

GUIA BÁSICA PARA EL DISEÑO DE SISTEMAS DE EVALUACIÓN DE PROGRAMAS Y POLÍTICAS SOCIALES

Ángel Martínez Moreno

Esta guía pretende constituirse en un instrumento de trabajo que ofrece, en un solo documento los elementos básicos a partir de los cuales construir un sistema de evaluación eficaz. La guía puede ofrecer algunos criterios de selección y definiciones, pero la elaboración de los mecanismos y de los contenidos concretos corresponde a los responsables de la evaluación.

(A lo largo de la presente guía hemos utilizado de manera, a veces indistinta y en ocasiones alternativa, las palabras proyecto y programa. Sabemos que un programa se compone de proyectos y los proyectos de actividades. No es que consideremos que el concepto de programa y el de proyecto sean intercambiables entre sí, sino que lo que se dice aquí acerca de la evaluación se puede aplicar tanto a uno como a otro. Cuando no sea así, se indicará específicamente)

LAS DIMENSIONES BÁSICAS DE LA EVALUACIÓN

Dimensión descriptiva:

La evaluación ha de ser capaz de facilitarnos el conocimiento de la realidad identificando sus elementos de forma ordenada y dotando de sentido los acontecimientos de los que trata. Para ello ordena y clasifica los datos, los hechos, las situaciones, etc.

Dimensión analítica:

La evaluación ha de ser capaz de ofrecernos imágenes de la relación entre los diferentes elementos que configuran la realidad. Para ello establece vínculos de causa efecto entre diferentes factores, descompone situaciones en otras más sencillas, plantea explicaciones acerca de cómo se relacionan unos elementos con otros, etc.

Dimensión valorativa:

Se le da valor a los hechos, fenómenos, o elementos asociados a una situación dada. Se mide, se utilizan estándares, se comparan los resultados obtenidos con los objetivos propuestos, etc.

PARA QUÉ SE EVALÚA

El sistema de evaluación que se diseña ha prever concretamente cuáles van a ser los usos de la evaluación. En gran medida el **para qué** va a determinar el **qué** y el **cómo**. Hemos de considerar aquí que el **para qué** de la evaluación presenta habitualmente dos caras: una explícita, programática y habitualmente proclamada; y otra implícita que hace las veces de lo que se ha venido a llamar agenda oculta. Aunque parezcan contrapuestas, e incluso incompatibles, no lo son, sino que simplemente tienden a comportarse de tal manera que coexisten y finalmente, las dos influyen en los resultados. En la práctica las dos están presentes y ha de ser consideradas.

La agenda oculta ha de evaluarse en función de criterios que le son propios y que, por su misma naturaleza no pueden ser explicitados aquí, sino que corresponde a los agentes titulares de la agenda el hacerlo.

En referencia a la agenda explícita puede plantearse la evaluación para dos tipos de acciones:

1.- Legitimadora: medir la eficacia y eficiencia de un proyecto

- Identificar hasta qué punto un proyecto o programa es capaz de resolver los problemas o situaciones que le dieron origen.
- Identificar el efecto diferencial de las diferentes actividades que componen el programa y el valor que añaden a los resultado y por tanto a la eficacia final.
- Valorar la relación coste –beneficio entre los esfuerzos planteados por la puesta en marcha y realización del programa de intervención y los resultados obtenidos (productividad).

2.- Operativa: facilitar la toma de decisiones

- Determinar la oportunidad de continuar con el proyecto o abandonarlo.
- Plantear formas alternativas de uso de los recursos disponibles
- Mejorar las prácticas de intervención

- Modificar el desarrollo del proyecto o mantenerlo
- Identificar formas de aplicación a realidades similares a partir de las experiencia acumulada en el desarrollo de los proyectos.
- Pilotar el proyecto mientras se encuentra en ejecución a través de mecanismos de realimentación y control
- Asignar recursos a programas competitivos entre sí
- Aceptar o no las implicaciones de una hipótesis o teoría.

