

TEMA VIII: LA EVALUACIÓN

OBJETIVO:

Conocer conceptos, métodos y técnicas de evaluación y capacitarse para evaluar servicios

5.1 CONCEPTO DE EVALUACIÓN

5.2 ASPECTOS METODOLÓGICOS

5.3 DE LOS OBJETIVOS A LOS INDICADORES: EL PROBLEMA DE LA MEDICIÓN

5.4 LA MEDICIÓN DE LA EFICIENCIA:

- **Análisis costo–beneficio**

5.5 LA UTILIZACIÓN DE LA EVALUACIÓN

1. CONCEPTO DE EVALUACIÓN

Podemos considerar la definición de Weiner (1990): Evaluación es una investigación aplicada a estudiar la estructura, el proceso y/o los resultados de un proyecto. Cuando hablamos de resultados incluimos efectos e impactos. Y al hablar de proyecto, incluimos programa o servicio. Y entendemos que los resultados y el proceso se miden en función de objetivos que implican valores sobre lo que es deseable o no deseable, sobre lo que se quiere o no se quiere conseguir.

Decimos que la evaluación es una investigación porque las técnicas que se utilizan son científicas, las propias de la investigación en ciencias sociales; es decir, que al hacer un diseño de evaluación se tiene en cuenta la validez y la fiabilidad y se utilizan grupos de control o de comparación, etc.

La evaluación es una investigación orientada a dar herramientas para poder tomar decisiones razonadas sobre la puesta en marcha, el mantenimiento, la modificación o la terminación de un proyecto. En consecuencia, si no hay posibilidades verosímiles o razonables de que los resultados de la evaluación puedan ser utilizados para las finalidades que se persiguen, vale más ahorrar tiempo y dinero y no evaluar.

Existe un alto grado de diversidad respecto a las ideas, los usos y las conceptualizaciones sobre la evaluación. El uso del concepto de evaluación varía según los diversos criterios que se empleen a la hora de enjuiciar tanto programas comprensivos como políticas sectoriales. La evaluación asume, consiguientemente, formas diferentes según el abanico de posibilidades metodológicas. En el siguiente cuadro se definen esquemáticamente distintos tipos de evaluación. Cabe establecer un rasgo común a todos ellos como *el modo sistemático de obtener información sobre las actividades, características y resultados de las políticas, programas o acciones puestos en vigor*.

Si bien la evaluación de la eficacia pone el énfasis de su análisis en objetivos y resultados (*objective output*), la eficiencia concentra su atención en la relación generada entre recursos y resultados (*input–output*). En este contexto deben tenerse en cuenta dos importantes consideraciones:

- para buscar una valoración de su eficacia o eficiencia, los responsables políticos, los gerentes y los administradores de programas deben especificar cuales son los destinatarios de sus políticas y qué criterios se utilizarán para medir los efectos y posibles cambios introducidos en aquellas políticas o programas. La

estrategia de intervención propuesta, en conexión con la política o el programa seleccionado, debería estar igualmente propuesta de antemano. Sin embargo, se puede comprobar a menudo cómo programas de política social o de salud se caracterizan por la simple enumeración de fines descritos muy genéricamente y hasta vagamente. A veces se denota, simple y llanamente, la ausencia de una expresa declaración de fines. En estas circunstancias el ejercicio de la evaluación es ambiguo cuando no imposible.

- La producción de indicadores de rendimiento medibles, válidos, sensibles y fiables es esencial tanto para la evaluación orientada a la eficiencia como a la eficacia del programa o la política en vigor. De hecho, la medición del resultado aparece, bien sea como denominador o como numerador, en ambos modos de evaluación. La necesidad de mejorar la medición de resultados producidos por la puesta en vigor de políticas o programas es de una crucial importancia para realizar evaluaciones que sean metodológicamente aprovechables.

