

La evaluación de programas sociales: definición y diseño

Laura Guzmán¹

SECCION I

DEFINIENDO LA INVESTIGACIÓN EVALUATIVA

1.1 INTRODUCCION

La evaluación de programas sociales o investigación evaluativa (usaremos ambos términos indistintamente) es una actividad relativamente reciente. Si bien podemos identificar esfuerzos muy tempranos por establecer si determinadas acciones y programas servían para reducir la incidencia de ciertos problemas sociales, no es sino hasta después de la Segunda Guerra Mundial cuando la evaluación de programas recibe un mayor impulso y adquiere una fisonomía propia como actividad investigativa.

Es durante los años 50's cuando se comienza a desarrollar programas sociales a gran escala en los Estados Unidos, Europa y otras naciones industrializadas, así como programas de desarrollo en los países del Tercer Mundo (Rossi & Freeman, 1985). Por ejemplo, organismos internacionales y gobiernos en Asia, África y América Latina impulsaron en estas regiones programas de planificación familiar, nutrición, salud pública y desarrollo comunal, que tenían componentes evaluativos tendientes a medir cambios como resultado de las acciones del programa (Levine et al., 1981; Freeman et al., 1980).

La investigación evaluativa prosperó muchísimo durante las décadas del 50 y 60, principalmente por el desarrollo de métodos de investigación y estadística aplicables al estudio de los problemas sociales, procesos sociales y relaciones interpersonales. Al mismo tiempo, la necesidad de producir métodos cada vez más sofisticados para evaluar programas sociales y los avances en el campo de la computación estimularon el desarrollo de metodologías de investigación aplicables a las necesidades de quienes necesitaban evaluar la eficacia y eficiencia de programas sociales, especialmente para la recolección sistemática de datos mediante procedimientos de encuesta y manejo analítico de grandes cantidades de variables mediante estadísticas multivariadas (Freeman & Rossi, 1985).

Los años 70 se caracterizaron por ser una época de cambios en la orientación de los programas sociales y naturaleza de la evaluación. Durante esta época se comenzaron a producir serios cuestionamientos acerca de la conveniencia de continuar expandiendo los programas gubernamentales. Diversos sectores comenzaron a presionar para que las instituciones documentaran, cada vez con mayor intensidad, cómo gastan sus recursos, la efectividad de los servicios y la conveniencia social de continuar prestando estos servicios a determinadas poblaciones (Austin et al., 1982; Freeman & Rossi, 1985).

¹ Profesora Escuela de Trabajo Social, Universidad de Costa Rica, 1991.

La evaluación de programas se convirtió , entonces, en una instrumento para determinar la necesidad de ciertas políticas y programas, controlar el desarrollo de las acciones programadas y los resultados obtenidos, y para producir información oportuna para administradores, planificadores y otras personas que toman decisiones en el gobierno, fundaciones e instituciones responsables de la prestación de servicios sociales para decidir cuáles actividades deben realizar y apoyar.

Esta tendencia se incrementó durante los 80's. Políticas neoliberales han llevado a una privatización cada vez mayor de importantes servicios sociales. El financiamiento de programas sociales por parte de organismos internacionales y gobiernos nacionales es cada vez más limitado, obligando a juntas directivas, administradores y personal técnico-profesional a dedicar mayores esfuerzos a documentar el impacto social de programas en curso y la necesidad de otros, buscar nuevas fuentes de financiamiento que requieren mecanismos de control cada vez más estrictos, y a reducir los "experimentos sociales".

Para el Trabajo social, la evaluación se ha convertido en una actividad necesaria y fundamental para determinar la efectividad y eficiencia de su quehacer cotidiano, corregir oportunamente problemas y cambiar el curso de programas y servicios sociales. Los constantes recortes presupuestarios y las crecientes críticas de diversos sectores acerca de la efectividad y eficiencia de muchas intervenciones sociales han llevado a que muchos profesionales en Trabajo Social hayan tomado conciencia de la importancia de garantizar que los servicios que prestan respondan a las necesidades de la clientela a la que sirven y que la calidad de los mismos sea la mejor. Pero más importante aún, las experiencias de los últimos años le han permitido a quienes trabajan en la organización, administración y prestación de servicios sociales desarrollar una perspectiva más realista acerca de los factores que obstaculizan una ejecución exitosa en muchos programas sociales y acerca de la magnitud de los resultados que se pueden esperar de éstos.

Los limitados recursos y este mayor realismo han contribuido a incrementar la necesidad de realizar evaluaciones que permitan maximizar esfuerzos para manejar la multiplicidad de problemas sociales y humanos que todavía persisten. Esta preocupación, así como las demandas por distinguir programas útiles de otros que son ineficientes e infectivos y para planear, diseñar y ejecutar esfuerzos innovadores que tengan un impacto significativo en la calidad de vida de los grupos a los que sirven, exigen que aquellos que están involucrados en la toma de decisiones, financiamiento, planeamiento y ejecución de los programas generen respuestas a varias interrogantes, entre ellas:

- ❖Cuál es la naturaleza y Ámbito del problema social que requiere de una intervención?
- ❖Qué intervenciones podrían realizarse para reducir significativamente el problema?
- ❖Están llegando los servicios al grupo-meta?
- ❖Se está efectuando la intervención de la forma como se planeó?

- ❖ Está siendo efectiva?
- ❖ Cuánto cuesta?
- ❖Cuál es su costo con relación a su efectividad y beneficios?

La respuesta a estas preguntas es necesaria de parte de trabajadores sociales y otros científicos sociales involucrados en la prestación de servicios, no sólo relacionados con programas de gran envergadura sino también en aquellos que tienen una incidencia local o especializada. La capacidad para responder a estas interrogantes se convierte, entonces, en el meollo de la investigación evaluativa.

1.2 QUE ENTENDEMOS POR “EVALUACIÓN DE PROGRAMAS”

La evaluación de programas, llamada también investigación evaluativa es la aplicación sistemática de los métodos y procedimientos de investigación social para:

- a) Identificar problemas en la concepción (conceptualización), estructura (diseño) y puesta en marcha (ejecución) de programas sociales; y
- b) Determinar si un programa está cumpliendo con (alcanzando) los objetivos propuestos, los efectos que está teniendo en la población que recibe los servicios y si está funcionando adecuadamente (como se esperaba cuando el programa se planeó).

El fundamento esencial de la evaluación reside en la comparación, pues los parámetros usados para medir la actividad e impacto del programa se comparan con pautas y pruebas empíricas, o bien, con programas similares para establecer el éxito relativo del programa que está siendo evaluado. Por ello, se propone contribuir a la solución de problemas mediante la determinación de cuales estrategias programáticas funcionan mejor y bajo cuales condiciones. La evaluación trata de responder a la pregunta: Qué fue lo que ocurrió como resultado de esta estrategia de intervención que no hubiera ocurrido en ausencia del programa? (Moroney & Grub, 1978).

En otras palabras, la evaluación supone el uso de métodos y técnicas de investigación con el propósito de valorar y mejorar el planeamiento, control, efectividad y eficiencia de los programas sociales. Pero es mucho más que una mera aplicación de metodologías de investigación. Es una actividad política y administrativa que provee insumos para la toma de decisiones, asignación de recursos y planificación, diseño, ejecución y continuación de programas para mejorar las condiciones de vida de diversas poblaciones (Freeman & Rossi, 1985; Raymond, 1985).

En los niveles de formulación y administración de la política y programas sociales, la evaluación se ocupa de la identificar, recolectar, procesar, analizar y diseminar información sobre los programas, información que necesitan quienes tienen que tomar decisiones (Moroney & Grub, 1978). Estas personas pueden ser agentes externos a la institución o parte de ella, dependiendo de agentes externos a la institución o parte de ella, dependiendo de según se relacionen con el programa. La utilidad de la evaluación de un programa estaría

condicionada, entonces, por la capacidad que ésta demuestre para proveer a directores (as) y administradores (as) de información oportuna para mejorar el funcionamiento institucional y en particular, el de los servicios sociales.