3.- Social

En algunas modalidades de intervención, la evaluación constituye un factor fundamental de implicación de los agentes sociales en los programas y proyectos. En este sentido, la evaluación constituye un instrumento de participación; permite que diferentes agentes que se encuentran actuando en distintos proyectos dentro de un mismo programa conozcan de primera mano los resultados obtenidos hasta el momento, se refuerce el sentido de pertenencia, se planteen alternativas de coordinación entre los responsables de diferentes actividades y proyectos; se generen sinergias; y se origine una visión compartida que resulta imprescindible en todos los procesos participativos.

Independientemente de estas funciones, en ocasiones la evaluación se convierte en un elemento de fiscalización sobre aquellos que han gestionado o dirigido un proyecto. Cuando esto ocurre estamos asistiendo a una utilización perversa de la evaluación. Desgraciadamente este uso ha sido más frecuente de lo que a primera vista pudiera parecer y constituye uno de los fenómenos responsables de la desconfianza de los profesionales y de los participantes en programas sociales hacia las instituciones y los responsables de la realización de evaluaciones externas y supuestamente objetivas.

| |
|---|
| <p style="text-align: center;">¿QUIÉN PUEDE ESTAR INTERESADO EN LOS RESULTADOS DE LA EVALUACIÓN?</p> |
|---|

Dado que la evaluación ha de tener consecuencias, es preciso identificar a todas aquellas personas o instancias organizativas que se verán afectadas por esas consecuencias.

Todos aquellos que van a utilizar los resultados o verse afectados por los mismos son también instancias que pueden actuar como informadores o, al menos, han de estar informados acerca de los procedimientos y contenidos relevantes que

determinarán la evaluación. Ello es así porque la evaluación es, básicamente, un proceso que consiste en observar la realidad desde diferentes perspectivas y con distintos niveles de detalle, y los resultados de lo observado han de ser reconocidos como reales y útiles por todos aquellos que han de verse afectados por ellos (validez aparente). Así, los instrumentos que se utilicen no sólo han de cumplir todos los requisitos científicos (como fiabilidad, validez, objetividad, parsimonia, etc.) sino que, además, han de resultar creíbles para los usuarios de la evaluación.

Cada una de las partes interesadas en los resultados de la evaluación pueden tener intereses legítimos diferentes al respecto. En este sentido los intereses de los técnicos, políticos, administradores, usuarios, etc. no tienen por qué coincidir entre sí, pero el sistema de evaluación ha de satisfacer los distintos requerimientos. Para cada uno de ellos se deberán elegir preguntas, indicadores y métodos que, a la vez que cumplen los requerimientos científicos, se adecuan a las características específicas de las partes interesadas. De lo contrario, lo que se pone en peligro es el uso mismo de los resultados de la evaluación.

¿QUÉ SE EVALÚA?

1.- Los objetivos

Todo proceso de evaluación se legitima en primer lugar por el contenido de lo evaluado. En este sentido, el contenido que, en principio, aparece como fundamental son los objetivos del programa. El programa tiene sentido si es capaz de resolver los problemas que lo originaron (objetivos). Pero, en la práctica, la evaluación de los objetivos no resulta tan sencilla como en principio pudiera parecer dado que en ocasiones no están formulados de manera que se puedan medir o se producen reduccionismos absurdos en las unidades de medida. El fenómeno del reduccionismo se produce cuando se confunde el objetivo con un solo indicador que lo señala, de forma que se confunden los medios con los fines y se persiguen cuotas, generalmente numéricas, que son capaces de desplazar el centro de atención de los programas.

Habitualmente podemos distinguir tres tipos de objetivos:

- **Objetivos estratégicos:**

Son propios de grandes programas que a su vez se componen de proyectos y actividades. En general hacen referencia a la relación de la organización, territorio o colectivo humano de referencia con su entorno.

Suelen plantearse en un horizonte temporal de dos o más años (dependiendo de las características del entorno) Estos objetivos han de plantearse de forma amplia y flexible, de manera que no queden obsoletos ante cambios imprevisibles. Proporcionan criterios de decisión que se aplicarán, con diferentes niveles de concreción, a las diferentes programas y proyectos.

- **Objetivos tácticos**

Poseen un horizonte temporal que generalmente coincide con el ciclo anual o con el ciclo de un proyecto en el contexto de un programa (aunque no necesariamente). Han de ser coherentes con los objetivos estratégicos y concretarlos, de manera que se les exige que sean cuantificables a través de indicadores precisos y objetivables.