Tipos de evaluación

Análisis coste-beneficio	¿Cuál es la relación entre los costes del programa a realizar y los resultados del mismo (beneficios) expresados en términos monetarios?
Análisis coste-efectividad	¿Cuál es la relación entre los costes del programa y los resultados (cuando los resultados no se pueden traducir en términos monetarios?)
Evaluación de la eficacia	¿Hasta qué punto es efectivo el programa desarrollado en la consecución de sus fines?
Evaluación de la eficiencia	¿Cuáles son los recursos aportados al programa en relación a la ratio personal/usuarios y a otros aspectos relativos a los niveles de actividad del programa puesto en marcha?
Evaluación externa	La evaluación se lleva a cabo por personas no integradas directamente en el equipo que desarrolla el programa, con el fin de conseguir una mayor objetividad evaluativa
Evaluación interna	Realizada por miembros del equipo que realiza el programa o política puesta en marcha
Evaluación formativa	¿Cómo puede mejorarse el programa que se está llevando a cabo?
Evaluación del impacto	¿Cuáles son los efectos directos e indirectos producidos con la aplicación del programa?
Evaluación de resultados	¿Hasta qué punto los resultados que se esperaban han sido cumplidos satisfactoriamente? ¿Cuáles son los efectos producidos por el programa en la conducta de los usuarios?
Evaluación del proceso	¿Cuáles son los aspectos más débiles y más consistentes en las actividades del día a día? ¿Puede mejorarse tal proceso de actividades?
Evaluación conclusiva	¿Debería continuarse el programa en el futuro?

Dos enfoques específicos en el desarrollo de los estudios de evaluación que se relacionan con las dimensiones formativa y conclusiva (*summative*) de la evaluación:

- EVALUACIÓN FORMATIVA

- Persigue obtener información relativa al grado de cumplimiento de un programa o política respecto de sus fines.
- Se ocupa del proceso en sí.
- Requiere un alto grado de conocimiento de los múltiples aspectos presentes en los programas o políticas a desarrollar.
- Ayuda a observar con mayor claridad las metas propuestas.
- Ayuda a identificar los cambios factibles que se requieren para mejorar la eficacia en las actividades del programa en marcha.

• EVALUACIÓN CONCLUSIVA

- Se ocupa del acopio y presentación de información para resúmenes y juicios valorativos del programa o políticas concretas llevadas a cabo.
- Se orienta a los resultados finales.
- Busca contrastarlos efectos y los costes del programa con aquellos otros de programas alternativos que persiguen las mismas metas.
- Ayuda a decidir si un programa debe continuarse o no y, en caso afirmativo, en qué medida.

Existen diversas razones para la utilización de un enfoque formativo o conclusivo, dependiendo de los objetivos y necesidades del programa, política o acción a desarrollar. En general, los responsables en el nivel estatal, así como los patrocinadores institucionales, suelen mostrarse más interesados en la obtención de un tipo de información conclusiva. Pretenden conocer, ante todo, si el programa funcionó adecuadamente y si los efectos fueron los deseados, así como si se produjeron efectos secundarios, inesperados y negativos. De acuerdo al tipo de información conclusiva, necesitan determinar si el programa debería proseguirse, ser ampliado, reducido o, simplemente, abandonado. Mientras que el personal directamente implicado en la provisión a los usuarios de determinados programas, políticas o actuaciones, estarán generalmente más inclinados a obtener un tipo de evaluación formativa.

5.2 ASPECTOS METODOLÓGICOS

No existe un único método de evaluación adecuado a todas las situaciones. Sin embargo, la considerable variedad de técnicas pueden ser agrupadas en tres grandes categorías:

- **TÉCNICAS DE SUPERVISIÓN.** Pretenden registrar lo sucedido durante el desarrollo del proyecto. Este tipo de técnicas basa su actuación en procesos de seguimiento para la revisión de las operaciones del programa, tales como estudios del cumplimiento de cronogramas y auditorías de tipo administrativo y de responsabilidad ejecutiva.
- **TÉCNICAS DE COSTE ANALÍTICO.** Su finalidad es la de tasar el valor relativo de un proyecto en relación a sus costes. Estas técnicas incluyen: contabilidad de costes y análisis de coste-beneficio y de coste-eficacia.
- **TÉCNICAS DE INVESTIGACIÓN SOCIAL.** Tratan de descubrir relaciones de causalidad entre acciones del programa y resultados cuantificables. Las técnicas de investigación social más frecuentemente usadas para evaluar medidas de política social incluyen experimentos, diseños cuasiexperimentales, encuestas y estudios de casos concretos.