Según Bigman (1961), la evaluación de programas tiene seis propósitos fundamentales:

- a) Descubrir si los objetivos del programa se están cumpliendo y cómo;
- b) Identificar éxitos y fracasos logrados por el programa;
- c) Explicar los factores que determinan el éxito o fracaso de un programa;
- d) Orientar el desarrollo de experimento en los que se apliquen técnicas para aumentar la efectividad de determinadas alternativas (modalidades, estrategias, métodos) de intervención;
- e) Sentar las bases para investigaciones evaluativas futuras y establecer las razones de por qué ciertas técnicas de intervención social son más efectivas que otras;
- f) Redefinir los medios que deben emplearse para alcanzar objetivos y replantear (redefinir, reorientar) metas a la luz de los resultados de la investigación.

1.3 LA PRACTICA DE LA INVESTIGACIÓN EVALUATIVA: PROBLEMAS Y PERSPECTIVAS

EL problema práctico más serio que enfrenta un(a) evaluador(a) que se propone aplicar los procedimientos de la investigación social a la evaluación de programas sociales, es el de que las evaluaciones deben realizarse en un ambiente que cambia constantemente por razones políticas, financieras y sociales, que no puede ser controlado por quienes estudian los programas. Algunas de estas características son (Freeman & Rossi, 1985):

- a) Cambios en la influencia política, recursos y prioridades de quienes tienen a cargo los programas sociales.
- b) Cambios en los intereses e influencia de los diferentes grupos de presión.
- c) Modificaciones en las prioridades y responsabilidades de las organizaciones e instituciones que se encuentran implementando programas.
- d) Problemas no previstos durante la intervención o la prestación de los servicios sociales, que obligan a cambiar el curso de la intervención.
- e) Problemas no previstos con el diseño de la evaluación.

El diseño de evaluaciones deben tomar en cuenta consideraciones de tipo práctico y científico. La evaluación supone la aplicación del método científico, pero busca dar respuesta a problemas prácticos que enfrentan los programas sociales con el propósito de quienes los ejecutan, administran o dirigen puedan tomar decisiones acertadas y oportunas.

Esta doble condición ha planteado serias controversias acerca de si la evaluación es una actividad cualitativamente diferente de la investigación científica, y hasta dónde quien evalúa debe preocuparse de realizar su trabajo

aplicando diseños de investigación. La polémica más ilustrativa es tal vez la planteada por Campbell y Cronbrach acerca de la investigación evaluativa.

Donald Campbell (1969) efectuó un extenso análisis sobre las aplicación de enfoques experimentales al estudio del impacto de los programas sociales. Como resultado, llegó a desarrollar una perspectiva ideológica acerca de la investigación evaluativa que se puede resumir en los siguientes términos. Las decisiones en cuanto a políticas y programas sociales tienen que generarse a partir de la prueba continua de modelos para mejorar las condiciones de vida de las personas. Por esta razón, los esfuerzos de cambio social deben apoyarse y fundamentarse en la experimentación, siendo la tecnología de la investigación social el medio que posibilita esto. Con base en estos argumentos, Campbell trató de imponer el método experimental en la investigación evaluativa.

En contraposición, Lee Cronbach (1982), estadístico y respetado investigador, opinó que la investigación evaluativa es un arte a diferencia de la investigación científica, en donde cada evaluación representa un esfuerzo muy particular de proveer a quienes tienen que tomar decisiones con información útil y oportuna. Ello contrasta con las investigaciones científicas cuyo diseño y ejecución persiguen cumplir con los requisitos establecidos por la comunidad científica. Para Cronbach, lo que diferencia a una de la otra son sus propósitos y objetivos; aun cuando ambas recurran a la lógica y procedimiento de la investigación, los estudios científicos buscan cumplir con un conjunto de reglas, mientras que las evaluaciones tienen que realizarse de manera que respondan a los intereses de las políticas y programas formulados y ejecutados para que sean de utilidad para tomar decisiones de acuerdo a recursos disponibles, entorno político y limitaciones.

Si analizamos detenidamente ambas posiciones, encontraremos que las divergencias no son tan abismales, ya que ambos están concibiendo como investigación científica sólo aquella que siga las reglas y procedimientos del positivismo lógico. Sabemos hoy día que no es posible –ni siquiera para quienes realizan investigaciones bajo esta concepción de ciencia- abstraerse de las limitaciones que enfrentan quienes hacen “investigación aplicada”. Es posible llevar a cabo evaluaciones rigurosas en el ámbito institucional, pese a los problemas que enumeramos anteriormente y sin necesidad de caer en posiciones extremas que consideran lo científico como sinónimo del modelo experimental.

Boruch et al. (1978) lograron, por otra parte, inventariar un gran número de “verdaderos” experimentos llevados a cabo como evaluaciones, demostrando con ello que este tipo de investigación puede también efectuarse cuando existen las condiciones para ello. No obstante, debemos enfatizar que no siempre encontraremos ni el ambiente ni las condiciones –mucho menos en el medio institucional latinoamericano- para llevar a cabo evaluaciones de este tipo.

1.4 ¿CUANDO SE REQUIERE HACER UNA EVALUACIÓN?

La investigación evaluativa se preocupa de medir la capacidad de un programa social para cumplir con los objetivos que le fueron definidos, así como también del por qué este fracasa o logra alcanzar estos objetivos, aportándole a administradores, planificadores y técnicos información necesaria para controlar el desarrollo de los programas (monitoreo), reestructurarlos y hacerlos más efectivos (Weiss, 1972; Raymond, 1985).

Las evaluaciones son muy útiles cuando la institución necesita identificar formas de cómo manejar los servicios y metodologías utilizadas en su prestación o cuando requiere probar ideas innovadoras acerca de cómo resolver determinados problemas sociales. Las evaluaciones son útiles también cuando queremos constatar los resultados del programa, decidir si se amplía, reduce su ámbito de intervención o elimina (Freeman & Rossi, 1985).

Es muy importante que las y los trabajadores sociales veamos la investigación evaluativa como una actividad necesaria en cualquier programa o sistema de prestación de servicios. No se requiere efectuar complejos y largos estudios para que ésta rinda sus frutos; tampoco lograremos dilucidar todo lo que ocurre con el programa o lo que debemos cambiar o corregir. Pero sí nos permitirá reducir incertidumbres en varias áreas.

Una pregunta que mucho nos hacemos es, cuándo hacer una evaluación. Cuando evaluar es una decisión que dependerá de un conjunto de factores, algunos de los cuales serán analizados en esta sección.

Las evaluaciones se realizan por tres razones fundamentales:

1.4.1 Como punto de partida para la conceptualización y diseño de intervenciones sociales (formulación de políticas y programas). Este tipo de evaluaciones tiene carácter diagnóstico y las definimos como “diagnósticos situacionales” (Pichardo, 1986).

Las intervenciones sociales son respuestas a problemas colectivos que han sido reconocidos como “cuestiones sociales” y que como tales, requieren de una acción intencionada, organizada y colectiva para su solución.

La evaluación diagnóstica debe dar respuesta a por lo menos estas tres preguntas (Freeman & Rossi, 1985):

- ❖ ¿Cuál es la magnitud y distribución del problema objeto de intervención y/o la población afectada por éste?
- ❖ ¿Se ha diseñado el programa de acuerdo a los objetivos propuestos? Existe una coherencia entre las necesidades de los grupos afectados y los objetivos del programa? ¿Se han maximizado las posibilidades de éxito del programa?
- ❖ ¿Cuáles son los costos actuales y proyectados y cuál es su relación con los beneficios y su efectividad?

1.4.2 Para medir los resultados de la ejecución de un programa. Esto es lo que algunos llaman “monitoreo” o control de la respuesta de los servicios que presta la institución.

Son varias las razones por las que se realiza este tipo de control en un programa en ejecución.

En primer término; una dirección y administración adecuada de los servicios sociales debe garantizar que las actividades se realicen de la manera más eficiente. Aquellos(as) funcionarios(as) utilicen inapropiadamente los recursos financieros y humanos asignados al programa, arriesgan su posición y el futuro del programa.