Habitualmente suelen referirse a actividades que se relacionan con la consecución de resultados que implican la coordinación de diferentes funciones y procesos.

- **Objetivos operativos.**

Tienen un horizonte temporal de corto plazo. Se relacionan con actividades a las que se exigen rendimientos específicos, por lo que su formulación ha de ser concreta y medible de forma objetiva.

Como puede observarse fácilmente cada nivel de objetivos se concreta en el nivel inferior, por lo que el sistema de objetivos y metas de la organización demanda un elevado grado de coherencia entre los diferentes tipos de objetivos que se propone.

De todas formas, cabe considerar que son precisamente los objetivos los que constituyen el punto de partida y justificación de la propia existencia del programa. Por tanto cabe precisar, con el mayor detalle posible, cuales son los resultados mínimos previsible (objetivos mínimos) que es necesario conseguir para que se considere aceptable el programa.

Finalmente, puede plantearse la existencia de objetivos diferentes en un mismo programa para distintos grupos poblacionales, contextos territoriales, o niveles organizativos, etc. Ello implica que no sea necesaria la uniformidad de los objetivos respecto a todos los segmentos de población, y que pueda considerarse un valor operativo el reconocimiento de la diversidad.

2.- Los resultados libres de objetivos

Precisamente para evitar el fenómeno del reduccionismo se ha planteado la necesidad de identificar los resultados obtenidos por un programa y no previstos inicialmente en el mismo. Efectivamente, dado que la realidad social es compleja, ello implica que sea también en gran medida impredecible y que cuando incorporamos una novedad (programa) a un sistema social, pueda provocar efectos, tanto esperados como no esperados. El evaluador ha de ser capaz de identificar los cambios no esperados que un programa provoca o potencia que se que provoquen a fin de identificarlos, medirlos en la medida de lo posible e incorporarlos posteriormente a las previsiones del mismo programa o de otros programas futuros.

Los resultados libres de objetivos, es decir, aquellos no previstos formalmente, pueden ser de tres tipos básicamente:

- **Indiferentes:** es decir, que no afectan significativamente a la realidad social de la que nos estamos ocupando.
- **Multiplicadores:** es decir que potencia aspectos deseables en la dirección de las intenciones de cambio del programa
- **Perversos:** es decir que potencia o activa aspectos no deseables y/o contrarios a las intenciones de cambio del programa.

3.- El impacto.

La evaluación del impacto puede ser considerada como la conjunción de la evaluación de los objetivos y la libre de objetivos, es decir, los resultados totales conseguidos por el programa. En realidad, en la práctica, si se trabaja con poblaciones o zonas geográficas muy extensas es posible que la evaluación del impacto posibilite la aparición de fenómenos de reduccionismo a indicadores. Es por ello que puede resultar conveniente la identificación de impactos diferenciales sobre cada territorio o sobre cada colectivo humano concreto.

4.- La situación

En este punto es habitual encontrar en la literatura que las referencias a la evaluación de la situación se considera equivalente a la evaluación de las necesidades. Preferimos referirnos a la situación y no a las necesidades porque estimamos que estas pertenecen a un nivel de elaboración teórica que se encuentra un paso más adelante que la situación. Efectivamente, hacer referencia a las necesidades es tanto como haber identificado las carencias, o las ausencias

que se perciben en un determinado colectivo o situación, se definan como se definan esas carencias (con criterios normativos, comparativos, expresados o experimentados, en terminología de Bradshaw). Lógicamente, para poder iniciar un programa es necesario identificar necesidades; sin éstas el programa de intervención no se legitima.

Sin embargo nuestra intención al plantear necesidad de evaluar la situación es la de identificar los elementos dinámicos del sistema social de referencia y sobre todo responder a una pregunta que resulta fundamental: ¿qué ocurre si no se interviene? o ¿qué ocurre si no se interviene con ese nivel de profundidad? Hemos de considerar que, en muchas ocasiones la intervención externa inhibe respuestas de los miembros del colectivo destinatarios en la dirección de la satisfacción de las propias necesidades. En realidad ninguna intervención debería considerarse justificada sino es después de haberse planteado sistemáticamente qué se prevé que ocurra en caso de no intervención (diagnóstico de no intervención)

5.- El diseño.