5.3 DE LOS OBJETIVOS A LOS INDICADORES: EL PROBLEMA DE LA MEDICIÓN

Un problema relativo a la eficacia es LA MEDICIÓN: los instrumentos para cuantificar están íntimamente ligados a la conceptualización y desarrollo de indicadores de resultados medibles, válidos, sensibles y fiables. Constituye uno de los aspectos metodológicos más espinosos a los que se enfrenta la evaluación.

Hay una instancia central en el proceso de evaluación: **determinar el grado en que se han alcanzado las**

finalidades de un proyecto.

Para ello, el objetivo general del proyecto debe ser dimensionalizado en subconjuntos de objetivos específicos, que a su vez tendrán metas, cuyo logro será medido a través de indicadores. Las variaciones en los valores que se verifican en las unidades de análisis, permiten cuantificar este proceso.

Operacionalmente, el indicador es la unidad que permite el alcance de una meta. Pero la relación entre unidad y meta es de carácter probabilístico y no de implicación lógica, por lo cual resulta conveniente incrementar el número de indicadores de una meta para aumentar así la probabilidad de lograr una medición adecuada.

Este proceso puede analizarse tanto desde la lógica de la deducción como desde la lógica de la inducción. En el primer caso, se va de los objetivos a los indicadores; en el segundo caso, el proceso parte de los indicadores

La evaluación aprovecha los avances alcanzados por la metodología de la investigación, que ha desarrollado procedimientos y técnicas que permiten transformar los conceptos abstractos en variables mensurables.

Lazarsfeld (1979) ofrece una secuencia para la transformación de los conceptos en variables:

- Se parte de una representación literaria del concepto: de una abstracción derivada de la observación de un conjunto de hechos particulares (p.e. noción de inteligencia...)
- La segunda fase consiste en la especificación del concepto, que implica su división en las dimensiones que lo integran. Dichas dimensiones pueden ser deducidas analíticamente a partir del concepto general que las engloba o empíricamente a partir de la estructura de sus interrelaciones. De todas formas, un concepto corresponde, casi siempre, a un conjunto complejo de fenómenos y no a un fenómeno simple y directamente observable. La deducción analítica de las dimensiones lleva a la participación del concepto que formalmente debe ser: 1) exhaustiva, significa que las dimensiones deben cubrir todo el concepto; 2) excluyente, que las dimensiones no deben superponerse o interceptarse entre sí (p.e.: inteligencia abstracta; inteligencia verbal; inteligencia manual).
- La tercera etapa: elección de indicadores que permitan medir cada una de las dimensiones del concepto. Conviene utilizar una cantidad elevada de indicadores, para garantizar que se está midiendo el concepto que se pretende medir y disminuir el efecto negativo que deriva de haber elegido un mal indicador.
- El cuarto paso: formación de índices para sintetizar los datos resultantes de las etapas anteriores. En la construcción de índices hay que resolver dos problemas. Primero, decidir qué datos de dimensiones parciales han de integrarse en el índice. Teniendo en cuenta que un índice debe referirse explícitamente a TODAS las dimensiones parciales del espacio del correspondiente fenómeno, pero sólo a éstas. En segundo lugar, se debe indicar de qué manera se han de poner en relación mutua esos datos (asignación de valores numéricos, ponderación, combinación (Maintz, 1975, 62).

Los más corrientemente usados son:

- Índices sumatorios simples, en los que cada dimensión o cada indicador tiene el mismo peso.
- Índices sumatorios ponderados, en los que en función de consideraciones teóricas o derivadas del análisis empírico, se asigna un peso diferente a las dimensiones o indicadores.

