impactos estimados. Por ejemplo, un programa de reinserción social de jóvenes en conflicto con la ley puede llegar a mostrar que la intervención tuvo un alto impacto. Un análisis más detallado de quienes son los usuarios del programa puede llegar a identificar que en este grupo hay un alto número de

personas que son casados y que tienen hijos – respecto al grupo de comparación – y que este puede ser el factor que explique la reinserción social observada – motivada por los deseos de estar junto a la familia – y no la intervención misma.

- *Sesgo de selección.* Este ocurre cuando los procedimientos de selección de usuarios del programa incorporan – inadvertidamente – a personas cuyas características personales, capacidades individuales o condiciones pre – existentes se relacionan más favorablemente con resultados e impactos positivos. De este modo, el impacto observado puede estar relacionado con esas capacidades, características o condiciones y no con la intervención misma. Por ejemplo, un programa de capacitación y habilitación laboral que incorpora a jóvenes deseosos de incorporarse al mercado laboral puede mostrar mejores resultados e impactos – en términos de egresados del programa que han encontrado trabajo – que una intervención enfocada en jóvenes en riesgo social. En este caso, puede ser, justamente, el deseo de los jóvenes de trabajar lo que explique el efecto observado y no la implementación del programa. La autoselección también es una variante de este tipo de sesgos, que se da normalmente en programas a los que hay que postular. Puede ser que la motivación por recibir los beneficios del programa sean la explicación de los resultados y efectos observados y no el programa mismo. En este caso la pregunta que surge es ¿sin la predisposición positiva hacia el programa que muestran los usuarios habría tenido la intervención los mismos resultados e impactos?.

❖ Problemas de Validez Externa en el Cuasi – Experimento

De acuerdo a Stock y Watson (2007) los siguientes son los principales problemas de validez externa que enfrenta un diseño cuasi – experimental.

- *Diferencias poblacionales.* Las diferencias en las características de los grupos estudiados y de la población de interés o referencia – a la que se le quisiera aplicar el tratamiento si se verifica un impacto positivo – hace que las conclusiones que arroja el cuasi – experimento no sean aplicables a la población de interés o referencia.
- *Diferencias en los contextos* del grupo estudiado respecto de la población de interés o referencia. Aún cuando los grupos estudiados y de referencia puedan ser comparables, las diferencias contextuales que enfrentan unos y otros pueden hacer que las conclusiones del cuasi – experimento no sean generalizables a la población de interés o referencia.

4.7. Evaluación de Programas de Cobertura Total

Las secciones precedentes de esta Unidad han mostrado modos alternativos para evaluar el impacto de programas con cobertura parcial. En los métodos reseñados el impacto se estima a partir de la comparación de quienes reciben la intervención, con quienes no la reciben. Este último grupo es obtenido de la población objetivo que no ha sido expuesta al programa. La cuestión entonces es como evaluar programas cuando hay

insuficiente población objetivo sin recibir el tratamiento. Este es el caso de programas que alcanzan cobertura total.

Siguiendo a Rossi y Freeman (1993) los tipos de evaluación más comunes para este tipo de programas son las evaluaciones basadas en controles reflexivos y/o controles sombras.

a) Controles Reflexivos

Reciben esta denominación aquellas evaluaciones de impacto en que la población objetivo sirve como su propia comparación. Lo que se compara es la situación de la población usuaria del programa, antes de recibir la intervención y después de haberla recibido. También puede darse a través de mediciones antes del inicio del programa y sucesivas mediciones durante la ejecución y luego del término del programa. A esta práctica se le llama comparaciones simples antes/después, en el primer caso, o análisis de series de tiempo – cuando los datos de las distintas mediciones provienen de grupos con distinta composición – o panel de datos – cuando los datos provienen de mediciones tomadas de las mismas personas, en distintos momentos –, en el segundo caso. El impacto se deduce de comparar los dos conjuntos de mediciones: las mediciones que se toman después – o sucesivas mediciones durante la ejecución – de la intervención son comparadas a las mediciones que se tomaron antes del inicio del programa. La diferencia entre ambos tipos de mediciones identificará el impacto del programa.