Se aplica a la formulación del programa y sirve para determinar el grado de idoneidad y de factibilidad de las acciones propuestas para resolver la situación problemática. Se han desarrollado múltiples formularios de preguntas y *check list* para realizar de forma más o menos rutinaria la evaluación del diseño que resultan útiles aunque no sean suficientes. ¿cómo se relacionan unas actividades con otras? ¿cómo se aprovechan los recursos de forma sinérgica? ¿cómo contribuyen diferentes actividades a un mismo objetivo? ¿de qué manera los resultados de una actividad influyen en la ejecución de otra? etc.

6.- El desempeño.

Se propone medir la forma en que se está comportando el programa a fin de establecer comparaciones con las previsiones hechas en la fase de formulación. Se miden también los efectos no esperados o previstos.

Cuando se realiza de forma simultánea a la ejecución del programa se trata de identificar, no sólo las diferentes desviaciones (*gap*) respecto a lo previsto, sino también los resultados no previstos a fin de identificar oportunidades de mejora a lo largo de la ejecución del programa y de detección y anulación de los posibles efectos perversos en los momentos iniciales de los mismos.

Cuando se realiza al finalizar la ejecución del programa se trata de identificar la contribución diferencial de las actividades y recursos a la obtención de resultados, valorar la secuencia de ejecución, identificar las relaciones causales de los efectos no previstos, etc. En este sentido, la evaluación del desempeño constituye un

instrumento de pilotaje del programa que se complementa con los sistemas de control de gestión (que por otro lado pueden ser considerados como parte del sistema de evaluación).

7.- El aprendizaje

La experiencia en sí misma no vale nada. Lo que realmente es útil de la experiencia es lo que somos capaces de aprender de ella. En este sentido cada proyecto que realizamos o en el que participamos aporta un conocimiento a la persona o al grupo. Convertir la experiencia en aprendizaje constituye un proceso de capitalización de la experiencia. En algunas dinámicas vinculadas a la investigación-acción el aprendizaje constituye un objetivo de muchas de las fases del proceso. Incluso en programas de corte más tradicional algunos proyectos pueden llegar a tener como finalidad principal la creación de conocimiento.

A fin de facilitar la evaluación de los aprendizajes podemos identificar cuáles son los mecanismos de incorporación de conocimientos nuevos a la persona o al grupo:

1.- Incremento

Se compara la nueva información con la previamente disponible. Se añaden datos de conocimiento de forma coherente con los esquemas previamente existentes.

2.- Reestructuración

La nueva información no se ajusta a los esquemas disponibles o la actual estructura de los datos no es considerada satisfactoria. Los esquemas han de ser reestructurados.

3.- Ajuste

Los datos y los conceptos existentes se afinan para conseguir mayores cotas de precisión, facilitar la generalización o determinar una mayor especificidad.

4.- Creación de esquemas

4.1.- Generación de patrones: nuevos esquemas basados en patrones antiguos

4.2.- Inducción: construcción de nuevos patrones combinando patrones recurrentes antiguos.

¿QUIÉN EVALÚA?

1.- Evaluación interna.

Decimos que la evaluación es interna cuando los responsables de llevarla a cabo son las personas responsables de la ejecución del proyecto o programa.

2.- Evaluación externa

La evaluación es externa cuando los responsables e la misma son personas ajenas tanto al diseño como a la implementación y gestión del programa.

3.- Evaluación mixta

Cuando la responsabilidad de la evaluación recae simultáneamente en personas comprometidas con el diseño y/o gestión del programa y otras completamente ajenas al mismo.

4.- Evaluación participativa

Decimos que la evaluación es participativa cuando en ella aparecen como agentes los usuarios o personas pertenecientes a la población objetivo del programa. Para que sea participativa la evaluación, el rol de los usuarios ha de ser superior al de simples informantes, incluso informantes clave. Han de poder incidir en la fijación de criterios de valoración o en la aplicación de los mismos. Naturalmente, la participación contiene grados, dependiendo del protagonismo que asuman los implicados.