En el caso del concepto de inteligencia, los indicadores para la Inteligencia abstracta serían los tests A, B y C; para la inteligencia verbal, los tests D, E y F; de la inteligencia manual los tests G, H e I. Cada test entrega resultados que fluctuarían entre 1 y 100, que es el puntaje máximo obtenible. Sobre esta base, se pueden ordenar los resultados obtenidos, construyendo una escala como la siguiente:

Muy alto de 851 a 900

Alto de 801 a 850

Superior al promedio de 751 a 800

Promedio de 701 a 750

Inferior al promedio de 651 a 700

Muy inferior al promedio de 601 a 650

El rango verdadero de variación de la variable ha sido de 600 a 900 para los sujetos de la población estudiada. La resultante del test se denomina PUNTAJE DE LA UNIDAD DE ANÁLISIS que, comparada con la escala previamente construida, permite ubicar a los integrantes del universo dentro de las categorías establecidas por la misma.

Alternativamente, se puede construir un índice sumatorio ponderado, que difiere del anterior solamente en que las dimensiones y variables están ponderadas desigualmente (en la anterior, el factor de ponderación era siempre igual a 1).

Si la teoría indica que cada tipo de inteligencia tiene diferente importancia, podrían fijarse factores de ponderación para cada dimensión, que reflejen esa diferencia; p.e., 3 para la abstracta, 2 para la verbal y 1 para la manual. Se estaría asumiendo no sólo que hay un orden entre distintos tipos de inteligencia, sino que existe igual tipo de distancia entre los mismos. Esto quiere decir que en lugar de operar con una escala ordinal, se utiliza una de naturaleza intervalar. Supone que la inteligencia abstracta es tres veces más importante que la manual y que ésta tiene la mitad de jerarquía que la verbal.

Puede suceder también que los tests que miden cada tipo de inteligencia no tengan la misma relevancia para apreciar lo que supuestamente miden. Por ello, se ponderan desigualmente los resultados que entregan, estableciendo que los tests A, D y G tienen un factor de ponderación igual a 3; los tests B, E y H igual a 2 y los C, F e I igual a 1.

Los puntajes obtenidos derivan del siguiente cálculo: la dimensión I A tiene un factor 3, el test A = 3, el test B = 2 y el test C = 1. Recordando que el puntaje máximo obtenible en cada test es igual a 100, se tiene:

Factor de Factor del Puntaje máximo Puntaje máximo

la dimens. indicador del test teórico

Test I A=3 Test A=3 100 $3 \times 3 \times 100 = 900$

Test B=2 100 $3 \times 2 \times 100 = 600$

Test C=1 100 $3 \times 1 \times 100 = 300$

La escala a construir se elaboraría por un procedimiento idéntico al descrito para los índices sumatorios simples.

Un objetivo general no se descompone en objetivos específicos ni éstos en metas concretas a través de la deducción, sino a partir de prioridades políticas y de la racionalidad técnica.

Por ello, la descomposición de un objetivo en metas no constituye una partición, que es un proceso formal, sino que traduce la importancia que asignan quienes toman las decisiones a la población afectada por el proyecto y al tipo de necesidades que éste va a satisfacer.

Ello no obsta para que la evaluación se realice en proyectos formulados con base en la teoría social, que comparte con ella buena parte de la metodología utilizada.-->[Author:sokbo7]

5.4 LA MEDICIÓN DE LA EFICIENCIA

ANÁLISIS COSTO-BENEFICIO

El ACB posibilita la evaluación de proyectos, determinando si son o no rentables, mediante la comparación **de los flujos de**

Beneficios y de costos que su implementación implica.

Para la identificación de los costos y beneficios del proyecto que son pertinentes para su evaluación, es necesario definir una situación base o situación *sin proyecto*; la comparación de lo que sucede con proyecto versus lo que hubiera sucedido –sin proyecto, definirá los costos y beneficios pertinentes del mismo (Fontaine, 1984,27).

La evaluación puede ser realizada desde dos ópticas diferentes:

- la privada, que considera una unidad económica concreta, pudiendo ser ésta un individuo o una organización.
- la social, en la que se considera a la sociedad en su conjunto (normalmente el país), o algún ámbito espacial definido dentro del mismo (región, provincia, etc.)

Tanto en los proyectos económicos como en los sociales, junto con la evaluación social, debe realizarse la evaluación privada, con la perspectiva individual. A partir de esta última, se determina la estructura de los incentivos que el proyecto debe ofrecer a sus participantes.