El supuesto detrás de los controles reflexivos es que no hay eventos contaminantes, estímulos o cambios en la población objetivo, aparte de

aquellos producidos por la intervención, que hayan ocurrido entre las observaciones en análisis. Así, se asume que cualquier cambio que ha ocurrido a través del tiempo es efecto neto de la intervención.

El problema con este tipo de comparaciones es que, cuando el tiempo transcurrido entre los dos tipos de mediciones es apreciable, es más probable que algunos procesos que ocurren en el período de tratamiento – y que no están relacionados con la intervención – puedan influir en sus resultados, bien ampliándolos o reduciéndolos. En general, las evaluaciones basadas en diseños antes/después entregan resultados que tienen un bajo nivel de credibilidad porque, en contraste a los diseños experimentales y cuasi – experimentales, no es posible tomar en cuenta los efectos de otras variables o situaciones que pueden estar presentes o coexistiendo con la intervención. Los análisis de series de tiempo pueden lograr mayor confiabilidad, dado que sucesivas mediciones son efectuadas a través del tiempo, con lo que se logra trazar la tendencia de cambios que la población usuaria del programa ha ido mostrando a lo largo de la implementación.

La evaluación de impacto basada en controles reflexivos se presenta esquemáticamente de la siguiente forma:

$$\begin{array}{l}
 \text{Impacto} \\
 = \\
 \text{Neto}
 \end{array}
 =
 \left\{ \begin{array}{c} \text{Resultado} \\ \text{de los} \\ \text{participantes} \\ \text{después} \\ \text{del} \\ \text{tratamiento} \end{array} \right\}
 -
 \left\{ \begin{array}{c} \text{Resultado} \\ \text{de los} \\ \text{participantes} \\ \text{antes} \\ \text{del} \\ \text{tratamiento} \end{array} \right\}
 +
 \left\{ \begin{array}{c} \text{Efectos} \\ \text{de otros} \\ \text{procesos} \\ \text{durante} \\ \text{el} \\ \text{tratamiento} \end{array} \right\}
 +
 \left\{ \begin{array}{c} \text{Efectos} \\ \text{del} \\ \text{Diseño} \end{array} \right\}$$

Fuente: Rossi y Freeman (1993)

En general, el uso de controles reflexivos no es recomendado en circunstancias en las cuales diseños cuasi – experimentales son posibles de aplicar. Controles reflexivos pueden ser útiles en evaluaciones de programas como un primer paso simple, rápido y de bajo costo, especialmente si no hay indicios que permitan suponer que los resultados de la población objetivo hubieran cambiado sin la intervención. Asimismo, los simples diseños antes/después serían apropiados principalmente para evaluaciones de impacto de corto plazo de programas de cobertura total.

b) Controles Sombras

Los diseños de evaluación basados en controles sombras se basan en la opinión de expertos, administradores del programa o de la población objetivo misma para estimar si el programa produce los efectos deseados. El término controles sombras apunta a mostrar que este tipo de evaluaciones normalmente se basa en opiniones y no en evidencia sustantiva (Rossi y Freeman 1993).

Aunque estas evaluaciones son comunes, dada su simplicidad, rapidez y bajo costo, ellas no son recomendables dado que es imposible estimar el efecto neto en ellas. Los diseños basados en controles sombras pueden ser usados en programas pequeños, de importancia menor y solo cuando hay una buena posibilidad de asumir que la evaluación podrá identificar cambios relevantes que se hayan producido a consecuencia de la implementación del programa. En este caso, el supuesto es que, al menos, este tipo de evaluación podrá estimar los efectos brutos – aunque muy difícilmente podrá identificar los efectos netos.

- *Evaluación de expertos.* Normalmente uno o más prestigiados expertos en un campo del conocimiento o área de actividad profesional son contratados para analizar un programa y emitir un juicio fundado sobre los resultados alcanzados, su efectividad y/o impacto. El trabajo de los expertos consiste en analizar el desarrollo del programa, examinar los registros y archivos, recoger información de diversa fuente – por ejemplo, entrevistas a usuarios, funcionarios o entorno del programa – formarse un juicio acerca de la intervención y evacuar un informe de evaluación.