Segundo, quienes auspician el programa requieren evidencias de que los dineros se están gastando bien. Los datos provenientes de evaluaciones que se realizan regularmente constituyen una de las armas más importantes que puede emplear quien dirige un programa para documentar la efectividad de éste, justificar cómo se emplea el personal, solicitar más recursos o defender el desempeño de su programa en contraste con otros similares.

Por otra parte, existen muchos programas que no se ejecutan de acuerdo a su diseño original o no están llegando a la población meta. En estos casos, una evaluación de “monitoreo” o control puede alertarnos sobre estos problemas.

Este tipo de evaluaciones tratan de responder a dos preguntas:

- ❖ Está llegando el programa a la población meta o área objeto de atención?
- ❖ Se está llevando a cabo la intervención según como se especifico en el diseño del programa?

1.4.3 Para determinar la utilidad del programa. La utilidad de un programa se puede determinar de acuerdo al tipo de preguntas que se intente responder en la evaluación. El énfasis de tales interrogantes puede centrarse en el esfuerzo, la efectividad o impacto o la eficiencia.

a) Definiendo esfuerzo: El esfuerzo tiene que ver con los insumos asignados a un programa, esto es, la cantidad y tipo de actividades que se emplean para lograr los resultados (Austin, 1982). Hace referencia a las actividades de y al tiempo empleado por el personal, la asignación y empleo de recursos materiales (v.g espacio de oficina, equipo, etc.).

La medición del esfuerzo de un programa intenta responder a las preguntas siguientes (Austin, 1982):

- ❖ Qué y cuánto está haciendo la institución a través del programa para lograr los objetivos?
 - A cuánto asciende la clientela atendida;
 - El tiempo dedicado por el personal a cada una de las actividades programadas, y

- El numero de servicios prestados (por persona) por mes?

Este tipo de información no ilustra sobre la calidad de los servicios, pero sí indica que algo está ocurriendo en el programa.

b) Definiendo efectividad: Las instituciones generalmente buscan obtener información que vaya más allá del esfuerzo de un programa. También desean conocer sobre la efectividad o impacto de este. Por ejemplo, han logrado los Hogares CREA recuperar permanentemente a los drogadictos inscritos en su programa? Está logrando el Programa de Hospital de día prevenir el deterioro físico, mental y social de los ancianos?

Es sumamente difícil defender la implementación y continuación de un programa a menos de que demostremos su impacto. De allí la importancia de evaluaciones de impacto. No obstante, estos resultados deben evaluarse contra los costos de los insumos, ya que por lo general se tiende a no apoyar programas muy caros con relación a sus resultados, independientemente de la necesidad de los servicios. Así por ejemplo, diversas universidades han eliminado sus programas de orientación estudiantil, pues el costo de los mismos es muy elevado frente a los limitados beneficios que parecen obtener los estudiantes.

La evaluación de impacto es la medición del grado en que un programa genera cambios en la dirección deseada o prevista. Contiene un conjunto de objetivos y criterios de éxito claramente especificados y definidos operacionalmente. UN PROGRAMA QUE TIENE IMPACTO ES AQUEL QUE TIENE ALGUN MOVIMIENTO O CAMBIO HACIA LOS OBJETIVOS QUE ORIENTAN LA INTERVENCIÓN (Freeman & Rossi, 1985).

El diseño y conducción de una evaluación de impacto requiere de un plan para la recolección de la información que sea riguroso para demostrar de manera precisa y sistemática que los cambios se han generado como resultado de la intervención y no por otros factores (Austin, 1982; Freeman & Rossi, 1985).

Las evaluaciones de impacto o efectividad tratan de responder a las preguntas siguientes:

- ❖ Ha logrado el programa responder a los objetivos propuestos?
- ❖ Pueden los resultados explicarse por algún proceso alternativo o no previsto por el programa?

Una evaluación de impacto recurre, por lo general, a diseños experimentales o cuasi-experimentales para determinar la efectividad del programa.

El diseño experimental es un plan para conducir un experimento, en el que unidades experimentales (v.g. personas, comunidades, grupos) reciben un "tratamiento" (v.g. asignadas a un programa social o son expuestas a innovaciones educativas) o en otros casos, no reciben tratamiento alguno. En el primer caso, estas unidades constituyen el 'grupo experimental' y en el segundo el 'control'. Este plan incluye la estrategia y procedimientos relacionados con cómo asignar cada unidad a cada uno de estos grupos y cómo medir las respuestas de estas unidades al tratamiento (Riecken &

Boruch, 1974). En un capítulo posterior discutiremos algunas técnicas utilizadas para hacer las asignaciones de sujetos a cada grupo, cómo hacer el análisis de los datos recogidos durante el experimento, cómo analizar estos resultados y algunos de los problemas más difíciles que enfrentan los(as) trabajadores(as) sociales cuando recurren a este tipo de diseños.

Los diseños cuasi-experimentales se emplean cuando no es posible asignar al azar unidades experimentales para producir los grupos experimental y control requeridos en un verdadero experimento. Cuando este es el caso, existen sin embargo, modos de organizar tratamientos y mediciones y modalidades para analizar los datos que son útiles para evaluar impacto. Estos son los modelos cuasi-experimentales. Estos se distinguen de los verdaderos experimentos en que no hacen una asignación al azar de las unidades a los tratamientos (programa o intervención) (Riecken & Boruch, 1974). Este tipo de diseño es particularmente útil cuando no existen las condiciones para asignar unidades al azar. Este tema será tratado en un capítulo posterior en mayor detalle.

Indistintamente del tipo de diseño que se utilice para la evaluación, es particularmente importante el empleo de comparaciones, ya que nos proponemos evaluar la ocurrencia de cambios como resultado de la intervención. Por ejemplo, se requiere comparar el comportamiento de cada cliente antes y después del tratamiento, las tasas de delincuencia se pueden medir antes, durante y después de realizar un programa preventivo; así también, los participantes en talleres para líderes juveniles podrían ser observados manejando grupos antes y después de la capacitación.

c) Definiendo eficiencia: La eficiencia se refiere a la relación entre costos y beneficios. La evaluación de eficiencia se ha convertido en una actividad importante debido a que los recursos disponibles para programas sociales son cada vez más escasos, debiendo las instituciones competir por ellos. Constantemente deben tomarse decisiones relativas a continuar o discontinuar un programa, financiarlo o no y ampliarlo o reducirlo con respecto a otro.

Este tipo de evaluación trata de responder a preguntas tales como:

- ❖ Cuánto cuesta la prestación de servicios y el otorgamiento de beneficios por participante (el costo por participante según el resultado del servicio)?
- ❖ Hace el programa un uso eficiente de los recursos en comparación con otras formas alternativas de empleo de los recursos?

El tipo de técnicas que se emplean para hacer evaluaciones que respondan a estas preguntas son de dos tipos: análisis de costo-beneficio y análisis de costo-efectividad (Freeman & Rossi, 1985).

Las preguntas sobre eficiencia tiene que ver con costos en dinero, pero también con el uso del tiempo, equipo e infraestructura. Por ejemplo, podemos reducir el tiempo que dedicamos a entrevistas de ingreso? Cómo hacer un mejor uso del espacio de oficina disponible? Cómo redefinir las funciones del personal de manera que se logren los objetivos del programa? Hay duplicación de esfuerzos?

Una interrogante relacionada con la eficiencia intenta obtener información acerca de la relación entre los insumos del programa (esfuerzo) y los resultados de éste (efectividad) (Austin, 1982). Generalmente, se plantean evaluaciones de eficiencia cuando los recursos son limitados y deben limitarse o reorganizarse los servicios. En estos casos, se tiende a enfatizar en la prestación de servicios a grupos antes que a individuos o privatizando los servicios sociales.