Una evaluación social muy satisfactoria no sirve si los potenciales beneficiarios carecen de interés en participar en el proyecto. Tampoco tiene sentido un proyecto con una alta tasa de rentabilidad social, si los beneficiarios no pueden vivir con los ingresos que el proyecto genera.

El ACB permite determinar los costos y beneficios que se deben tener en cuenta en cada una de las perspectivas consideradas previamente. Por otro lado, mediante la actualización, hace converger los flujos futuros de beneficios y costos en un momento dado en el tiempo (valor presente o actual) tornándolos comparables. Ello, empero, no reemplaza a la teoría sustantiva sobre el fenómeno en cuestión, que determina las dimensiones relevantes que se deben incluir en el análisis.

5.6 UTILIZACIÓN DE LA EVALUACIÓN

Leviton y Hugues consideran que existen cinco condiciones que afectan al sistematismo y consistencia de la utilización de la evaluación:

- relevancia
- comunicación entre evaluadores y usuarios de la evaluación
- procesamiento de la información por parte de los usuarios
- plausibilidad de los resultados de la evaluación
- participación y apoyo de los usuarios.

De los trabajos de Leviton y Hugues (1981), Weiss y Buculavas (1980a/b), Solomon y Shortell (1981) y Rossi y Freeman (1987), pueden extraerse un número de provechosas consideraciones generales respecto a la utilización de la evaluación:

- Dado que las evaluaciones se realizan por encargo de grupos específicos de individuos e instituciones, los evaluadores deben estar familiarizados con los valores, estilos cognitivos e intereses de los ejecutivos o decisores (*decision-makers*), así como de los compromisos en juego por parte de los responsables en la implementación factual de los programas.
- Los resultados de la evaluación deben acompañarse temporalmente con el desarrollo de las políticas puestas en vigor y deben estar disponibles en el momento oportuno. Objetivo, pues, de los evaluadores, es el de encontrar el equilibrio apropiado entre el ritmo temporal y la accesibilidad a las conclusiones evaluativas mediante la realización de análisis metódicos y comprensivos. A veces, sin embargo, las características de la audiencia a la que se dirige el evaluador hacen más apropiada la presentación informal –incluso verbal– de los informes que la publicación de sus hallazgos en una reputada revista académica o mediante un sesudo *dossier*.
- Los apartados de utilización y difusión deberían ser parte del diseño evaluativo. De esta manera los resultados evaluativos también se tendrían en cuenta y serían maximizados al incorporar el aspecto formativo de los usuarios de los mismos. Comporta el examen de los puntos fuertes y débiles de los análisis realizados, así como la articulación del modo en que la información generada por la evaluación puede ser transmitida por los ejecutivos y responsables de los programas/políticas/acciones a las instancias a las que tienen que rendir cuentas. Para ello es muy aconsejable una anticipación en la determinación de las posibles críticas.
- La evaluación debe incluir una valoración de su propio uso. Los evaluadores y ejecutivos (*decision-makers*) compartirían de tal manera un acuerdo común respecto a los propósitos por los cuales se lleva a cabo el estudio evaluativo, así como respecto a los criterios bajo los cuales su utilización posterior se consideraría como fructífera.

Hasta tiempos recientes, no pocos investigadores consideraban que su responsabilidad se circunscribía meramente a diseñar estudios de evaluación, recoger información y publicar los resultados. Lo que hiciesen o dejasen de hacer los responsables (*policy-makers*), no era asunto suyo. El rigor metodológico y la calidad técnica eran las principales preocupaciones de los investigadores. El uso de la evaluación no era asunto de su incumbencia. El rigor metodológico suponía generalmente un diseño experimental, información de carácter cuantitativo y un detallado análisis estadístico. El problema de los investigadores no era si los responsables en la toma de decisiones entendían o no los análisis complejos que frecuentemente se llevaban a cabo (Patton 1986).

En la actualidad se mantiene que un mayor rigor metodológico no soluciona por sí sólo el problema de la utilización –a menudo subutilización– de la evaluación. Producir evaluación es una cosa y otra bien diferente es utilizarla apropiadamente.