Este tipo de evaluaciones puede presentar sesgos que, en algunos casos, pueden llegar a ser considerables. Es normal que entre los expertos haya desacuerdos sobre los temas de su dominio, por diversidad de enfoques conceptuales, aproximaciones metodológicas, conclusiones de sus propios estudios anteriores, entre otros aspectos. Así, los expertos cargan con sus propios sesgos y ellos influyen en el modo en que perciben el programa, en el juicio que se

forman acerca de su funcionamiento y en el informe de evaluación que presentan.

A pesar de sus debilidades, hay ocasiones y circunstancias en que este es el único tipo de evaluación posible, dada la premura con que se requiere disponer de una "opinión" acerca del programa, la baja disponibilidad de presupuesto para la evaluación u otras restricciones. En estos casos la evaluación de expertos aparece como una opción de evaluación recomendable, aunque es necesario tener presente sus limitaciones.

- *Informes de Administradores de Programas.* Rutinariamente los administradores de programas reportan sobre la marcha de sus programas y el cumplimiento de metas. En ocasiones las autoridades políticas pueden basar sus decisiones – acerca de expansión o reforma del programa, por ejemplo – en este tipo de informes. Sin embargo, las conclusiones de estos reportes no son altamente confiables, como instrumentos de evaluación. Es altamente improbable que un administrador de programas informe a sus superiores que el programa no ha logrado resultados o no ha tenido impactos significativos, dado que ello podría poner en riesgo su propia posición. Pero esta es la hipótesis normal de trabajo de toda evaluación adecuadamente desarrollada.

Así, los informes de Administradores de Programas deben ser vistos como reportes de gestión y no como resultados de una evaluación del programa. Esto llama, otra vez, la atención sobre la necesidad de disponer de evaluaciones independientes.

- *Percepción de Usuarios.* Aunque recabar antecedentes acerca de la opinión que tienen los usuarios sobre el funcionamiento del programa es muy útil para efectuar ajustes que se orienten a lograr mayor satisfacción de los beneficiarios, ello no puede ser entendido como una evaluación rigurosa que identifica los cambios que han ocurrido en la realidad producto de la intervención.

Por otro lado, es preciso indicar que los usuarios pueden desarrollar conductas estratégicas en sus respuestas. Si perciben que de sus respuestas se pueden derivar incrementos de beneficios, o evitar que éstos se reduzcan o eliminen, sus respuestas estarán en consonancia con sus intereses.

Adicionalmente, la pregunta clave para identificar el impacto no puede ser respondida, con objetividad, por los usuarios: ¿cómo estarían en la actualidad si no hubieran participado en el programa?

5. IDEAS CENTRALES DEL TEXTO

Al finalizar esta “exposición – revisión” de los métodos más comúnmente utilizados para evaluar programas sociales, es necesario retener tres ideas, por parte de quienes se están introduciendo a este apasionante campo.

1. Que la evaluación de programas es una actividad u oficio en la que se aplican métodos, procedimientos y técnicas de recolección de datos e información y de análisis, aceptados en las Ciencias Sociales, al estudio de una intervención, programa o proyecto que se implementa en una determinada realidad social.
2. Ello requiere que, en su trabajo, el evaluador se empeñe en:
 - a) Desarrollar la evaluación con rigor científico, para evitar sesgos conciente. Ello también requiere que el evaluador documente y describa con precisión cómo ha llevado la evaluación. Ello permitirá a los destinatarios primarios y directos del informe, a los interesados en el programa y a la comunidad de evaluadores y analistas de política pública advertir la fortaleza metodológica y la validez de las conclusiones del reporte de la evaluación.
 - b) Atender cuestiones prácticas, como las referidas al debate ético que se puede dar en torno a la implementación de un programa piloto – ya reseñado en el documento – o las referidas a la duración de la evaluación y costos que involucra

realizar una evaluación que cumpla los estándares requeridos en una investigación en Ciencias Sociales.