No existen criterios fijos para determinar la conveniencia diferencial de cada tipo de evaluación. En todo caso se trata de una cuestión relacionada con los efectos buscados: a mayor participación mayor implicación de los usuarios pero mayores restricciones en la toma de decisiones para los responsables institucionales de los programas. Asimismo siempre es necesario considerar el hecho de que la participación constituye un proceso que se aprende, es decir, que no es algo que simplemente ocurre por efecto de la voluntad sino que requiere una serie de competencias, habilidades y actitudes mínimas que garanticen el éxito en la participación.

¿CUÁNDO SE EVALÚA?

1.- Evaluación *ex ante*

Se lleva a cabo cuando todavía no ha comenzado la implementación del programa o proyecto. Habitualmente incorpora el diagnóstico de situación, el análisis de necesidades, y la evaluación del diseño.

Aunque algunos autores consideran que la evaluaciones ante o es una verdadera evaluación sino parte del proceso de planificación de los programas, consideramos que su inclusión facilita la creación de una necesaria mentalidad evaluadora en los agentes vinculados a la planificación y gestión de las intervenciones.

Al mismo tiempo su relevancia hay que buscarla en el hecho de que la evaluación *ex ante* constituye un instrumento que permite a los agentes implicados ponerse de acuerdo en el análisis de la realidad y en la adecuación de los mecanismos que van a ser activados para transformarla.

Complementariamente ha de tenerse en cuenta que la evaluación *ex ante* puede incluir el análisis y la valoración de los impactos conseguidos en otros programas similares o referidos a la misma población objetivo. Es por ello que supone aprendizaje y legitimación de los programas propuestos.

2.- Evaluación *in vía*

Se realiza de forma simultánea a la ejecución del programa en momentos predeterminados en el propio diseño de la intervención. Incorpora también el panel de indicadores de control de gestión.

La evaluación *in via* constituye un instrumento imprescindible de monitorización y pilotaje de los programas y proyectos. Permite el cuestionamiento de las actividades, de la eficiencia de los recursos que se han habilitado, de los objetivos operativos, de las secuencias de actividad propuestas, la identificación de las desviaciones respecto de las previsiones, etc.

3.- Evaluación *ex post*

Se plantea a la finalización del proyecto o programa. Su función es cuestionar todos los aspectos de la intervención, los que se refieren a los resultados y los que se refieren al proceso. Finalmente ha de ser capaz de identificar los elementos básicos del balance coste-beneficio.

Se trata de un tipo de evaluación capaz de cuestionar los objetivos tácticos y estratégicos de los programas.

4.- Evaluación diferida

En realidad se trata de una modalidad de la evaluación *ex post* pero realizada después de un periodo temporal más menos prolongado (habitualmente 6 meses, o entre 1 y 2 años).

Su funcionalidad hay que buscarla en las características propias de los sistemas sociales en los que, muy a menudo, los efectos de la aplicación de un programa de intervención no son visibles de forma inmediata sino después de periodos prolongados de tiempo. Así ocurre por ejemplo en todos los procesos en los que son relevantes los cambios de actitudes de las personas, el planteamiento de sistemas de participación social, que requieren entrenamiento y legitimación en el proceso, o en los planes de desarrollo local o comunitario, por ejemplo.

¿CON QUÉ SE EVALÚA?

Los instrumentos de evaluación son múltiples y en gran medida coinciden con las herramientas que pueden ser utilizadas asimismo en la investigación. En la práctica la evaluación es una modalidad de la investigación.

Pueden identificarse diferentes dimensiones a partir de las cuales clasificar los diferentes instrumentos de evaluación. Así puede considerarse hasta que punto el instrumento es intrusivo, en que medida facilita la participación de los destinatarios del programa, etc.

No es esta guía el lugar en el que se han de desarrollar los diferentes instrumentos y métodos de evaluación, pero podemos nombrar algunos de uso frecuente: cuestionarios, entrevistas (abiertas, semiestructuradas, estructuradas) Técnicas de Grupo Nominal, grupos de discusión, grupos focalizados, análisis documental, observación, listas de comprobación, estudio de casos, incidentes críticos, etc.