1.4.4 Las evaluaciones se pueden efectuar también según la etapa en que se encuentre el programa:

a) Programas innovadores: Este tipo de programas es poco común. Según Freeman & Rossi (1985), un programa es innovador cuando no ha sido ejecutado y evaluado de alguna de estas formas:

- La intervención está apenas en proceso de desarrollo, esto es, no se cuenta aun con la suficiente evidencia de que tiene impacto como programa.
- El sistema de prestación de servicios o partes de éste no han sido aprobados todavía. Un programa de este tipo sería la idea de poner a estudiantes de secundaria a impartir información y educación nutricional para ancianos.
- La población meta es nueva o ha sido ampliada. Una intervención de esta índole sería el ofrecer a padres de niños con retardo mental entrenamiento mediante medios audiovisuales y una especialista que se desplaza una vez a la semana a la escuela. Este servicio se ofrece como alternativa para reducir costos de transporte a quienes viven en el perímetro del Área Metropolitana.

b) “Afinamiento” de programas en ejecución: A veces es conveniente evaluar cambios en un programa en marcha para mejorar su eficacia o eficiencia, en otras palabras, para aumentar la magnitud del impacto o para reducir los costos por unidad de impacto. Veamos un ejemplo. El personal de un programa que ofrece tutorías semanales a niños pobres ha constatado que a través de éste se han mejorado algunas habilidades educativas en los niños, pero consideran que el programa sería mucho más efectivo si se les ofrecieran otras oportunidades de enseñanza que fueran más allá de una tutoría semanal. En consecuencia, se introduce y evalúa un programa de “tareas” o asignaciones. Otro ejemplo lo constituye un programa de terapia diaria para alcohólicos que estuvieron hospitalizados por desintoxicación, que debido a los altos costos del tratamiento ambulatorio, se reduce el periodo de sesiones diarias de tres meses a dos. Se plantea, entonces, una evaluación para determinar si este cambio permite alcanzar los mismos resultados que con la modalidad anterior.

Las razones para “afinar” un programa es marcha pueden ser diversas. Además, a veces es difícil establecer una línea divisoria entre lo que podemos considerar innovaciones o ajustes para mejorar la eficacia y eficiencia de un programa. Algunas veces, estos cambios pueden resultar muy costosos y generar consecuencias inconvenientes, como por ejemplo, alterar la calidad del servicio que se le presta a la población beneficiaria. Por estos motivos, todo esfuerzo en esta línea debe ser analizado previo a ejecutar los cambios y la

evaluación para determinar posibles implicaciones para todos los grupos involucrados (población meta, personal, dirección, etc.).

c) Evaluación de programas consolidados: Los programas que ya tienen algún tiempo de funcionamiento, inclusive años, bien pueden ser sujetos de evaluación. Las razones son varias. Diversos sectores podrían estar interesados en conocer sobre su impacto o el funcionamiento del sistema de prestación de servicios, así como la relación costo-beneficio para justificar su continuación, expansión o conclusión. Pueden afectar también, cambios en la disponibilidad de recursos, la orientación política de la Administración de turno o una reducción importante de la población afectada por el problema que motivó la intervención.

1.4.5 Algunos principios para recordar: Conviene recordar algunos principios básicos a considerar cuando nos encontramos en la etapa de seleccionar una alternativa de evaluación (Austin, 1982).

- ❖ El tipo de preguntas que orienten el problema de evaluación deben ser evaluables, en otras palabras, posibilitar la recolección de información sobre el problema en estudio. Hay preguntas muy generales o que generan mucha susceptibilidad (v.g. cambios en el comportamiento sexual de las personas como resultado de nuestra intervención) que dificultan la obtención de datos significativos.
- ❖ Debe existir más de una respuesta posible al problema, de manera que los informantes se sientan libres de expresar sus puntos de vista sin tener que ceñirse a alternativas de respuesta muy limitadas. Por ejemplo, si le preguntamos a una mujer que participa en un programa de internas en libertad condicional que tan satisfecha se siente con los servicios que le ofrece Trabajo Social, nos daría una respuesta que posiblemente no prueba nada, pues no dirá que sienta que pueda afectarle su condición de libertad.
- ❖ Quienes tienen que tomar decisiones con respecto al problema pueden querer y necesitar más información de la que se encuentra disponible. Esto debe valorarse con ellos(as), pues si consideran que hay suficientes datos o que pueden depender en mecanismos informales de retroalimentación que han venido funcionando en la institución, la evaluación sería una pérdida de tiempo.
- ❖ Quienes están involucrados en la toma de decisiones deben desear información para ellos que les sea útil para la toma de decisiones, y no solamente para satisfacer demandas externas.
- ❖ Es necesario que estas personas puedan articular cómo usarían los resultados de cualquier evaluación para garantizarnos que ésta no se quedará descansando en la gaveta de algún escritorio.

1.5 COMO SE USAN LAS EVALUACIONES?

Así como los objetivos y diseño de una evaluación deben considerar el estado de desarrollo del programa y las demandas que hacen los diversos sectores involucrados, es conveniente considerar también el uso que se le dará. En términos generales, podemos afirmar que la utilidad fundamental de una evaluación radica en la información que aporta para la toma de decisiones a diferentes niveles. Freeman y Rossi (1985) nos ofrecen algunas propuestas en este sentido.

1.5.1 Para tomar decisiones definitivas respecto a la aprobación de un programa o su terminación. Es poco probable que este tipo de decisiones se tomen exclusivamente con base en los resultados de una evaluación; sin embargo, si fuese elaborada con rigurosidad y sus resultados son contundentes, bien podría jugar un importante papel en la decisión.

1.5.2 Para desarrollar una racionalidad que oriente las acciones. Las evaluaciones pueden influenciar consideraciones de tipo político, práctico y relacionadas con recursos a tomar en cuenta por quienes tienen que tomar decisiones técnicas, legales o políticas sobre un programa. Por ejemplo, la presión por desinstitucionalizar niños abandonados genera una controversia sobre aspectos políticos, legales y técnicos. La puesta en marcha del primer programa de este tipo podría ir acompañado de una evaluación que determine las consecuencias de la desinstitucionalización en los niños, sus familias y los miembros de las comunidades en donde aquellos residirían. Los resultados podrían tener importantes implicaciones en la legislación futura, así como en el tipo de decisiones legales que tomen los organismos responsables por ley.

1.5.3 Legitimación y demostración de que el programa responde. Las evaluaciones pueden servir tanto a quienes defienden como a quienes adversan un programa. Datos acerca de la calidad de las intervenciones, su extensión, impacto y costos pueden ser utilizados por amigos y enemigos.

La legitimación de un programa puede ser necesaria en algún momento de su desarrollo. Por ejemplo, el Ministro de Salud podría manifestar su preocupación respecto a si los centros de atención infantil están llegando al mayor número de niños que necesitan los servicios y operando con costos muy elevados. La administración del programa podría emplear informes mensuales con esta información para evaluar el alcance de y eficiencia en la prestación de los servicios.

Este tipo de evaluaciones de evaluaciones pueden ser útiles también para alertar a directores y administradores sobre los “puntos débiles” en los programas que puedan ameritar modificaciones, expansiones o limitaciones en la intervención.

1.5.4 Estudios para el análisis administrativo y de políticas.

Muchas de los estudios efectuados en el campo han sido sobre evaluaciones de políticas.

Las evaluaciones de políticas se refieren al impacto de determinada política en grandes sectores de la población. Esta cualidad hace que puedan generar cambios mayores en la organización y actividades de grupos que están prestando servicios sociales y jugar un papel fundamental en la asignación de recursos financieros, personal y otros recursos.

La evaluación administrativa focaliza, en cambio, en aquellos componentes relacionados con el manejo del programa, uso del tiempo y recursos por parte del personal, efectos de cambios burocráticos u otros en esta línea.

Las diferencias entre estudios de política o administrativos dependerán, más que todo, del enfoque y énfasis que imponga la persona o equipo realizando la evaluación.

1.6 RESUMEN DE CONCEPTOS UTILIZADOS EN ESTA SECCION

1.6.1 Evaluación comprensiva: Es el análisis que abarca la conceptualización y diseño de las intervenciones sociales, el control o monitoreo de la ejecución del programa y la constatación de la utilidad de éste.