3. El evaluador debe comprender que está en el centro de una red de intereses, que el resultado de su trabajo es esperado con ansias por actores con intereses distintos y – con frecuencia – contrapuestos, y por la comunidad de evaluadores. Así, el rol del evaluador es el de un testigo experto que expone, ante la comunidad, los resultados de una evaluación conducida con rigor y basada en evidencia comprobable.

La finalidad de la evaluación de programas es verificar la utilidad, beneficio y contribución de cualquier intervención. En último término, las evaluaciones se hacen para verificar si un determinado programa o proyecto produjo, produce o producirá un mayor bienestar para los ciudadanos.

GLOSARIO

Análisis costo - beneficio: es un método de evaluación que busca identificar los beneficios que reporta o reportará un proyecto o programa, los costos en que se incurrirá para producir esos beneficios y la cuantía de la diferencia entre ambos, ambos expresados normalmente en dinero, con la finalidad de poder efectuar comparaciones entre distintas alternativas de la misma intervención o programas distintos. Ello permitirá identificar cual es el programa que genera mayores beneficios, descontados sus costos.

Análisis costo - efectividad: es un método de evaluación de programas que permite comparar los costos de un programa con los beneficios resultantes, no expresados en la misma medida.

Cobertura: identifica la proporción en que la población seleccionada para recibir los beneficios de un programa realmente es atendida por el programa.

Diseño cuasi - experimental: hace referencia a aquel diseño de evaluación en que la conformación de los grupos de tratamiento y comparación no es hecha al azar.

Diseño experimental: es aquel tipo de diseño de investigación de evaluación de impacto que distribuye la muestra, entre un grupo de tratamiento y otro de comparación, al azar.

Evaluación de diseño: es aquella que tiene como propósito estimar, a priori, si un proyecto o programa podrá abordar el problema social al que está dirigido con una razonable probabilidad de éxito.

Evaluación de eficiencia: provee una estimación acerca de la relación entre costos y beneficios o costos y resultados que un programa pretende alcanzar. Ello permite efectuar comparaciones entre diversos programas acerca de la eficiencia relativa de ellos.

Evaluación de impacto: es el tipo de evaluación que busca identificar efectos – deseados y no deseados – que se producen como consecuencia de la implementación de un programa o proyecto.

Evaluación de programas: es la aplicación sistemática de metodologías de investigación social para juzgar, analizar, y mejorar las formas en que programas y políticas sociales son conducidos, desde las etapas más tempranas de definición y diseño, hasta las etapas de desarrollo, implementación y obtención de resultados e impactos.

Evaluación de resultados: este tipo de evaluación se orienta a determinar si los logros que exhibe el programa son pertinentes y consistentes con los objetivos que se plantearon en la etapa de diseño de la intervención.

Evaluación ex - ante: es aquella evaluación que se lleva a cabo antes del inicio de la implementación y puesta en práctica de un programa o proyecto.

Evaluación ex – post: es aquel tipo de evaluación que se implementa una vez concluida la fase de implementación del programa o, al menos, concluida una etapa del programa.

Evaluación intermedia: es un tipo de evaluación que se desarrolla durante la ejecución del programa o proyecto y que se orienta a identificar si hay circunstancias que estén impidiendo que la intervención marche como se preveía o cumpla las metas que se habían presupuestado para los distintos momentos de la ejecución.

Focalización: se refiere a la identificación de aquellas personas que serán o no serán beneficiarias de un programa social.

Indicadores: es una unidad de medida que expresa normalmente una relación entre dos dimensiones y que representa la expresión cuantitativa de un cierto fenómeno o realidad que busca ser influido por la acción del proyecto o programa. Ello permite comparar el estado de la realidad identificada por el indicador en distintos momentos del tiempo, o con realidades próximas.

Monitoreo de la cobertura: busca determinar si el programa está llegando a su población objetivo.

Monitoreo de la implementación: busca establecer si en la ejecución real del programa han habido desviaciones respecto de las especificaciones establecidas en el diseño que estén teniendo incidencia en los logros del programa.

Problemas a considerar en la evaluación: se refiere a las circunstancias que, potencialmente, pueden generar sesgos en el desarrollo de los trabajos de evaluación de programas y en los resultados que éstos arrojen.