Es importante que los instrumentos sean capaces de ofrecernos combinaciones adecuadas de aspectos cuantitativos y cualitativos. En todo caso, la investigación cuantitativa debe incorporar siempre una investigación cualitativa que permita describir y comprender el contexto en el que se ha realizado la investigación evaluadora..

DISFUNCIONES A EVITAR EN LA EVALUACIÓN

- Evaluar sin referentes teóricos de ningún tipo. En ese caso todos los resultados se convierten en una simple colección de anécdotas. Si no se dispone de una teoría formal, al menos hay que construir una serie de hipótesis capaces de guiar nuestros razonamientos aunque sea de forma provisional
- Valorar sólo al sujeto o sujetos objetivo de intervención de manera independiente del contexto. Descontextualización.
- Valorar solo los conocimientos que el sujeto o sujetos han obtenido o dicen que han obtenido, como resultado de un programa de intervención independientemente de los comportamientos cotidianos (ejemplo: información sobre el fumar o cambiar la conducta de fumar; colectivos que hablan de participación pero en los que las personas de hecho no participan).
- Valorar sólo los comportamientos concretos observables sin considerar si el sujeto cuenta o no con información procesos o destrezas o las actitudes para realizar un comportamiento de forma consistente a través de situaciones diversas.
- Valorar sólo cuantitativamente dando prioridad a la cantidad de comportamiento y olvidándose de la cualidad de los efectos de los programas de intervención.
- Valorar para controlar y fiscalizar
- Valorar sólo desde dentro de una organización o equipo. Ello hace muy difícil tratar de buscar interpretaciones o resultados alternativos a los oficialmente presentados. No se evalúa desde fuera. La realidad es poliédrica y no puede observarse sistemáticamente
- Mal uso de la valoración de programas que se evidencia en: comunicación deficiente de los resultados a los potenciales usuarios, utilización de la información por parte de usuarios mal informados, etc.
- No evaluar los efectos perversos u otros efectos indirectos de los programas

- Demasiado énfasis en la validez interna y estadística en la estructura de los diseños de los proyectos de valoración, olvidando a menudo la transferibilidad o validez externa y significado práctico y la validez ecológico-social de los programas.
- Ceder ante los grupos de interés, rechazar toda crítica al sistema establecido y proponerse extraer el máximo provecho posible del sistema sociopolítico.
- Utilizar la evaluación como instrumento de legitimación de políticas

BIBLIOGRAFÍA DE REFERENCIA

Alvira, F.: *Metodología de la evaluación de programas*. Centro de investigaciones sociológicas, Cuadernos Metodológicos, nº 2. Madrid, 1991

Cembranos, F., Montesinos, D. Y Bustelo, M.: *La animación sociocultural: una propuesta metodológica*. Editorial Popular, 1988

Cohen, E. y Franco, R.: *Evaluación de proyectos sociales*. Edit. Siglo XXI, 1993

Espinoza, M.: *Evaluación de proyectos sociales*. Humanitas, Buenos Aires, 1986

Fernández – Ríos, L.: *Manual de psicología preventiva*. Edit. Siglo XXI, , 1994

Fantova, F.: *La gestión de organizaciones no lucrativas*. Editorial CCS, Madrid, 2001

García, G.y Ramírez, J.M.: *Diseño y evaluación de proyectos sociales*. Certeza, Zaragoza, 1996

Lopez Ceballos, P., Merlo, J. y García, H.: *Un método de evaluación formativa en el campo social*. Editorial Popular, 2001

Pérez Serrano, G.: *Elaboración de Proyectos Sociales. Casos prácticos*. Narcea, 1994

Pichardo, A.:*Evaluación del impacto social*. Lumen Humanitas, Buenos Aires, 1997

Piñero, A.: *Establecimiento de un sistema de evaluación de los servicios sociales. Indicadores*. Generalitat Valenciana. Conselleria de Treball i Seguretat Social, 1991

Vedung, E.: *Evaluación de políticas públicas y programas*. Instituto Nacional de Servicios Sociales. Madrid, 1997