1.6.2 Análisis de la conceptualización y diseño del programa: Se refiere a estudios sobre: (a) la extensión y localización del problema de intervención, (b) formas en cómo la población meta se puede definir operacionalmente, y (c) si la intervención propuesta es apropiada.

1.6.3 Análisis de costo-beneficio: Estudios sobre la relación entre costos y resultados del proyecto, expresados generalmente en términos monetarios.

1.6.4 Análisis de costo-efectividad: Estudios sobre la relación entre los costos y resultados del proyecto, expresados generalmente en costos por unidad de resultado alcanzado.

1.6.5 Sistema de prestación de servicios: Se refiere a la organización del personal, procedimientos y actividades, infraestructura y materiales necesarios para ofrecer servicios a través de un programa

1.6.6 Investigación formativa: Es el diseño y desarrollo de pruebas para maximizar el éxito de una intervención.

1.6.7 Evaluación de impacto: Es la medición o determinación del grado en que un programa genera cambios en la población meta en la dirección deseada o prevista.

1.6.8 Intervención: El esfuerzo diseñado o realizado por un programa para producir los cambios esperados en la población meta.

1.6.9 "Monitoreo" o control: Constatación de si una intervención está: (a) operando de conformidad con su diseño, y (b) llegando a la población meta definida en el programa.

1.6.10 Evaluación de la utilidad del programa: Es el estudio de la efectividad (impacto) y eficiencia (relación costo-beneficio o costo-efectividad) en el programa.

1.6.11 Población meta: Son las personas, familias, organizaciones, comunidades o cualquier otra unidad hacia la cual se dirige la intervención.

1.6.12 Problema de intervención: Condiciones, deficiencias o carencias hacia dónde se dirige la intervención.

SECCION II

DISEÑANDO EVALUACIONES

2.1 INTRODUCCION

En la sección anterior definimos la investigación evaluativa. Nos toca ahora responder a la pregunta: COMO SE DISEÑA UNA EVALUACIÓN?

En esta sección nos ocuparemos de introducirlos(as) al proceso general de la evaluación de programas, destacando los componentes principales de cada una de las principales fases del proceso. Dado que existen diversos modelos de evaluación que pueden emplearse según el estado de desarrollo de la política o programa que deseamos estudiar o el tipo de información disponible acerca del programa, continuaremos con una descripción de algunos de los modelos o diseños de mayor utilidad para quienes tienen a cargo programas de Trabajo Social o están involucrados en la prestación de servicios sociales de diversa índole.

2.2 EL PROCESO DE LA EVALUACIÓN DE PROGRAMAS O INVESTIGACIÓN EVALUATIVA

Diversos autores(as) han tratado de identificar las fases en el proceso de investigación evaluativa (Austin, 1982; Raymond, 1985). Con el propósito de guiar a quienes se proponen diseñar y ejecutar estudios en este campo, hemos sintetizado estos aportes en seis fases, a saber:

- 1) Identificación de los objetivos del programa;
- 2) Determinación de las medidas de resultado (variables dependientes);
- 3) Identificación de las variables independientes e intervinientes;
- 4) Definiendo el diseño de investigación;
- 5) Diseño de los procedimientos para recoger la información; y
- 6) Recolección, análisis e interpretación de los datos.

Procederemos ahora a describir los elementos que caracterizan a cada una de estas fases.

2.2.1 Identificación de los objetivos del programa

Uno de los pasos más difíciles en la evaluación de programas es el de la determinación de los objetivos y metas del programa. Por qué esta dificultad? Es usual encontrar programas sociales cuyos objetivos están planteados en forma vaga y general, lo que hace difícil determinar si están siendo alcanzados o no. Algunos ejemplos usuales de estos objetivos son "mejorar el nivel de

salud mental”, “fortalecer la vida en familia” o “mejorar el funcionamiento social”.

Debemos pensar que todo programa debe ser evaluado en algún momento. Además, es necesario recordar que quién evalúa no siempre participa en la formulación de los objetivos del programa. Por estas razones, es un requisito plantear siempre estos objetivos en términos claros, específicos y medibles. Se requiere que sean lo suficientemente claros para que, quien evalúa sepa exactamente lo que tiene que buscar; específicos para que puedan utilizarse métodos de investigación y análisis apropiados.

Una vez definidos, se hace necesario identificar subobjetivos, esto es, las actividades o pasos necesarios para lograr estos objetivos. Este proceso de desagregación puede llevarse a cabo por niveles hasta lograr una jerarquía de objetivos, en la cual cada uno de éstos corresponde a una división, departamento u oficina responsable de cumplir con ese objetivo. La división del trabajo en una organización que se dedica a la prestación de servicios sociales es tal que los métodos de intervención empleados en un nivel constituyen los objetivos del nivel inferior. Asimismo, los objetivos de cualquier nivel constituyen los métodos del nivel superior siguiente. De esta manera, un programa social se puede dividir en un cadena de eventos en la cual cada uno de ellos es el resultado del que viene antes de éste y una condición necesaria del que le sigue. Por esta razón se afirma que la evaluación de programas consiste en demostrar que los objetivos de cada nivel están siendo logrados a través de los métodos empleados.

2.2.2 Determinación de las medidas de resultado

Las medidas de resultados, llamadas también las variables dependientes, permite evaluar el grado en que están siendo alcanzados los objetivos del programa. Existen muchos tipos de medidas para determinar el éxito de un programa. La selección de estas medidas dependerá en gran medida de la orientación del programa de acuerdo a lo definido en sus objetivos. Estas medidas de resultado pueden reflejar actitudes, valores, conocimientos, comportamientos, productividad y servicios ofrecidos, entre otros, aunque por lo general enfatizan cambios en los clientes (Raymond, 1985).

A veces los cambios propuestos no pueden medirse directamente. En estos casos se requiere establecer medidas de resultado relacionadas con los objetivos que se aproximen a estos cambios o que los reflejen. Así por ejemplo, un programa dirigido a mejorar la autoestima de niños agredidos puede no estar en capacidad de evaluar directamente el concepto psicológico de autoestima, pero puede administrarle a cada niño una escala que permita medir el concepto. Este tipo de medidas indirectas implican el uso de premisas o supuestos cuya validez es discutible, por lo que es recomendable recurrir, siempre que sea posible, a medidas directas. Cuando ello no sea posible, conviene emplear medidas o instrumentos cuya validez y confiabilidad ya han sido probadas en estudios similares en vez de crear nuevos.

Por lo general, el éxito de un programa social no puede evaluarse a través de una sola medida. En estos casos es útil usar varias medidas. Por ejemplo, un programa de Trabajo Social interesado en prevenir la ocurrencia de factores de riesgo social en población anciana tendría que desarrollar varias medidas de resultado como: auto reportes de los(as) ancianos(as) acerca de cómo se sienten, grado de funcionamiento físico, grado de desarrollo de enfermedades, observaciones del medio en que vive la persona (casa, vecindad inmediata para determinar barreras arquitectónicas y condiciones generales de la vivienda que impliquen riesgo).

Este es un caso típico en donde no contamos con una única medida que pueda servir como indicador del éxito del programa. No obstante, en estos casos es esencial que todas las medidas empleadas sean complementarias entre sí.

2.2.3 Identificación de las variables independientes e intervinientes

La persona que evalúa debe determinar no sólo el éxito del programa sino también por qué ocurren cambios como resultado del programa. Para ello hay que identificar y cuantificar las variables independiente e intervinientes. Las variables independientes pueden ser los objetivos de la agencia o programa, métodos, tipo de personal, localización, extensión y duración del servicio, así como también características de la clientela como sexo, edad, condición socioeconómica o actitudes (Raymond, 1985)

Las variables intervinientes son de dos tipos. El primero se denomina variables de operación del programa. Trata sobre la forma en cómo funciona el programa. Ejemplos de lo anterior serían las variables como frecuencia en la asistencia al servicio, grado de aceptación del programa por el personal profesional y participación de la clientela en otros programas sociales similares. Estas variables permiten identificar aquellos factores que afectan o condicionan el resultado de la intervención que se realiza a través del programa.