Con el fin de lograr una utilización más efectiva de la evaluación, los investigadores y técnicos evaluadores deberían preocuparse principalmente del examen de los métodos óptimos de exponer los resultados de la evaluación. Se necesita realizar un trabajo más imaginativo con el fin de enjuiciar la actitud de los patrocinadores que buscan información y de aquellas otras personas en situación de aplicar posteriormente los resultados de la investigación. Mediante la informática, la presentación de los resultados puede hacerse de modos diferentes para maximizar las expectativas tanto de los profesionales técnicos de los servicios sociales como de los políticos o ejecutivos integrados en los diversos niveles del proceso de toma de decisiones.

Patton (1987) propone para los evaluadores la asunción de una serie de compromisos, sensibilidades y cualidades personales especiales con el fin de maximizar el uso de la evaluación. Entre ellos deben destacarse los siguientes:

- Compromiso personal con respecto al uso de la evaluación
- Cualidades para ayudar a formular a los no-investigadores cuestiones de carácter evaluativo sustancial
- Sensibilidad política y dotes diplomáticas
- Cualidades para la comunicación interpersonal

- Un amplio abanico de métodos alternativos de evaluación
- Cualidades para dirigirse a audiencias de composición muy variada mediante distintos niveles de sofisticación en la presentación de los análisis e investigaciones
- Creatividad.

Patton llama la atención sobre los cambios de énfasis que tienen lugar en el transcurso de la evaluación conforme ésta avanza:

- Naturaleza de los métodos de evaluación que deben incorporar la posterior utilización de la misma; conveniencia y posibilidades de desplegar al completo el proceso evaluativo, e influencia de las prácticas políticas implicadas en el desarrollo de la evaluación.
- Traslación de un predominio de los métodos puramente cuantitativos y experimentales hacia opciones de carácter múltiple, de acuerdo a la especificidad del contexto en el que se desarrolla la evaluación.
- Ampliación del ámbito de la evaluación desde el interés concreto de saber si un programa/política/acción ha conseguido sus objetivos hasta la atención de un creciente número de cuestiones, propósitos, enfoques, modelos y usos, a veces interrelacionados entre sí.
- Progresivo rechazo –implícito o explícito– de los usuarios de la evaluación hacia informes de corte académico y que no se ajustan a sus expectativas concretas.
- Mayor profesionalización de los evaluadores que han incrementado notablemente sus conocimientos mediante las experiencias portadas por la gran expansión de las políticas de servicios sociales durante las décadas de los años setenta y ochenta en Europa occidental y Norteamérica. En este proceso de maduración, los profesionales de la evaluación han focalizado su interés en el análisis de los valores corporativos en concurrencia, así como en los conflictos específicos presentes en cada ámbito de evaluación. Todo ello desde una perspectiva superadora de la meta interpretación instrumental –lógica y racional– de datos e informaciones pulcramente recopilados.

La evaluación, pues, debe ser utilizable. Debe ayudar al desarrollo del programa sin impedir su acción. Por eso, además de una metodología consistente, la correcta utilización de la evaluación depende en buena medida en comprender qué se espera de ella y en definir los límites dentro de los cuales la información por ella generada debe aplicarse. Para ello es decisiva la articulación de procesos participativos que integren tanto a evaluadores como a usuarios de la evaluación.

EL PERT (*Program Evaluation and Review Technique*)

El PERT permite representar y analizar una secuencia de tareas; en este sentido es particularmente útil para la planificación de la ejecución de un proyecto o de un programa. Sin embargo, puede ser utilizado para analizar cualquier tarea que se describa por actividades y acontecimientos.

El método consiste en representar las diferentes actividades sobre un gráfico de tipo red, de tal manera que la secuencia de acontecimientos pueda ser vista en su totalidad. Esto permite al planificador, preocupado por el tiempo y por el coste de las operaciones, la verificación de la posible aceleración del proceso. El diagrama sirve también para el seguimiento (*monitoring*) y para el control de las actividades durante el proyecto.