Tipos de Evaluación: los trabajos de evaluación de programas pueden ser clasificados según la etapa de implementación del programa en que la evaluación se lleva a cabo, la finalidad para la cual se hacen los trabajos de evaluación y la ubicación del evaluador respecto del programa que evalúa. Así, según la etapa del programa en que se hace la evaluación habrán evaluaciones ex - ante, evaluación intermedia y evaluación ex - post. Según la finalidad perseguida por la evaluación de programas se distinguen evaluación de diseño, evaluación de eficiencia, seguimiento o monitoreo, evaluación de resultados y evaluación de impacto. Si el énfasis está puesto en los instrumentos de evaluación se distinguen evaluaciones cuantitativas, evaluaciones cualitativas y evaluaciones mixtas. Y si la clasificación se hace considerando la perspectiva o ubicación del evaluador, los tipos de evaluaciones que se distinguen son las evaluaciones internas, externas, conjuntas y participativas.

Seguimiento y monitoreo: es un tipo de evaluación que se desarrolla durante la ejecución del proyecto o programa y que tiene por finalidad proveer información de retroalimentación sobre los avances y problemas detectados durante la implementación.

REFERENCIAS

Baker, Judith (2000), "Evaluación del impacto de los proyectos de desarrollo en la pobreza. Manual para profesionales". Washington DC: Banco Mundial.

Budnick, Frank S. (1993), "Applied Mathematics for Business, Economics and the Social Sciences". USA: McGraw – Hill.

CEPAL (2007), "Capítulo VII. Término del Proyecto y Evaluación de Resultados", Serie Manuales N° 7, Material Docente sobre Gestión y Control de Proyectos – Programa de Capacitación BID/ILPES. Santiago, Chile: CEPAL. Sitio web visitado Marzo 2007 (<http://www.eclac.cl/publicaciones/xml/2/5542/lcl1321eCap7.pdf>)

CESC, Universidad de Chile (2005), "Los qué cómo y cuando de la evaluación: recomendaciones para el diseño de evaluaciones de proyecto de prevención comunitaria del delito", en Boletín +Comunidad +Prevención. Santiago, Chile: CESC, Octubre 2005.

Cohen, Ernesto y Rolando Franco (2000), "Evaluación de Proyectos Sociales" 5ta Ed. México: Siglo XXI Editores.

Cruz, José Miguel (2003), "Violencia y Democratización en Centroamérica: el impacto del Crimen en la legitimidad de los regímenes de posguerra"; en América Latina Hoy, año/vol 35, Universidad de Salamanca, España; pp. 19 – 59.

FMI, Fondo Monetario Internacional (2006), "Ficha Técnica para le Crecimiento y la Lucha contra la Pobreza (SCLP)", Agosto 2006, en sitio web (visitado en marzo 2007):

<http://www.imf.org/external/np/exr/facts/spa/prgfs.htm>

FOSIS (1995), "Evaluación de Impacto del FOSIS (1991 – 1994), Segundo Borrador. Informe del Panel de Expertos. Santiago, Chile: FOSIS, 24 de Mayo de 1995.

Frechtling, Joy (2002) "The 2002 User – Fiendly Handbook for Project Evaluation", Division of Research, Evaluation and Communication of the National Science Foundation. USA: NSF.

Grosh, Margaret (1994), "Administering Targeted Social Programs in Latin America. From plenitude to practice", World Bank Regional and Sectoral Studies. Washington DC: The World Bank.

Jansson, Axel (1994), "Preparación y Evaluación Privada y Social de Proyectos de Salud". Santiago: PIAS, Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

MIDEPLAN (2007), "La Evaluación de Resultados en el SIN", en www.sni.mideplan.cl. Sitio web visitado en Marzo 2007.

Rossi, Peter y Howard Freeman (1993), "Evaluation. A Systematic Approach" 5th Ed. California: SAGE.

Stock James y Mark Watson (2007), "Introduction to Econometrics", Second Edition. USA: Pearson – Addison Wesley.

Valadez, Joseph y Michael Bamberger (1994), "Monitoring and Evaluating Social Programs in Developing Countries. A Handbook for Policymakers, Managers, and Researchers". Washington DC: The World Bank.