La cantidad de variables intervinientes que pueden ser medidas a la hora de evaluar un programa puede ser muy grande. Sin embargo, son pocos los programas que tienen la capacidad para evaluar todas las variables; tampoco es necesario considerarlas todas. Es recomendable, más bien, seleccionar aquellas cuya importancia ha sido reconocida por estudios previos, así como también a través de la experiencia de quienes trabajan en el programa o de acuerdo a lo que indica la teoría sobre cada caso.

El segundo tipo de variables intervinientes se denominan “puentes”, ya que están relacionadas con el logro de los subobjetivos del programa. Como señalamos antes, un programa social se organiza alrededor de una cadena de objetivos y subobjetivos. Teóricamente, para que un objetivo pueda ser alcanzado, el nivel inferior de subobjetivos tiene que haber sido logrado. Por lo tanto, éstos pueden ser considerados como variables “puente”, en donde la medición de estas variables prueba la validez de la teoría que sirve de fundamento al programa.

Un ejemplo es el siguiente. En un programa cuyo objetivo general es mejorar el nivel de salud de un grupo de ancianos(as), un subobjetivo o variable “puente” sería darle a los(as) beneficiarios(as) del programa tres tiempos de comida diarios para mejorar su dieta. El constatar hasta dónde todos los(as) ancianos(as) reciben las tres comidas para relacionarlo con el objetivo general, aporta información acerca de la relación entre los subobjetivos y el objetivo general.

2.2.4 Utilizando diseños de investigación

La aplicación de un determinado diseño de investigación depende del propósito de la evaluación, el énfasis de la misma (esfuerzo, efectividad o eficiencia) y de la naturaleza del programa que está siendo evaluado. Para orientar mejor la comprensión del tema, utilizaremos la clasificación de Raymond (1985) de estos diseños según se aplican a la evaluación de programas.

a) Diseños de desarrollo de hipótesis: estos describen una unidad de investigación a profundidad durante un periodo de tiempo previamente establecido. La descripción de una institución o programa le permite al (la) trabajador(a) social desarrollar ideas, interrogantes o hipótesis para un estudio posterior. Este tipo de diseños son particularmente útiles para alcanzar propósitos formativos o proveer un diagnóstico preliminar acerca de la efectividad de un programa. Por ejemplo, en estudio interseccional puede emplearse para determinar los efectos de un programa de tratamiento para alcohólicos crónicos. La trabajadora social definiría primero el concepto “alcohólico crónico” a través de la revisión de reportes retrospectivos de los clientes acerca de la ingesta de alcohol. esta información podría emplearse entonces para desarrollar una escala sobre el abuso del alcohol o para aislar las características particulares de las personas descritas como alcohólicos crónicos con el propósito de utilizarlas en investigaciones futuras.

b) Diseños cuantitativo-descriptivos: Estos permiten generalizar resultados a la población, pues recurren a procedimientos basados en la selección al azar. Tales diseños intentan responder a preguntas simples y descriptivas, generalmente a través de frecuencias o porcentajes relativos a una variable, tal y como es identificada por otra variable. Por ejemplo, un trabajador social podría recurrir a un diseño de este tipo para estudiar la satisfacción de la clientela con un programa de capacitación a padres. Una parte de quienes asistieron a las sesiones de capacitación son seleccionados al azar para administrarles una prueba de satisfacción paternal/maternal. Los resultados de la prueba reflejarán, entonces, la proporción de personas que están satisfechas con el estilo y contenido de las sesiones. Como la muestra fue seleccionada al azar, los resultados pueden generalizarse a todas las personas que asisten al curso.

c) Diseños de asociación: Estos examinan la presencia o dirección de relaciones entre dos o más variables. Por ejemplo, para este caso podríamos usar un diseño de pretest-postest aplicado a un grupo para evaluar el resultado de un programa de consejería estudiantil. El “ajuste estudiantil” definido de acuerdo a criterios predeterminados, se mide antes y después del proceso de

consejería puede utilizarse para complementar el estudio cuantitativo de los cambios ocurridos durante la ejecución del programa.

d) Diseños de causa-efecto: Estos le permiten al(la) investigador(a) identificar las variables independientes que generan cambios en las variables dependientes durante el desarrollo de la evaluación. Estos diseños le permiten a los(as) trabajadores(as) sociales concluir que los cambios ocurridos durante el desempeño del programa, ocurrieron de hecho como resultado de éste. El diseño denominado serie de tiempo interrumpida implica, por ejemplo, una serie de mediciones indicador de resultado antes, durante y después de concluido el programa.

Así, en un programa dirigido a elevar la autoestima en mujeres agredidas, un grupo de éstas podría participar en un taller de sensibilización. A las participantes se les aplicaría un instrumento con una escala de autoestima varias veces antes de comenzar el taller, durante y varias veces después de realizado. Si la autoestima aumenta mucho durante el taller y poco o nada antes y después del mismo, se podría concluir que los cambios son el resultado de éste.

El tipo de diseños descritos anteriormente ponen el énfasis en la evaluación de resultados. No obstante, se hace necesario muchas veces ir más allá del examen de la eficacia o efectividad del programa, estudiando la eficiencia del mismo. Como se discutió en la Sección I, una de las técnicas más comunes para esto es el análisis de costo-beneficio. Este tipo de análisis se concentra en los productos del programa, confronta el peso de cada alternativa con su efectividad, y compara los costos de cada una de acuerdo a su eficacia.

2.2.5 Diseñando los procedimientos para recoger la información

Una vez que hemos optado por el diseño para la evaluación y formulado sus componentes, debemos proceder a desarrollar una metodología para recolectar la información que permita responder a la o las preguntas que orientan la evaluación. En cuanto a esto, debemos insistir que los datos que se produzcan a través de la investigación evaluativa deben ser datos útiles, que sirvan para la toma de decisiones con respecto al programa.

Los instrumentos que se elaboren deben garantizar que los datos recogidos sean validos y confiables. La confiabilidad se refiere a la naturaleza estandarizada del instrumento (sea un test o un cuestionario) que asume que éste podrá producir información consistente a través de varias aplicaciones. Si por ejemplo, el cuestionario contiene preguntas cuya formulación es vaga, los(as) entrevistados(as) no comprenderán lo que se pregunta, y sus respuestas serán erráticas y arbitrarias. En estos casos, el instrumento no es confiable.

La validez se refiere a si nuestro instrumento mide lo que queremos medir. Es perfectamente posible que las personas respondan una pregunta consistentemente de una manera diferente a los que el(la) evaluador(a)

esperaba. Consecuentemente, los resultados sobre esa pregunta no pueden considerarse validos.

Tanto la confiabilidad como la validez son importantes en todo proceso de investigación. No obstante, hay que ponerle especial atención al problema de la validez ya que se relaciona directamente con la utilidad inmediata de los resultados. Si estos no son validos, al información no sirve para la toma de decisiones.

Como principio básico, conviene compartir con el personal las decisiones relativas a la evaluación del programa con el propósito de aumentar su comprensión de y compromiso con los resultados que se deriven de ella. Las y los funcionarios necesitan familiarizarse con algunas de las debilidades y fortalezas que encontramos en algunos de los métodos de recolección de datos y estar en capacidad de seleccionar aquellos que ofrezcan mayores ventajas en el marco de limitaciones impuestas por recursos insuficientes, falta de tiempo y la disponibilidad de la información. En síntesis, es mejor disponer de una respuesta probable y aproximada a la pregunta correcta que una respuesta cierta y definitiva a la pregunta incorrecta.

2.2.6 Recolección, análisis e interpretación de los datos

Una vez que se han definido las preguntas que orientan el problema de evaluación y escogido la metodología, hay que proceder a recoger los datos, analizarlos e interpretarlos. Esta fase implica ocho actividades (Austin, 1982):

- 1) Recolección y registro de los datos;
- 2) Organización de los datos;
- 3) Comparación de los datos;
- 4) Cuestionamiento de la información, esto es, identificación de las interrogantes o preguntas que se generan a partir de la información disponible;
- 5) Desarrollo de respuestas viables a las interrogantes. Estas respuestas constituyen las conclusiones de la evaluación, por lo que cada una debe estar debidamente fundamentada y verificada.
- 6) Formulación de lineamientos programáticos alternativos para cada respuesta a las interrogantes surgidas de los datos.
- 7) Elaboración de recomendaciones a partir de una priorización de alternativas.
- 8) Conversión de las recomendaciones en planes de acción.

Es recomendable involucrar al personal y administración del programa en parte o todo el proceso de análisis e interpretación de los datos. La recolección de

los datos puede ser una tarea que realice personal ajeno al programa; sin embargo, la interpretación se facilita cuando todas las personas involucradas en el éxito o fracaso de un programa aportan insumos acerca de lo que para cada una de ellas significa el “éxito” del programa. Es usual encontrarnos con evaluaciones que luego son cuestionadas por el personal o la administración debido a que realizan su propia interpretación de los datos.

Un principio sano: el énfasis debe estar en lo que los datos indican acerca de situaciones y relaciones específicas, en vez de juzgar al programa como “bueno” o “malo”. Un personal que esta involucrado en el proceso, vera más claramente las fortalezas y debilidades de la información y sabrá mejor como interpretar los resultados. El proceso de interpretación supone hacer inferencias sobre los datos para transitar del análisis a la acción.

2.3 PROBLEMAS QUE SE ENFRENTAN EN EL DISEÑO DE EVALUACIONES DE PROGRAMAS

En el diseño de evaluaciones sociales encontramos varios problemas de tipo de metodológico (Raymond, 1985):

2.3.1 Los objetivos del programa son por lo general poco claros y difíciles de medir.

2.3.2 Las medidas de resultado no siempre aparecen definidas en el programa y tienen que ser desarrolladas por quien evalúa.

2.3.3 Es problemático seleccionar al azar ya signar beneficiarios / sujetos a grupos experimentales y de control.

2.3.4 No siempre es posible manipular la variable dependiente.

2.3.5 Es difícil controlar los factores intrínsecos que puede afectar la validez interna de la evaluación.

Estos problemas, sin embargo, pueden controlarse y reducirse a través de un manejo y ejecución cuidadosos del diseño metodológico.

Existen otro tipo de problemas de orden no metodológico que pueden enfrentarse durante el proceso de evaluación. Uno tiene que ver con la determinación de los objetivos. Las organizaciones formales pueden tener objetivos informales que quieren alcanzar y que no están explicitados en documento alguno. Es necesario identificar la presencia de estos objetivos siempre que se vaya a evaluar efectividad o eficiencia.

Una de las mejores alternativas para lograr esto es a través del análisis sistémico (Etzioni, 1960). En vez de enfatizar el grado de éxito en alcanzar determinados objetivos formales, el enfoque sistémico busca establecer hasta dónde una institución o programa logra cumplir con sus objetivos bajo particulares condiciones.

Otro problema que enfrentan sobre todo los(as) evaluadores(as) externos es la resistencia de parte del personal. Este puede sentirse amenazado o puede operar bajo normas o tener intereses diferentes a los de la persona que evalúa o sentir que ésta viene a obstaculizar la prestación del servicio.

El personal puede concebir la evaluación como actividad que se realiza de una sola vez. La evaluación de programas tiene una naturaleza repetitiva, ya que sus resultados deben emplearse para mejorar el programa, volver a evaluar estas modificaciones para introducir los cambios necesarios para mejorar su eficacia y eficiencia. En realidad, la evaluación debe constituirse en un proceso sistemático y permanente del sistema formal de retroalimentación de la institución.

2.4 DETERMINANDO LA EVALUACIÓN DE UN PROGRAMA

La determinación de la evaluabilidad de un programa es una técnica de evaluación y un proceso administrativo de toma de decisiones diseñado para enfrentar el problema de programas que no logran producir los resultados esperados. Hay programas que pese a ser sistemáticamente evaluados, no logran mejorar con el tiempo. Algunos autores () plantean que esto ocurre como resultado de deficiencias en la descripción de los programas y la estructura administrativa gubernamental que limita la intervención de administradores y evaluadores. La técnica para la determinación de la evaluabilidad contribuye a sentar las bases para lograr que el programa mejore en su formulación y poder así, establecer cuál es el modelo evaluativo más adecuado para lograr una rigurosa evaluación o evaluaciones posteriores.

2.4.1 A qué se le llama determinación de la evaluabilidad de un programa social?

Esta técnica consiste en un proceso descriptivo y analítico del programa objeto de estudio que se sigue con el propósito de sentar las bases para una evaluación que pueda ser utilizada para la toma de decisiones. Fue desarrollada por el grupo de evaluación del Instituto Urbano (The Urban Institute) de los Estados Unidos de América entre 1966 y 1978. Este grupo sistematizó los problemas que todos sus miembros(as) habían enfrentado durante cinco años en el campo de la evaluación, identificando un conjunto de condiciones que deben reunirse por un(a) evaluador(a) antes de proceder a realizar una evaluación extensiva.

Este enfoque indica por obtener una descripción detallada del programa por parte de quienes tienen a cargo su administración. Esta es luego analizada sistemáticamente para determinar si reúne los siguientes requisitos:

- a) Está completa la descripción del programa?
- b) Es esta descripción aceptable a los términos de quienes tienen que tomar decisiones?

- c) Es esta descripción un fiel reflejo de cómo opera el programa en la realidad?
- d) Son factibles de alcanzar las expectativas que se tienen acerca del programa?
- e) La evidencia que requiere la administración para evaluar el programa, puede producirse confiablemente?
- f) La evidencia que requiere la administración para evaluar el programa, puede efectivamente recolectarse?
- g) Podemos esperar que el uso que pretende darle la administración a la información producida puede incidir en el desempeño del programa?

Todas estas preguntas deben ser respondidas por la persona que evalúa en conjunto con quienes tienen a cargo la administración del programa. En los casos en que existan vacíos o problemas de descripción, es necesario diseñar descripciones alternativas que resuelvan el problema.

2.4.2 Qué resultados podríamos obtener con esta técnica?

El propósito de la técnica es llegar a una descripción del programa que lo haga evaluable. Por lo tanto, podemos señalar tres tipos de conclusiones que harían un programa no evaluable:

- a) La comparación de la descripción del programa hecha por la administración con la lista de preguntas demuestra que aquella carece de algunos elementos, indicando que la administración aun tiene que tomar algunas decisiones importantes. Por ejemplo, un objetivo podría no haber sido definido en términos de medidas y expectativas. Por lo general, sin embargo, este tipo de problemas pueden resolverse con relativa facilidad.
- b) La comparación de las descripciones hechas por la persona que evalúa y quienes ejecutan el programa con la elaborada por la administración indica que esta última es no es razonable, su evaluación tendría un costo muy grande o no sería útil para la institución. Problemas de esta índole estarían reflejando que la descripción hecha por la administración es inadecuada o la del(a) evaluador(a) es incorrecta. En estos casos, quien evalúa debe presentar pruebas que demuestren que la descripción de la administración no es evaluable. Si no se llega a un acuerdo, entonces el(la) evaluador(a) necesita recoger más información sobre las actividades del programa para resolver el diferendo.
- c) La comparación de las descripciones del programa hechas por quien define la política y la administración indica que la descripción de esta última es inaceptable. Los desacuerdos entre lo que plantea quien define la política y la administración reflejan, por lo general, que ésta ha omitido algún objetivo o actividades importantes para quien define la política o

que ha incluido objetivos y actividades que no comparte aquel. En estos casos es necesario promover un proceso de discusión para que los diferendos se puedan resolver para completar los componentes que hacen falta.

Una vez que se llegan a conclusiones evaluables, debemos proceder a tomar una serie de decisiones –en conjunto con la administración del programa- en las áreas siguientes:

- ❖ Redacción definitiva de una descripción del programa que sea evaluable;
- ❖ Determinación de la información que se requiere obtener (comprar) para evaluar el programa;
- ❖ Establecer si se hace necesario introducir cambios en el diseño del programa;
- ❖ Determinar si conviene modificar las expectativas que tienen quienes definen la política sobre el programa;
- ❖ Cómo organizar y asignar personal para la ejecución de las decisiones.

2.4.3 Cómo administrar el proceso para determinar la evaluabilidad?

El equipo evaluador dedicara la mayor parte de su tiempo a la realización de las siguientes actividades:

- ❖ Coordinación del trabajo: planeamiento y calendarización de las actividades, control del personal y asignación de tareas, supervisión y control de la calidad del trabajo.
- ❖ Revisión de la literatura: lectura, resumen y procesamiento de documentos sobre el programa (programaciones, informes, memorandums, legislación, decretos, evaluaciones, etc.).
- ❖ Entrevistas a informantes claves: personas encargadas de la toma de decisiones, directores, administradores y personal adscrito al programa.
- ❖ Documentación de las descripciones del programa obtenidas de la investigación documental y entrevistas, para verificar la veracidad de las mismas.
- ❖ Observaciones en terreno para describir la intervención directa e indirecta del programa.
- ❖ Completar análisis para determinar la evaluabilidad.

- ❖ Documentación escrita de las conclusiones y evidencia presentada sobre las mismas.
- ❖ Organización y administración del proceso de toma de decisiones que sigue a la determinación de la evaluabilidad del programa.

Durante este proceso se espera que la administración del programa participe activamente en varias actividades, a saber:

- ❖ Aportando la descripción del programa;
- ❖ Valorando la validez de las descripciones y conclusiones elaboradas por la persona que evalúa.
- ❖ Orientando a quien evalúa acerca de la relevancia de las distintas actividades que se realizan en el programa y la dirección que lleva el trabajo.
- ❖ Decidiendo sobre las diversas opciones presentadas por el(la) evaluador(a) para resolver los problemas que afectan la evaluabilidad del programa.

2.4.4 Pasos a seguir para determinar la evaluabilidad de un programa

La determinación de la evaluabilidad de un programa implica cuatro pasos:

- 1) La descripción del programa;
- 2) La validación y análisis de la descripción del programa para determinar hasta donde puede ser éste evaluado;
- 3) La definición de alternativas, cuando éstas sean necesarias; y
- 4) Ejecutar las decisiones tomadas con respecto a la evaluación definitiva del programa.

El primer paso es crucial para el planeamiento de la evaluación y administración del programa. Generalmente nos encontraremos con serias discrepancias entre la conceptualización original del programa y cómo éste opera en la realidad. Estas contradicciones deben resolverse antes de iniciar cualquier evaluación. Dada la importancia que tiene en el proceso la descripción del programa, procederemos a detallar este primer paso.

2.4.5 La descripción del programa para efectos de establecer la evaluabilidad del programa

El fundamento de esta técnica reside en la descripción del programa realizada por la administración. Ello contiene dos elementos: programas reales y varias descripciones de programas reales.

Las descripciones son, por lo general, resúmenes muy condensados de programas reales, los que no siempre describen todas las actividades y resultados del programa. Estas carencias se incrementan si consideramos que las descripciones son muchas veces inexactas, pues reflejan lo que las personas piensan, quieren creer o quieren que otros creen acerca del programa. En la medida que las descripciones no incorporen características claves del programa o sean inexactas, las evaluaciones apoyadas en esa caracterización serán igualmente inexactas.

Para los efectos de la aplicación de la técnica, la descripción del programa debe cubrir tres grandes tipos de actividades:

a) De intervención directa: son aquellas diseñadas para cambiar o servir la sociedad de determinada manera. El contexto en que se realizan es aquel delimitado por los encargados de prestar el servicio y quienes lo reciben, por ejemplo: los centros de salud y su clientela, el I.M.A.S. y los adjudicatarios de su programa de vivienda, los inspectores de trabajo y los trabajadores.

b) De intervención indirecta: Son las actividades que lleva a cabo la administración para desarrollar y alterar intervenciones directas y mejorar su desempeño. Ejemplo de éstas serían sistemas de evaluación, programas de asistencia técnica, supervisiones regionales.

c) Para demostrar que el programa cumple: Son aquellas a través de las cuales la administración actúa para informar sobre el desempeño del programa y establecer expectativas realistas entre quienes están a cargo de la toma de decisiones. Estas se realizan a través de mecanismos formalmente establecidos como por ejemplo, control de presupuesto o informes regulares, así como también mediante canales informales como conferencias, publicaciones, reuniones y correspondencia.

Los siguientes componentes deben estar presentes en toda descripción de un programa:

- ❖ Una secuencia de eventos, la que sirve para definir en orden cronológico, la relación entre los principales eventos que ocurren en el programa. Esta secuencia describe la lógica del programa (supuestas relaciones de causa-efecto entre eventos). La secuencia entre eventos debe ser aceptable y bien definida, de manera que si el EVENTO A ocurre, podemos asumir que el EVENTO B tendrá lugar en consecuencia.
- ❖ Una descripción de cada evento. Estas se emplean para definir un evento, como por ejemplo, "los niños de las familias de tipo X leerán tan bien como las familias de tipo Y". Estas descripciones contienen los supuestos en que se funda el programa. Hay que evitar descripciones

de eventos que incluyen o implican dos eventos como “los niños de las familias tipo X leerán tan bien como las familias de tipo Y como resultado de adoptar el curriculum tipo Z”. En estos casos, es necesario describir estos eventos en dos planteamientos.

- ❖ Medidas y comparaciones. Se utilizan para determinar el tipo de evidencia que demuestra que el evento ha ocurrido. Para el evento anterior, al medida puede ser una comparación del desempeño de ambos grupos de niños en una prueba de lectura estandarizada. Las comparaciones pueden ser de diverso tipo: un grupo contra otro, un grupo contra sí mismo en diferentes puntos de tiempo o ambos.
- ❖ Valores esperados. Se usan para definir la dirección del y la cantidad de cambio esperada. Si afirmamos que la capacidad de lectura aumentará en los niños, tendríamos que suponer que cualquier aumento es aceptable. Para evitar este tipo de conclusiones, sería preferible enunciar que la capacidad aumentará en un 25% en un año o que al finalizar el programa un grupo alcanzará al otro.
- ❖ Las actividades sirven para definir y ayudar a examinar la plausibilidad del diseño básico del programa. Consecuentemente, el desempeño programático puede ser explicado a través los tipos y cantidad de actividad
- ❖ Los sistemas de información sirven para identificar las fuentes de donde se tomará la información, principalmente para determinar si éstos existen o es necesario crear nuevos sistemas.
- ❖ El uso que se le pretende dar a la información producida por la evaluación debe especificarse para determinar si vale la pena su costo.

La descripción del programa debe reunir una serie de requisitos que ayudan a prevenir problemas futuros en el momento de evaluar. Estos son:

- ❖ La descripción debe estar bien definida. Este requisito se cumple cuando se satisfacen los elementos contenidos en el Cuadro No.1.
- ❖ Es aceptable, esto es, cuando coincide con las expectativas de quienes tienen la responsabilidad de tomar decisiones.
- ❖ Es válida cuando acertadamente refleja las actividades programáticas que actualmente se realizan en terreno.
- ❖ Tiene credibilidad cuando existe evidencia de que lo descrito es real.
- ❖ Accesibilidad de la información se refiere a que el sistema de información descrito no impone costos o responsabilidades políticas más allá de los que quienes toman decisiones están dispuestos a asumir.

- ❖ Confiabilidad de la información. Existe confiabilidad cuando el sistema de información descrito permite realizar observaciones sistemáticas y verificaciones antes de utilizarlo.
- ❖ Uso realista de la información. Esto es posible cuando el empleo de la información producida es factible en el contexto de los recursos existentes y a través de actividades que controla la administración del programa.

Existen tres tipos de modelos que se emplean para describir programas: modelos lógicos, modelos funcionales y modelos de medición. Estos pueden usarse separada o integradamente.

MODELOS LÓGICOS: Son los diagramas que representan la lógica del programa. Ejemplo: