

Dirección General de Relaciones Exteriores

Dirección General del Desarrollo

EuropeAid Oficina de Cooperación

Unidad Común de Evaluación

**METODOLOGÍA DE EVALUACIÓN
DE LA AYUDA EXTERIOR
DE LA COMISIÓN EUROPEA**



**BASES METODOLÓGICAS
DE EVALUACIÓN**

VOLUMEN 1

Ni la Comisión Europea ni ninguna persona que actúe en nombre de la Comisión es responsable del uso que pueda hacerse de la información presentada a continuación.

Existe una gran cantidad de información adicional sobre evaluación de la ayuda exterior de la Comisión Europea en Internet Europa: <http://ec.europa.eu/europeaid>

Luxemburgo: Oficina de publicaciones oficiales de las comunidades europeas

ISBN : 92-79-00680-0

© Comisión Europea, 2006

Se autoriza la reproducción con la condición de que se mencione la fuente.

Impreso en Francia, 2006.

Prólogo



La mejora constante de las prácticas de evaluación es esencial para la Comisión Europea, sobre todo en el ámbito de la ayuda exterior y el desarrollo. Al interés por rendir cuentas a los contribuyentes europeos del buen uso de los fondos públicos se suma la necesidad de extraer lecciones de las intervenciones pasadas. La Comisión Europea dispone al efecto de métodos de evaluación creíbles para poder garantizar la validez de las conclusiones y recomendaciones.

La importancia de la evaluación de los resultados de las políticas regionales y sectoriales, de los programas y de la eficacia de la programación quedó reafirmada en el marco de la reforma de la gestión de la ayuda exterior emprendida en 2000. La evaluación es ahora un instrumento esencial para el éxito de las políticas de ayuda exterior en su conjunto, y está integrada en el ciclo de programación.

Entre 2002 y 2005, la Unidad de Evaluación común a las Direcciones Generales EuropeAid, Desarrollo y Relaciones Exteriores de la Comisión Europea ha desarrollado un enfoque destinado a responder a las nuevas necesidades de evaluación definidas en 2000. Los resultados de este trabajo metodológico están disponibles en el sitio web de la Comisión y se resumen en 4 fascículos.

Estos documentos representan una ayuda útil para entender y dominar mejor la evaluación tal y como se ha puesto en práctica en la Oficina de Cooperación EuropeAid. Me complace constatar que la evaluación así aplicada ya constituye una referencia entre los donantes.

La iniciativa emprendida por la Comisión Europea se pone a disposición del público y será de especial interés para el conjunto de los actores implicados en el campo de la ayuda pública al desarrollo y de la cooperación con los países terceros.

Benita Ferrero Waldner

Síntesis de los comentarios del panel de expertos sobre el manual de evaluación de EuropeAid

Impresión general sobre el manual

Este manual es un conjunto muy completo de documentos que describe de manera bien detallada el proceso de evaluación y refleja un esfuerzo para presentar el proceso de evaluación de forma clara y estructurada y proporcionar indicaciones al fin de satisfacer las necesidades de los miembros de la organización. Asimismo el manual podría resultar muy útil para otros actores implicados de una manera o de otra en los procesos de evaluación de EuropeAid, tales como los consultores o los representantes de las autoridades nacionales de los países socios.

El manual es equilibrado, se le atribuye el mismo peso a todos los elementos principales y la atención que se presta a cada aspecto del proceso de evaluación es proporcional a su importancia dentro del mismo. La mayoría de las secciones que componen el manual tienen un tamaño similar: de una a dos páginas. El mismo trata con frecuencia conceptos complejos con una brevedad admirable. Al final de las arborescencias, algunas veces los textos son más largos, hasta unas cuarenta páginas, lo que sin embargo es comprensible ya que los textos más largos se encuentran al final de una línea de investigación.

El manual está dirigido a diferentes tipos de lectores y puede ser abordado de diversas maneras en función del lector y de sus centros de interés, lo que resulta útil y lo hace flexible y evolutivo.

El contenido del manual refleja en su mayor parte un consenso emergente dentro de la comunidad de la evaluación acerca de la manera de conducir una evaluación. Actualizado al año 2005 debería mantener su vigencia para EuropeAid durante varios años.

Nuestros comentarios deben ser interpretados a la luz de un trabajo de gran envergadura que refleja una perfecta comprensión de la evaluación así como las competencias fundamentales para transcribir estos conocimientos y comunicarlos a través del manual así como de una fuente de aprendizaje on-line.

Cuestiones de política

Algunos de los comentarios están relacionados con la política global de evaluación tal como se la refleja en lo que está escrito –o lo que no está escrito– en el manual. Hacerle enmiendas para tener en consideración nuestras observaciones dependerá de la posición que tome EuropeAid respecto de estas cuestiones de fondo.

Orientaciones y consejos prácticos sobre las evaluaciones conjuntas

A menudo se recomienda que los donantes unan sus fuerzas en materia de evaluación debido a varias razones. En primer lugar, el deseo de aliviar la carga del proceso de evaluación a los países en vías de desarrollo, ya que el personal responsable de los proyectos así como otros funcionarios dedicarían demasiado tiempo a las evaluaciones externas.

Otra razón es que los proyectos y programas a menudo son financiados y/o puestos en marcha de manera conjunta, lo que justificaría que su evaluación sea también de manera conjunta. El Grupo de Trabajo sobre la Evaluación del CAD/OCDE ha recomendado con firmeza que se debe mejorar la coordinación entre los Estados miembros.

Durante las dos o tres últimas décadas, la cooperación para el desarrollo ha cambiado, y si bien el apoyo presupuestario, la asistencia sectorial y la ayuda programas se han sumado a la ayuda proyectos, los denominados “nuevos” enfoques tienden a adquirir cada vez más importancia. Así mismo, los enfoques de cooperación más amplios son los más aptos para la cooperación entre donantes y son los que podrían producir un mayor efecto en los países socios.

No obstante ello, la cooperación entre donantes implica cierto compromiso. En la medida en la que cada donante tenga su propia manera de codificar y de regular el proceso de evaluación, toda evaluación conjunta requerirá un alto nivel de adaptación recíproca. El manual no brinda ningún consejo ni orientación sobre la manera en la que los responsables de la evaluación deben abordar las evaluaciones conjuntas. En el futuro, es probable que los responsables de evaluación deban gestionar cada vez más evaluaciones en colaboración con colegas de otras instituciones, por lo tanto se verán obligados a adaptarse. Es una tarea difícil, que plantea la cuestión de saber si determinados aspectos del

proceso de evaluación son menos “negociables” que otros, y de ser así, cómo se podría establecer una colaboración no obstante ello.

En nuestra opinión, el manual se beneficiaría si se lo completara con una sección suplementaria sobre la evaluación conjunta y sugerimos además que esta última no se describa como un problema o una complicación adicional sino como una oportunidad de reforzar el proceso de evaluación, de realizar economías de escala y de hacer que la cooperación con el país beneficiario sea más eficaz.

Enfoque de la evaluación participativa

Otro acontecimiento significativo en el campo de la investigación sobre la evaluación es la evaluación participativa. El concepto surgió a través de trabajos de investigación acerca de la utilidad de los resultados de la evaluación. Cuando la primera función de la evaluación es formativa, ha resultado particularmente útil implicar a los diversos participantes / partes interesadas en los procesos de definición de preguntas, desarrollo de métodos, recolección de información y elaboración de conclusiones. Cuanto más las partes interesadas sientan que se apropian de los resultados intelectuales del proceso, más probabilidades habrá de que actúen teniendo en cuenta las recomendaciones de la evaluación. El término de partes interesadas aquí hace referencia a las organizaciones implicadas en la cooperación para el desarrollo tales como las ONG, las consultoras, los poderes públicos a nivel central o descentralizado, etc.

Cuando otros actores toman parte en el proceso, a veces resulta necesario adaptarlo. El grupo de referencia descrito en el manual puede, por ejemplo, desempeñar otros roles y funciones; también se podrían crear órganos suplementarios. El equipo de evaluación podrá asumir otros roles y necesitar de competencias adicionales, quizás recurriendo a facilitadores del proceso más que a operadores. Al igual que para la evaluación conjunta, pensamos que el manual debería contener algunos consejos sobre las estrategias de evaluación participativa, alentando a los responsables de la evaluación a experimentar con los enfoques participativos e indicándoles cómo ajustar y adaptar el proceso a fin de fomentar la participación de otros actores.

Normas éticas

La comunidad de la evaluación le ha dedicado una gran atención a la calidad en materia de evaluación. Las “Normas de Evaluación de Programas” de 1994 (*Programme Evaluation Standards*) constituyen una de las publicaciones pioneras más influyentes sobre el tema. Aborda el tema a través de cuatro dimensiones: la utilidad, la factibilidad, la precisión y la oportunidad. El manual que hemos examinado trata en detalle sobre los aspectos del proceso de evaluación que contribuyen a la realización de evaluaciones factibles, precisas y útiles, pero contiene pocos elementos relativos a la ética en el campo de la evaluación.

Hay buenas razones para pensar que los temas éticos serán, a largo plazo, particularmente espinosos en la evaluación de la cooperación para el desarrollo. Las relaciones asimétricas de poder, el predominio de modalidades de pensamiento y de acción de tipo “donante – beneficiario”, los sistemas de incentivos a menudo perversos en torno a la ayuda, así como las diferencias trans-culturales hacen que la evaluación de la ayuda resulte difícil y esté sujeta a complejas elecciones éticas. El manual podría completarse con una sección sobre lo que significan las normas éticas y la razón por la cual dichas normas son importantes en la ayuda al desarrollo. Podría además describir algunos escollos éticos a los que los responsables y los equipos de evaluación pueden enfrentarse, y aquellos en los que algunas veces caen.

Integración con los sistemas conexos

Según lo comprendido, EuropeAid, en su calidad de organización, aplica un sistema de Gestión Basada en los Resultados (*Results Based Management*). La manera en la que este sistema en su totalidad, así como la toma de decisiones dentro del mismo, se relacionan con el sistema de evaluación no es clara. Existe una ambigüedad en si estos documentos están diseñados o no para trabajar con perspectivas de una gestión basada en los resultados (GBR). Ciertas secciones dejan suponer la aceptación de la GBR como marco conceptual para este documento, mientras que el contenido de otras partes sugieren lo contrario, en particular en las secciones sobre los indicadores y los niveles meta. Este punto exige a una atención particular.

Pensamos que además es necesario realizar un debate para establecer de manera explícita el vínculo que liga los documentos aquí presentados con las funciones de seguimiento y auditoría. No

se habla sobre los vínculos que serían necesarios para crear un sistema de seguimiento y de evaluación en un país, para los programas e incluso para los proyectos. De hecho, los documentos establecen distinciones entre la evaluación, el seguimiento y la auditoría, pero no se menciona nada acerca de las similitudes y los vínculos que podrían nacer a partir de una mayor coordinación entre estos enfoques. En el pasado, las evaluaciones de la UE a menudo se han visto obstaculizadas por la limitada utilidad de los datos de seguimiento. Sería una pena no aprovechar esta ocasión para coordinar mejor dichas actividades.

Observaciones finales

Un elemento central de la tarea asignada al grupo de expertos es la evaluación de la “calidad” de sus documentos. La “calidad”, sin embargo, es un concepto marcadamente abierto a interpretaciones. Existiría entonces:

1. una definición más limitada de la *calidad científica* y
2. una definición más amplia de la calidad como utilidad final del documento para sus destinatarios.

De acuerdo con la primera definición (1.), el trabajo realizado por EuropeAid es de buen nivel. Se pueden hacer algunas críticas pero, en general, tanto la exhaustividad de los textos como su precisión son apreciables.

Con respecto a la segunda acepción (2.), estimamos que queda bastante camino para recorrer. La “motivación” es casi tan importante como la “credibilidad” o la “diseminación”, y hasta ahora hemos visto pocos elementos relativos a la motivación de los responsables de la evaluación. El documento les proporciona herramientas para realizar su trabajo, pero no hace mucho para convencerlos de que se trata de un trabajo útil, creativo, interesante y que por lo general vale la pena realizar.

Según lo que comprendimos, muchos de los responsables de evaluación en las Delegaciones deberán asumir varias responsabilidades, con el riesgo de que la evaluación se transforme para ellos en una carga de trabajo adicional, lo que convendría minimizar. Aun si el sitio web no puede por sí solo cambiar esta situación, debería por lo menos dar testimonio de una cierta comprensión del problema.

Expertos

Kim Forss, Andante, Estocolmo

Kim Forss trabaja para Andante - tools for thinking AB. Se dedica a la investigación y a la consultoría en el campo de la investigación en materia de evaluación y políticas relativas a las relaciones internacionales y al desarrollo organizativo. Es la autora de las comunicaciones sobre la calidad y la utilización de los resultados de evaluación con la perspectiva de una mejor gestión democrática.

John Mayne, Ottawa

John Mayne es un consultor independiente en materia de desempeño del sector público. Anteriormente se desempeñó ante la Oficina del Verificador General de Canadá y de la Secretaría del Consejo del Tesoro de Canadá. Trabajó en profundidad sobre la apreciación de la calidad de las evaluaciones. Además ha publicado varias obras sobre evaluación de programas, una de las más sobresalientes consiste en un libro internacional sobre la evaluación y la medición del desempeño.

Christopher Pollitt, Universidad de Rotterdam

Christopher Pollitt cuenta con una larga experiencia en investigación y enseñanza en materia de evaluación y gestión pública, primero en Brunel University (Reino Unido) y en la actualidad en la Universidad de Rotterdam (Países Bajos). Forma parte de varias redes de investigación en estas áreas relacionadas con la OCDE. Presidió la Sociedad Europea de Evaluación. Ha publicado un análisis crítico de una serie de guías de evaluaciones.

Ray Rist, Banco Mundial – OED, Washington

Ray Rist trabaja actualmente como asesor metodológico para la OED (Banco Mundial). Anteriormente ha trabajado en el *General Accounting Office* de Estados Unidos. Cuenta con un conocimiento acabado sobre las prácticas de evaluación a nivel mundial.

Pierre Spitz, INRA, París

Pierre Spitz trabaja para el Instituto Nacional de la Investigación Agronómica (INRA) en París. Su área de idoneidad es el desarrollo rural. Además, ha trabajado para organismos de las Naciones Unidas en el transcurso de la última década.

ÍNDICE

I. INTRODUCCIÓN	11
II. PROCESO DE LA EVALUACIÓN.....	13
1 Objeto de la evaluación	14
1.1 Proyecto, programa, estrategia	14
1.2 Sectores, temas y aspectos transversales	16
1.3 Campo de la evaluación.....	18
2 Momento de la evaluación	19
2.1 Caso de una intervención puntual	20
2.2 Caso de un ciclo de intervenciones.....	21
3 Utilización de la evaluación	22
3.1 Los usuarios de la evaluación	22
3.2 Los modos de utilización.....	24
3.3 La difusión de la evaluación	26
4 Los actores de la evaluación y el papel que desempeñan	28
4.1 El responsable de la evaluación	28
4.2 El grupo de referencia	30
4.3 El equipo de evaluación.....	32
4.4 Las partes interesadas	34
4.5 Asistencia a la calidad	34
III. MÉTODO DE LA EVALUACIÓN	35
5 Estrategia de la intervención	35
5.1 Razón de ser de la intervención	36
5.2 Reconstruir la lógica de intervención	37
5.3 Políticas complementarias.....	40
6 Preguntas de evaluación	42
6.1 Utilidad de las preguntas	42
6.2 Origen de las preguntas	43
6.3 Seleccionar las preguntas	45

6.4	Preguntas y criterios de evaluación	49
6.5	Formular una pregunta de evaluación.....	51
7	Referencial de juicio	54
7.1	Criterios de juicio	54
7.2	Nivel meta (<i>target</i>).....	56
7.3	Indicadores	58
7.4	De la pregunta al indicador	62
8	Diseño metodológico	64
8.1	Introducción	64
8.2	Tabla de diseño para cada pregunta.....	65
8.3	Optimización del diseño global	69
8.4	Desarrollo de las herramientas	71
9	Recolección de información	73
9.1	Plan de colecta de información	73
9.2	Dificultades frecuentes y soluciones	74
9.3	Fiabilidad de los datos.....	76
10	Análisis.....	79
10.1	Estrategias de análisis.....	79
10.2	El desarrollo del análisis	82
10.3	Validez del análisis	86
11	Juicio	88
11.1	Conclusiones y lecciones	88
11.2	Recomendaciones.....	91
12	Garantía de calidad.....	92
12.1	Reglas del juego.....	92
12.2	Aprobación de los informes	95
	Glosario	97
	Disponible en el sitio web Europa	101
	Lista de cuadros	101

I. INTRODUCCIÓN

Perspectiva General

La Comisión Europea ha desarrollado y formalizado su metodología de evaluación de la ayuda exterior dando prioridad a los resultados e impactos para mantener la calidad de sus evaluaciones en un nivel que corresponda a las buenas prácticas internacionalmente reconocidas en la materia.

En el pasado, la evaluación de la ayuda exterior europea se concentraba a nivel de los proyectos. Esta ayuda metodológica se propone facilitar el paso hacia una práctica de la evaluación orientada hacia los programas y las estrategias. Está destinada principalmente a:

- los responsables de las evaluaciones en la sede de la Comisión Europea y sus delegaciones,
- los equipos de evaluación externos.

Además, se encuentra a disposición de todos los responsables de la ayuda exterior europea y de la comunidad profesional dedicada a la evaluación.

Esta metodología de evaluación está disponible en tres idiomas (inglés, castellano y francés) en dos formatos, para la lectura y para la navegación en Internet respectivamente.

La versión Internet comprende numerosos ejemplos y análisis en profundidad y se la puede consultar a través del sitio web de la Comisión Europea:

<http://ec.europa.eu/europeaid/evaluation/index.htm>

La versión impresa está compuesta por cuatro volúmenes. Este primer volumen presenta los conceptos fundamentales y como se articulan. El segundo volumen es una “Guía de evaluación geográfica y temática”. Trata sobre la evaluación de todas las acciones comunitarias a la escala de un país o una región, o del conjunto de acciones relativas a un sector o a un tema a escala global. El tercer volumen “Guía de evaluación de proyectos y programas” está relacionado con los proyectos de gran envergadura, los proyectos pilotos, los programas multi-país y cualquier otro proyecto o programa para el que se requiera una evaluación. El cuarto volumen “Herramientas de evaluación”

presenta las principales técnicas aplicables para realizar la estructuración de la evaluación, la recolección y el análisis de la información así como la ayuda para la formulación de juicios.

Bases metodológicas

Este volumen presenta lo esencial de lo que se desea saber a fin de gestionar o realizar una evaluación, con una terminología homogénea en tres idiomas (inglés, castellano y francés) y un vocabulario que contribuye a los esfuerzos de armonización emprendidos en el marco del Comité de Ayuda al Desarrollo de la OCDE (ver pág. 100).

Ofrece soluciones completas, concretas y comprobadas para todos los problemas, incluso los más difíciles de solucionar, tales como la aclaración de la lógica de intervención, la redacción de las preguntas de evaluación, el diseño metodológico o el análisis de impacto.

Está estructurado en dos grandes secciones:

- El proceso: objeto de la evaluación, momento de la evaluación, utilización de la evaluación, los actores de la evaluación y los papeles que desempeñan.
- El método: la estrategia de intervención, las preguntas de evaluación (utilidad, factibilidad, elaboración), las referencias de juicio (criterios e indicadores), el diseño metodológico, la recolección de la información, el análisis, el juicio (conclusiones, lecciones y recomendaciones) y la garantía de calidad.

II.PROCESO DE LA EVALUACIÓN

¿Qué es evaluación?

OCDE - CAD

- Apreciación sistemática y objetiva de un proyecto, programa o política en curso o concluido, de su diseño, su puesta en práctica y sus resultados. El objetivo es determinar la pertinencia y el logro de los objetivos, así como la eficiencia, la eficacia, el impacto, y la sostenibilidad para el desarrollo. Una evaluación deberá proporcionar información creíble y útil, que permita incorporar las enseñanzas aprendidas en el proceso de toma de decisiones de beneficiarios y donantes.
- El término "evaluación" designa también un proceso sistemático y objetivo posible por el cual se determina el valor y el alcance de una acción de desarrollo planeada, en curso o concluida.

Dg Budget

- Juicio sobre las intervenciones según sus resultados, impactos y necesidades que buscan satisfacer (*The Communication on Evaluation: SEC (2000) 1051*)

Los propósitos principales son los siguientes:

- Contribuir a la concepción de las intervenciones y aportar los elementos necesarios para la definición de prioridades políticas
- Velar por una utilización eficiente de los recursos
- Mejorar la calidad de la intervención
- Dar cuenta de los resultados de la intervención (*accountability*).

Los capítulos que siguen están dedicados al proceso de evaluación visto bajo los siguientes cuatro ángulos: el objeto de la evaluación (lo que es evaluado), el momento de la evaluación, la utilización de la evaluación, los actores y sus papeles.

1 Objeto de la evaluación

Resumen

Proyecto, programa, estrategia

La intervención evaluada puede ser simple, por ejemplo un proyecto cuya meta es producir un efecto dado sobre un público determinado. Contrariamente, la intervención puede ser compleja, por ejemplo una estrategia que produzca sus efectos a través múltiples intervenciones. Dependiendo del grado de complejidad de la intervención, las preguntas tratadas y el método de evaluación pueden variar.

Sectores y aspectos transversales

Las intervenciones públicas por lo general están clasificadas por sectores o áreas. Además, una intervención dada o un conjunto de intervenciones pueden ser evaluadas desde el punto de vista de un aspecto transversal.

Campo de la evaluación

El campo es objeto de las investigaciones de la evaluación. Su perímetro está definido conforme a varias dimensiones: geográfica, temporal, institucional o reglamentaria.

Un segundo perímetro más amplio precisa las acciones complementarias y los elementos del contexto a tener en cuenta.

1.1 Proyecto, programa, estrategia

El término « intervención » designa lo evaluado, es decir uno o vario(s) proyecto(s), programa(s), estrategia(s), o cualquier otra forma de ayuda exterior.

Evaluación de un proyecto

Un proyecto es una operación no divisible, delimitada en cuanto al calendario y al presupuesto, y muy a menudo se encuentra bajo la responsabilidad de un solo operador.

Los proyectos más simples de evaluar son homogéneos (ver Cuadro 1), es decir, conformes a :

- Una actividad principal con una implementación en un solo contexto para estar orientados hacia un solo público
- Un solo resultado directo esperado para dicho público
- Un único impacto esperado a un nivel más global.

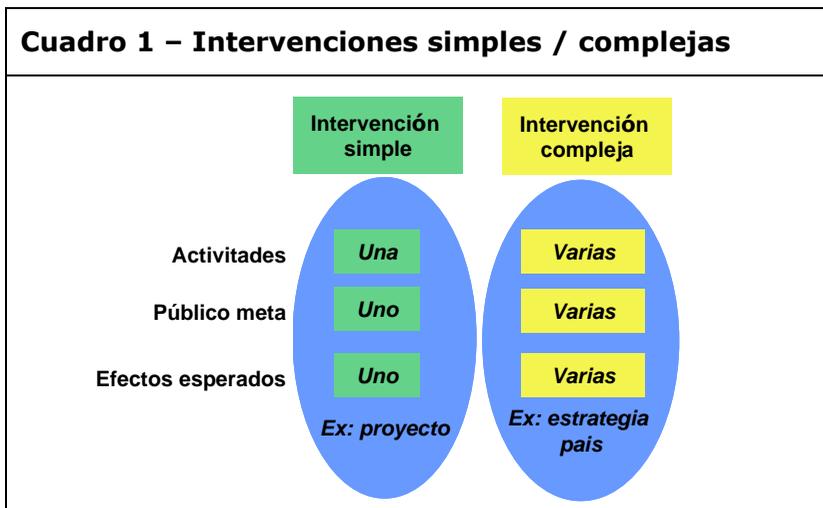
La evaluación de un proyecto simple se ve facilitada por el hecho de que se hace foco en una causa única (la actividad principal) y un solo efecto (resultado esperado para el público al que se dirige).

La evaluación de un proyecto privilegia las preguntas que interesan a los responsables, a los operadores y al público al que se dirige.

Evaluación de un programa

Un programa es un conjunto de intervenciones homogéneas, reagrupadas para alcanzar objetivos globales. Está delimitado en cuanto al calendario y al presupuesto, y muy a menudo se encuentra bajo la responsabilidad de un comité de pilotaje.

Un programa homogéneo está constituido por proyectos paralelos y similares entre ellos. La evaluación de un programa homogéneo puede realizarse por medio de síntesis de evaluaciones de proyectos o por extrapolación a partir de una muestra de proyectos.



Evaluación de intervenciones complejas

Las intervenciones complejas (por ejemplo una estrategia país) reagrupan intervenciones heterogéneas pero complementarias y puestas al servicio de un objetivo común. Se las concibe:

- Con varias actividades (subvenciones, inversiones directas, asistencia técnica, etc.) con puestas en ejecución en varios contextos (áreas geográficas diferentes, sectores diferentes) destinadas a varios públicos
- Con los resultados múltiples esperados para estos diferentes públicos
- Y con varios impactos esperados a nivel más global.

La evaluación de una intervención compleja no es ni la suma ni la síntesis de las evaluaciones de sus componentes. En efecto, se concentra en preguntas que son pertinentes en el nivel de los responsables políticos (efectos de sinergia, pertinencia de la repartición de los recursos entre los componentes, la contribución para llegar al objetivo global).

1.2 Sectores, temas y aspectos transversales

Las intervenciones públicas están comúnmente clasificadas por sectores. Además, se puede evaluar una intervención determinada o un conjunto de intervenciones desde el punto de vista transversal. En ambos casos, se habla también de una evaluación temática.

Sectores

Las intervenciones están clasificadas en un sector o en otro en función de la naturaleza de las actividades y de los productos.

Por ejemplo:

- Se clasifica una intervención que ayuda a formar a profesores de enseñanza primaria dentro del sector de la Educación
- Se clasifica una intervención que brinda asesoramiento y ayuda a agrupaciones de agricultores dentro del sector de la Agricultura.

Aspectos transversales

Se define un aspecto transversal (u horizontal) según los impactos y no las realizaciones, tal como lo muestran los dos ejemplos en el cuadro 2:

Cuadro 2 - Distinción de sectores y aspectos transversales		
	Ejemplo 1	Ejemplo 2
Realización	Profesores formados	Asesoramiento a las agrupaciones de agricultores
Sectores	<i>Educación</i>	<i>Agricultura</i>
Impacto	Las niñas tienen un mejor acceso a la educación	Nuevas prácticas de ganadería que limitan la desertización
Aspectos transversales	<i>Género</i>	<i>Medio ambiente</i>

Temas

Algunas veces se utiliza el término "tema" en el sentido de sector y otras en el sentido de aspecto transversal. Conviene entonces redefinirlo cada vez que se le emplea.

¡Atención!

Un mismo término puede designar de manera alternativa tanto un sector como un aspecto transversal. Por ejemplo, los derechos humanos pueden ser:

Un sector, cuando una evaluación se dirige a un proyecto de formación del personal de los tribunales

Un aspecto transversal cuando una evaluación examina los efectos de un proyecto de instalación de acceso a Internet en las zonas rurales pobres en relación al número de atentados a los derechos humanos.

1.3 Campo de la evaluación

El campo central de la evaluación corresponde al contorno preciso de la intervención sometida a la evaluación. Está definido de acuerdo con diversas dimensiones, tales como el territorio de referencia, el período considerado o el marco reglamentario, por ejemplo:

- Todas las financiaciones adjudicadas por la Unión Europea en América Latina desde 1998
- Las acciones de desarrollo local programadas en Rwanda en los últimos cuatro años.

Un segundo perímetro precisa las acciones complementarias y los elementos del contexto a tener en cuenta, sobre todo a los fines del análisis de la coherencia/complementariedad externa. Se trata del campo ampliado, contrariamente al campo central.

Delimitación del campo extendido

Las políticas complementarias

- Las intervenciones de las autoridades nacionales o de otros donantes que cubren los mismos territorios
- ... o que están dirigidas a los mismos públicos
- ... o que tratan los mismos problemas

Las co-financiaciones

- Las intervenciones co-financiadas con las autoridades nacionales
- Las intervenciones co-financiadas con otros donantes.

Otros elementos pertinentes de contexto

El campo central está especificado en los términos de referencia y el campo ampliado en el informe inicial.

2 Momento de la evaluación

RESUMEN

Caso de una intervención puntual

Se puede realizar una evaluación antes, durante o después de la intervención evaluada. Serán diferentes dependiendo del momento, las finalidades y su utilización.

La evaluación *ex ante* contribuye a garantizar la calidad del diseño.

La evaluación realizada durante o al final de la fase de ejecución apunta a extraer las lecciones de los primeros años de ejecución de la intervención y a hacer ajustes en el contenido de la intervención en curso según las realidades del terreno y/o la evolución del contexto.

La evaluación *ex post* se aboca principalmente a verificar los impactos obtenidos, a identificar y juzgar los impactos inesperados y a evaluar la sostenibilidad de los beneficios de la intervención.

Caso de una sucesión de intervenciones

Muchas intervenciones son organizadas por ciclos sucesivos y presentan una continuidad relativa de un ciclo al otro.

Estas constataciones llevan a escoger el momento de la evaluación considerando varios ciclos.

Es importante preocuparse del momento de la evaluación para

- Optimizar los recursos asignados a la evaluación optando por lanzar la evaluación en el momento en el que tendrá mayor valor añadido
- Responder ante las necesidades de los principales usuarios de la evaluación en el mejor momento
- Garantizar que una masa crítica de resultados y de impactos ya se haya materializado en el terreno y se preste a la recolección de información
- Evitar entrar en conflicto con ejercicios concomitantes de balance o auditoría.

2.1 Caso de una intervención puntual

2.1.1 Evaluación *ex ante*

Se realiza la evaluación *ex ante* antes de la adopción o antes de la ejecución de la intervención. Constituye una ayuda al diseño de la intervención. Aspira a mejorar la calidad del diseño de la intervención en la medida en que transpone las lecciones de las experiencias anteriores en el marco de la nueva intervención.

2.1.2 En el transcurso o al final de la ejecución

La evaluación realizada durante o al final de la fase de ejecución apunta a extraer las lecciones de los primeros años de ejecución de la intervención y a hacer ajustes en el contenido de la intervención en curso según las realidades del terreno y/o la evolución del contexto. A menudo comprende un balance de los productos acabados y un análisis de los primeros resultados directos e impactos obtenidos. Su intención es mejorar la intervención en curso y sus conclusiones pueden apoyarse en observaciones realizadas en el terreno.

2.1.3 Evaluación *ex post*

La evaluación *ex post* se realiza directamente después o mucho tiempo después de acabada la fase de ejecución. Está abocada principalmente a verificar los impactos obtenidos, identificar y juzgar los impactos inesperados y evaluar la sostenibilidad de los beneficios de la intervención.

Permite observar los cambios reales efectuados sobre el terreno, y si el plazo no es demasiado largo, analizarlos para estimar cuáles son atribuibles a la intervención.

La evaluación *ex post* a menudo tiende a rendir cuentas ante las instituciones que han asignado los recursos. Del mismo modo, ayuda a transferir la experiencia adquirida a otros países o sectores.

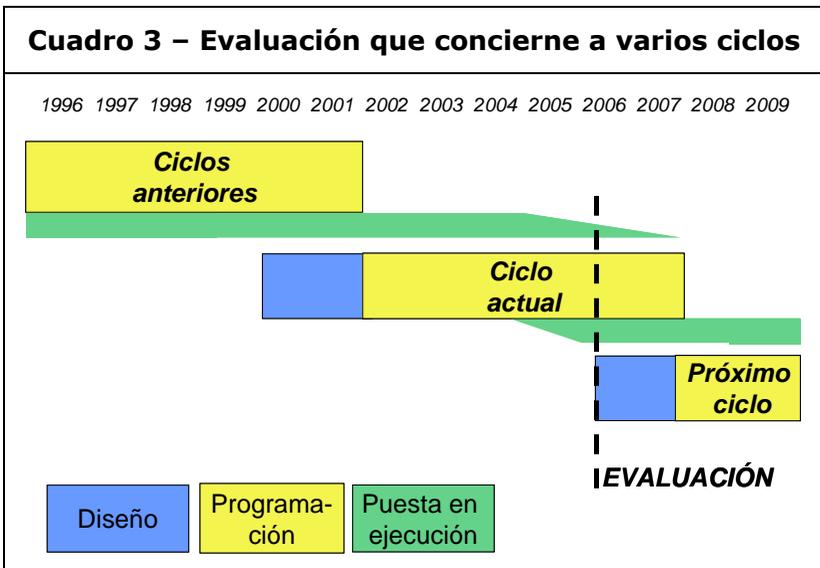
2.2 Caso de un ciclo de intervenciones

Muchas intervenciones están organizadas en ciclos sucesivos y presentan una relativa continuidad de un ciclo al otro.

A menudo se observa que un nuevo ciclo comienza incluso cuando la intervención del ciclo anterior no ha producido aún todos sus efectos.

Estos hallazgos llevan a elegir el momento de la evaluación considerando varios ciclos:

- Para extraer lecciones a partir de la observación de los efectos de las intervenciones de los ciclos precedentes
- Para alimentar la reflexión sobre la intervención del proximo ciclo
- Para observar la ejecución de la intervención del ciclo en curso y permitir una rápida retroalimentación.



3 Utilización de la evaluación

RESUMEN

Los usuarios de la evaluación

La evaluación está destinada a una variedad de usuarios tales como los responsables de tomar decisiones políticas y diseñadores de intervenciones, responsables y operadores encargados de la ejecución, socios, instituciones que hayan aportado fondos y ante las que se deba rendir cuentas, autoridades públicas que dirigen las intervenciones complementarias o similares, las organizaciones de la sociedad civil y las redes de expertos.

Los modos de utilización

Una evaluación puede ser utilizada para ayudar a tomar decisiones, para ayudar a formular juicios y / o para conocer y comprender la intervención y sus efectos.

La difusión de la evaluación

La difusión se refiere a la manera de dar a conocer el informe final de la evaluación, a todas las otras formas de dar a conocer los hallazgos, las conclusiones, las lecciones aprendidas y las recomendaciones. La difusión está dirigida a los servicios de la Comisión, las instituciones europeas, los actores externos, las redes de expertos, los medios de comunicación de masa y al público en general.

3.1 Los usuarios de la evaluación

La identificación de los usuarios de la evaluación es primordial al fin de optimizar la utilidad de la evaluación para los diferentes actores y en particular para:

- Asegurarse de que la evaluación responda a las expectativas de los usuarios previstos, en un formato adaptado a sus necesidades y de manera oportuna
- Asegurarse de que la evaluación tenga la credibilidad necesaria ante los usuarios previstos.

3.1.1 Responsables de tomar decisiones políticas y diseñadores

Los responsables de la toma de decisiones políticas y los diseñadores utilizan la evaluación para preparar el lanzamiento de nuevas intervenciones, la reforma de intervenciones existentes, la elección de las orientaciones estratégicas, las decisiones de asignación de recursos presupuestarios, humanos, etc.

Necesitan una información:

- Que sea directamente utilizable en el proceso de decisión
- Que llegue a tiempo
- Que responda a sus preguntas de manera clara, concisa y fiable.

Están interesados en lo que está en juego desde el punto de vista estratégico, en la coherencia externa y en los impactos globales, que constituyen la finalidad primordial de la intervención.

3.1.2 Responsables, socios y operadores

Los responsables tienen a cargo la ejecución y el seguimiento de la intervención, desde los servicios centrales hasta el terreno, los más próximos del público son los operadores. Los operadores del terreno pertenecen a la C.E. o bien a instituciones asociadas que compartan la responsabilidad de la ejecución.

Necesitan información lo antes posible para corregir las diferencias o validar los cambios.

Están interesados en los resultados directos de la intervención, en las necesidades y en las reacciones del público al que se dirigen, así como en las interacciones entre los socios.

3.1.3 Otros actores

Las instituciones que han aportado fondos esperan que se les rindan cuentas. Esta expectativa es válida tanto para el Parlamento como para el Consejo de Ministros, pero también para todos los co-financiadores. Los contribuyentes y los ciudadanos también son destinatarios de la evaluación.

Las autoridades públicas que dirigen intervenciones complementarias o similares son potenciales usuarios de la evaluación, en especial, bajo la forma de transferencia de

lecciones aprendidas. Lo mismo ocurre con las redes de expertos aludidos por la intervención.

Por último, la evaluación puede ser utilizada por los actores de la sociedad civil, en particular aquellos que representan los intereses de los públicos beneficiarios.

3.2 Los modos de utilización

Una evaluación puede ayudar a fin de:

- Tomar decisiones
- Formular juicios e informar la opinión pública
- Conocer y comprender mejor la intervención y sus efectos.

Estos tres modos de utilización no se excluyen entre sí: hay que comprender para poder juzgar y hay que juzgar para poder decidir.

3.2.1 Ayuda a la decisión

Se puede conducir la evaluación en beneficio de las personas que van a decidir o negociar un ajuste en la intervención evaluada. En este caso, se la utiliza para corregir la ejecución, para diseñar el siguiente ciclo de intervención, o para redefinir las orientaciones políticas.

La evaluación puede ayudar a la toma de decisiones de dos maneras diferentes:

- Formulando conclusiones de manera independiente y luego proponiendo recomendaciones
- Favoreciendo la implicación de los respectivos responsables de tomar las decisiones, o por lo menos de sus colaboradores cercanos, con el propósito de crear un fenómeno de apropiación y de retroalimentación (*feedback*) directa en el curso del proceso.

La asistencia a la decisión tiene diferentes formas, dependiendo del contexto de la decisión:

- Las recomendaciones pueden estar dirigidas a los responsables y/o a los diseñadores de la intervención
- Las recomendaciones pueden estar dirigidas a todos los actores y cofinanciadores de la intervención

- Por último, la evaluación puede estar concebida como una ayuda a la negociación y a la resolución de problemas entre las partes interesadas de la intervención, es decir, las instituciones responsables, los grupos de interés y los actores de la sociedad civil.

3.2.2 Ayudar en la formulación de los juicios

La evaluación puede ayudar a los usuarios a formar sus opiniones sobre los méritos de la intervención.

La formulación de una evaluación global es particularmente útil para rendir cuentas. En este caso, la evaluación examina los méritos de la intervención según los diferentes puntos de vista. Responde a las preguntas que son importantes para las instituciones financieras. El informe es accesible al público en general. Se pone de relieve tanto la independencia de la evaluación como la transparencia del juicio.

En esta instancia, se presta particular atención a la definición de los criterios de juicio (también llamados "criterios de evaluación razonada"). Sin embargo, el juicio en sí recién queda establecido definitivamente en el momento de la entrega del informe final y del debate sobre las conclusiones. La utilización de una evaluación para rendir cuentas supone entonces esperar hasta el final del proceso.

3.2.3 Conocer y comprender

Además de ayudar en la formulación de las decisiones y los juicios, que son por lo general las dos formas principales de utilización, la evaluación permite a los usuarios mejorar sus conocimientos de la intervención y entender mejor sus efectos. Indirectamente, contribuye a la transferencia de conocimientos así adquirida en beneficio de las redes profesionales que puedan tener un vínculo directo con la intervención evaluada.

Contrariamente a la retroalimentación (*feedback*), que alude directamente a los responsables de la intervención evaluada, la transferencia de las lecciones es un proceso indirecto que pasa por las redes de expertos tanto en el interior como en el exterior de la Comisión Europea.

La capitalización de los conocimientos a menudo comienza durante el proceso de evaluación, por intermedio de los expertos sectoriales implicados en el equipo de evaluación y el grupo de

referencia. Sin embargo, la transferencia de lecciones aprendidas comienza verdaderamente después de la entrega del informe final. La presentación de la evaluación en las redes especializadas es una etapa clave desde este punto de vista. Se realiza bajo la forma de seminarios o artículos técnicos.

3.3 La difusión de la evaluación

La difusión concierne al informe final de la evaluación, pero además a todas las otras maneras de dar a conocer los resultados, las conclusiones, las lecciones aprendidas y las recomendaciones. La difusión está dirigida a los servicios de la Comisión, las instituciones europeas, los actores externos, las redes de expertos, los medios de comunicación y el público en general.

3.3.1 Las disposiciones que hay que tomar

La difusión está prevista desde la etapa de los términos de referencia, anunciando en particular cómo se publicará el informe, y cuál será la función del equipo de evaluación en esta fase.

A lo largo del proceso, se tiene presente la calidad de la evaluación y de sus productos (informe, resumen, anexos) y la evaluación formal de la calidad al final del proceso.

Durante la última reunión del grupo de referencia, se identifican los principales mensajes pronunciados por la evaluación y los públicos a los que hay que llegar.

Tras la aprobación del informe, se completa el plan de comunicación eligiendo los mensajes a destacar y los canales de información mejor adaptados para cada público.

Por último, es importante obtener la cooperación necesaria para la instrumentación del plan de difusión y dividir los papeles a desempeñar entre el gerente de la evaluación, el equipo de evaluación (verificar que esta acción entre en su mandato y que esté mencionada expresamente en los términos de referencia) y los miembros del grupo de referencia y todo otro actor pertinente.

3.3.2 ¿Qué canales se utilizan para la difusión?

El informe de evaluación es difundido por Internet de manera que todo tipo de público pueda acceder al mismo. Sin embargo, se emprende una difusión más activa en el caso de públicos específicos:

- Se prepara un resumen de una a dos páginas acompañando el informe especialmente para la jerarquía del servicio responsable de la evaluación. El mismo señala las conclusiones y recomendaciones principales.
- Se envía el informe y/o su resumen a los servicios correspondientes y a los actores interesados.
- Asimismo se publica un resumen en las páginas web pertinentes, acompañado de un vínculo hacia el informe.
- Se difunde un resumen por Internet, destinado a la comunidad internacional de la ayuda al desarrollo. El mismo señala las lecciones aprendidas en la medida en la que sean transferibles.
- Se pueden redactar uno o varios artículos para el público en general o los medios especializados.
- Por último, el informe puede ser presentado en reuniones, talleres o seminarios.

Consejos prácticos para una evaluación

Una presentación breve (de 10 a 20 minutos) es suficiente para presentar lo esencial, pero se debe reservar más tiempo para las preguntas (20 a 40 minutos).

La presentación aborda los siguientes puntos:

- La intervención evaluada
- El método de evaluación
- Los mensajes resultantes de la evaluación
- Los límites metodológicos de dichos mensajes eventuales.

En el sitio Web: ejemplos de resúmenes y artículos sobre una evaluación

4 Los actores de la evaluación y el papel que desempeñan

RESUMEN

El responsable de la evaluación

El responsable de la evaluación (*evaluation manager*) es un miembro del servicio que está a cargo de la evaluación. Gestiona todo el proceso en nombre de la Comisión Europea.

El grupo de referencia

El grupo de referencia está presidido por el responsable de la evaluación y lo asiste en el pedido y la supervisión de la evaluación. Está compuesto por miembros de los servicios de la Comisión Europea y de los representantes del país socio cuando es posible.

El equipo de evaluación

El equipo de evaluación tiene la responsabilidad de la recolección y el análisis de datos así como de la formulación de juicios en respuesta a las preguntas de evaluación. Redacta el informe de evaluación del cual es responsable. Presenta periódicamente sus trabajos al grupo de referencia y al responsable de la evaluación y reacciona ante sus comentarios.

Las partes interesadas

Las partes interesadas (*stakeholders*) son los individuos, los grupos o las organizaciones que tienen intereses directos o indirectos en la intervención evaluada y en la evaluación en sí misma. Pueden o no estar afectadas por la intervención.

4.1 El responsable de la evaluación

El responsable de la evaluación (*evaluation manager*) es un miembro del servicio que está a cargo de la evaluación. Gestiona todo el proceso, de la A a la Z en nombre del comanditario de la evaluación.

El responsable de la evaluación se designa:

- Para asegurar la coherencia del proceso de evaluación, desde los términos de referencia hasta la difusión del informe y el seguimiento de las recomendaciones
- Para ser el referente administrativo y coordinar la actividad de los diferentes actores (grupo de referencia y equipo de evaluación)
- Para organizar, supervisar y evaluar la calidad en las diferentes fases del trabajo
- Para garantizar el buen desarrollo de la evaluación.

Lista de verificación: papel que desempeña el responsable de la evaluación

El responsable de la evaluación

- Constituye el grupo de referencia tras consultar con los jefes de los servicios aludidos
- Redacta los términos de referencia tras consultar con los miembros del grupo de referencia
- Procede a contratar el equipo externo de evaluación tras consultar con el grupo de referencia
- Ayuda al equipo de evaluación al acceso a la información y a los informadores
- Organiza el debate y la aprobación de las conclusiones y recomendaciones en interacción con el grupo de referencia y el equipo de evaluación
- Realiza el control de la calidad metodológica de los informes, que recibirá una doble verificación por parte de una segunda persona
- Difunde los resultados de la evaluación a los diferentes actores aludidos
- Interactuando con las autoridades competentes, realiza el seguimiento para verificar si se tienen en cuenta las recomendaciones / lecciones.

Se recomienda que el gerente de la evaluación reciba la ayuda del suplente en su trabajo diario, quien puede reemplazarle en caso de ausencia.

4.2 El grupo de referencia

El grupo de referencia es una interfaz entre el responsable de la evaluación y el equipo de evaluación externo.

El grupo de referencia da lugar a la expresión de una variedad de puntos de vistas relativos a la intervención evaluada.

Lista de verificación: papel que desempeña el grupo de referencia

Los miembros del grupo de referencia:

- Comentan los términos de referencia preparados por el responsable de la evaluación
- Representan y resumen los puntos de vista de los servicios de la Comisión y desempeñan una función de interfaz entre el equipo de evaluación y los servicios, complementando de esta manera los contactos bilaterales
- Se aseguran de que el equipo de evaluación tenga acceso a todas las fuentes de información y de documentación sobre las actividades emprendidas y las consulte
- Validan las preguntas de evaluación
- Comentan las notas e informes producidos por el equipo de evaluación. Los comentarios de cada miembro del grupo de referencia son recogidos por el responsable de la evaluación y luego transmitidos al equipo de evaluación
- Asisten en el proceso de retroalimentación (feedback) a partir de los hallazgos, conclusiones y recomendaciones de la evaluación.

Si la evaluación se lleva a cabo desde la sede, el grupo comprende:

- Los servicios de la Comisión relacionados con la intervención, incluyendo la Delegación afectada en el caso de una evaluación a nivel país / región
- Especialistas de un sector o de una cuestión transversal dentro de la Comisión
- La embajada del país asociado en el caso de una evaluación país.

Si la evaluación se lleva a cabo desde el país socio, el grupo se puede ampliar:

- A los servicios del gobierno
- A una selección de actores en materia de desarrollo
- A una selección de organizaciones de la sociedad civil
- A expertos.

La experiencia ha demostrado que es preferible que el grupo esté compuesto por un número limitado de miembros (entre 10 y 12) a fin de que su funcionamiento sea eficaz.

Tras haber identificado los servicios, instituciones u organismos a invitar al grupo de referencia, el jefe del servicio que gestiona la evaluación les envía una nota de invitación con el propósito de:

- Anunciar oficialmente el lanzamiento de la evaluación
- Solicitar la designación de una persona como miembro del grupo de referencia.

Cuadro 4 – Calendario y papel del grupo de referencia

Etapa	Reunión	Papel
Términos de referencia		Comentarios
Marco del equipo de evaluación		Comentarios
Preguntas de evaluación	1	Validación
Informe de inicio (<i>inception</i>)		Comentarios
Informe de documentación (<i>desk</i>)	2	Discusión Comentarios
Debriefing de la fase de terreno	3	Discusión
Proyecto del informe final, conclusiones y recomendaciones	4	Discusión Comentarios

En el sitio web: lista de verificación para organizar y animar una reunión del grupo de referencia

4.3 El equipo de evaluación

El equipo de evaluación es responsable de la recolección y del análisis de datos así como del informe, incluso de la formulación de juicios (también denominados “evaluación razonada”) y de la redacción de las conclusiones y recomendaciones. Interactúa con el grupo de referencia y el responsable de la evaluación. Le ofrece sus servicios al comanditario dentro del marco de un contrato.

Los miembros del equipo de evaluación externo son independientes de las organizaciones que han participado en el diseño y en la ejecución de la intervención evaluada.

Criterios para contratar al equipo de evaluación

- Conocimiento de los instrumentos de ayuda al desarrollo evaluados
- Conocimiento del campo de evaluación (sector, país, etc.)
- Presencia de expertos y de consultores locales
- Competencia gerencial del líder del equipo (organización, coordinación, diálogo con el cliente, etc.)
- Independencia del equipo con respecto a la intervención evaluada y ausencia de conflictos de intereses
- Competencia en materia de evaluación y, en particular, capacidad para dominar la formulación de juicios, redactar conclusiones y recomendaciones y realizar síntesis
- Dominio de herramientas de recolección y de análisis de la información
- Diversidad de los miembros del equipo en cuanto a perfiles y experiencia
- Precio.

El responsable fija el peso relativo de los criterios y lo especifica en los términos de referencia.

Asociar consultores locales

Cabe asociar consultores locales tanto como sea posible a fin de contribuir al desarrollo de una competencia local (construcción de capacidades nacionales de evaluación) además de beneficiarse de su buen conocimiento del terreno.

4.4 Las partes interesadas

Las partes interesadas (*stakeholders*) son los individuos, los grupos o las organizaciones que tienen responsabilidades y / o intereses directos o indirectos en la intervención evaluada. Pueden o no estar afectadas por la intervención.

Asociar a las partes interesadas en la evaluación sirve para:

- Tener en cuenta sus puntos de vista
- Aprovechar sus conocimientos técnicos y la información que poseen
- Interesarlas a la utilización de la evaluación.

4.5 Asistencia a la calidad

El responsable de la evaluación debe garantizar que las conclusiones tengan bases metodológicas y factuales sólidas.

En la Unidad de evaluación en la sede de la CE, los responsables poseen la formación y la capacidad necesaria para evaluar la calidad metodológica de los informes. Dos miembros de la Unidad realizan el seguimiento de todas las evaluaciones y participan en la evaluación de la calidad.

En el caso de las evaluaciones descentralizadas, el responsable utiliza los criterios de calidad propuestos en este sitio web, así como las indicaciones proporcionadas para completar la plantilla de evaluación de calidad, y si fuera necesario, recurriendo a la Unidad de evaluación.

En el sitio web: [plantilla de evaluación de la calidad y las indicaciones para completar la plantilla](#)

III.MÉTODO DE LA EVALUACIÓN

Los capítulos que siguen están dedicados a ocho componentes clave del método de evaluación presentados de modo lógico (y no cronológicamente).

Se detallan: la estrategia de intervención, las preguntas de evaluación (utilidad, factibilidad, elaboración), las referencias de juicio (criterios e indicadores), la concepción metodológica, la colecta de información, el análisis, el juicio (conclusiones, lecciones y recomendaciones) y la garantía de calidad.

5 Estrategia de la intervención

RESUMEN

La primera etapa del proceso de evaluación consiste en examinar la estrategia de intervención, es decir su razón de ser, su lógica y dónde se ubica respecto de las políticas complementarias.

Razón de ser de la intervención

La razón de ser de una intervención es satisfacer las necesidades, resolver los problemas o enfrentar los desafíos considerados como prioritarios en un contexto dado que no puedan ser tratados de manera más eficaz a través de una opción alternativa. Durante la evaluación, conviene recordar los principales elementos o bien reconstruirlos si los documentos de programación son imprecisos.

Lógica de la intervención

La lógica de la intervención identifica las actividades, productos, resultados y los diferentes niveles de impactos esperados. Es necesario recordar o reconstruir la lógica de la intervención para comprender cuáles son los efectos esperados y seleccionar las preguntas de evaluación a plantear.

Políticas complementarias

Es necesario identificar las políticas complementarias (otras políticas de la CE, del país socio, de otros donantes) y situar la intervención evaluada con respecto a ellas. Se trata de comprender dónde se encuentran las complementariedades, las potenciales sinergias, los riesgos de duplicación y las necesidades de coordinación.

5.1 Razón de ser de la intervención

La razón de ser de una intervención es tratar las necesidades, problemas o desafíos considerados prioritarios en un contexto dado, y que no pueden ser tratados de manera más eficaz por medio de una opción alternativa.

La razón de ser de la intervención ha de ser justificada en la etapa de la programación. En la etapa de la evaluación, conviene recordar los principales elementos o reconstruirlos si los documentos de programación no son precisos.

La evaluación identifica las razones por las que se eligieron las prioridades, por ejemplo: prioridades de la política en la que la intervención tiene lugar, urgencia de las necesidades a satisfacer, ventajas comparativas con respecto a las opciones alternativas.

El análisis procede en dos etapas de la siguiente manera:

Diagnóstico de problemas y respuestas

El informe de evaluación presenta de manera sucinta los siguientes elementos:

- Contexto de la intervención en el momento en el que fue iniciada
- Principales problemas (necesidades, desafíos) diagnosticados
- ¿Por qué se eligieron determinadas opciones estratégicas y no otras?

Ciertas preguntas de evaluación pueden tratar sobre los problemas abordados en la intervención, por ejemplo:

- ¿Había opciones alternativas para resolver los problemas identificados?

- Entre la fecha de la ejecución de la intervención y la de la evaluación, los problemas que habían justificado la intervención, ¿han cambiado de naturaleza o de amplitud?

Justificación de la intervención

Cada vez que sea pertinente, el informe de evaluación examina si las necesidades, problemas y desafíos podrían:

- Resolverse espontáneamente por medio de iniciativas privadas
- Resolverse de manera más eficaz por medio de otras iniciativas públicas.

Recomendaciones

Los documentos oficiales a menudo privilegian las opciones estratégicas que se retuvieron in fine. Es útil buscar a comprender cuáles fueron las alternativas.

5.2 Reconstruir la lógica de intervención

La lógica de intervención es el conjunto de actividades puestas en práctica, los efectos esperados (productos, resultados e impactos) de una intervención (ver Cuadro 8) así como de los supuestos que explican cómo las actividades van a conducir a los efectos en el contexto de la intervención.

Reconstruir la lógica de intervención es necesario para:

- Ayudar a aclarar los objetivos y a traducirlos en una jerarquía de efectos esperados de tal manera que ello permita su evaluación
- Sugerir preguntas de evaluación sobre dichos efectos
- Ayudar a evaluar la coherencia interna de la intervención.

La lógica de intervención puede ser "fidel" a los documentos de programación y a los documentos de origen de la política en la que se inscribe la intervención. En este caso, se deducen los efectos esperados de los objetivos establecidos en los documentos oficiales.

Al construir la lógica de intervención en el momento de la evaluación, se pueden tener en cuenta aquellos efectos que se

esperan implícitamente pero que no han sido evocados en los documentos iniciales. Eso significa un distanciamiento de la lógica "fiel", lo cual se debe especificar. El enfoque "fiel" es pertinente cuando se expresan los objetivos con la suficiente precisión y de un modo verificable.

Con frecuencia, la lógica de una intervención evoluciona a través del tiempo. Llegado el caso, ello justifica la reconstrucción de varias lógicas correspondientes a períodos sucesivos.

La lógica de intervención es una simplificación útil de la realidad, sin embargo, hay que estar siempre consciente de la complejidad del mundo real. Además de la reconstrucción de la lógica de intervención, resulta útil, identificar los principales factores externos que condicionan o que afectan la ejecución y los efectos. Es necesario también tener presente el hecho de que las explicaciones causales reales se revelan a menudo más complejas que los supuestos establecidos al inicio.

5.2.1 ¿Cómo reconstruirla?

Para reconstituir la lógica de intervención cabe proceder de la siguiente manera:

- Recabar y analizar los documentos oficiales a la base de la intervención
- Identificar las principales actividades
- Identificar los objetivos
- Traducir los objetivos para que adopten la forma de resultados e impactos esperados
- Conectar las actividades con los impactos esperados reconstituyendo relaciones de causa y efecto
- Verificar que las relaciones de causa y efecto sean lógicas, es decir, que se les pueda considerar plausibles a la luz de los conocimientos disponibles
- Debatir la lógica reconstruida con algunos informadores clave (diseñadores y responsables) y con expertos en el tema o país aludido
- Presentar y debatir la lógica de intervención en una reunión del grupo de referencia.

5.2.2 Las presentaciones más comunes

Marco lógico

Esta técnica consiste en producir una presentación en forma de matriz especificando los objetivos (primera columna), cómo se los deberá verificar (segunda columna: indicadores, y tercera columna: fuentes) y los principales supuestos acerca de los factores externos (cuarta columna).

Las líneas del marco lógico retoman una presentación simplificada de la lógica de intervención, en sólo cuatro líneas: actividades, resultados, finalidad (*purpose* = ventaja directa para el público al que está dirigida) y objetivo global.

Esta representación es conveniente para una intervención simple, tal como un proyecto o un programa bien acotado. En cambio, esta representación no logra captar correctamente la complejidad de una intervención heterogénea como lo es un programa de desarrollo rural integrado, una estrategia país o el apoyo a una política sectorial.

Debido a su simplicidad, el marco lógico permite detallar los indicadores que se deberán utilizar en cada línea, así como los factores externos.

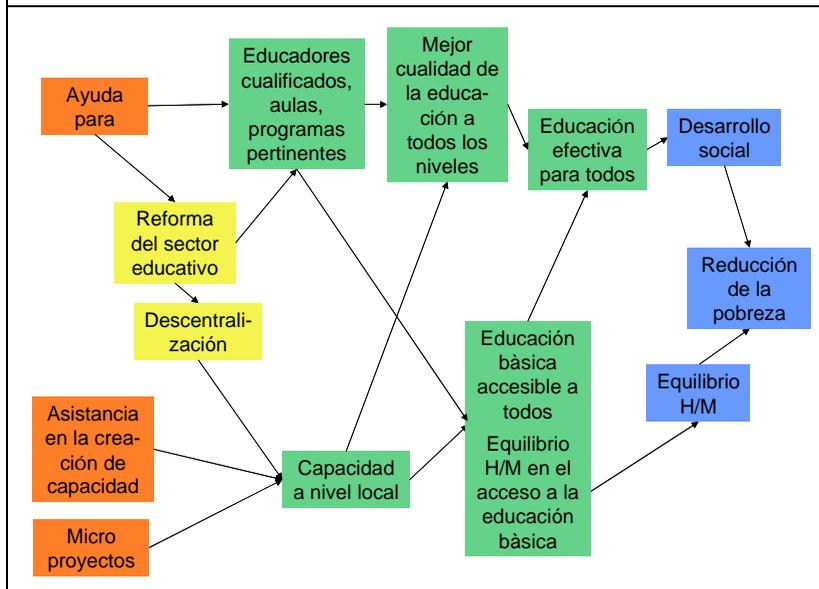
Diagrama de objetivos

Esta técnica (ver volumen 4) funciona a través de la identificación de objetivos oficialmente establecidos y una presentación gráfica que muestra las cadenas lógicas entre los objetivos, desde los más operativos hasta los más globales. La lógica de intervención está representada en forma de casillas y de flechas.

Una forma particular de representación consiste en el árbol de objetivos. Se aplica cuando cada objetivo de rango inferior contribuye a un solo objetivo de rango superior.

Diagrama de efectos esperados

Esta técnica (ver volumen 4) se asemeja al diagrama de objetivos ya que se apoya en objetivos oficialmente declarados. Sin embargo, los objetivos se traducen en efectos esperados antes de ser presentados en un diagrama. La traducción de los objetivos en efectos esperados permite trabajar con conceptos más concretos y más fáciles de verificar.

Cuadro 5 – Diagrama de los efectos esperados

Contrariamente al marco lógico, el diagrama puede tener cadenas de objetivos tan largas como sea necesario. Permite también poner en evidencia sinergias y relaciones complejas entre objetivos. Esta representación conviene en los casos de intervenciones complejas como pueden ser los programas integrados, las estrategias país o las políticas sectoriales.

Sin embargo, el diagrama no permite mostrar los factores externos, las condiciones y los riesgos como en la última columna de un marco lógico. Se definen estos elementos en la etapa de la concepción metodológica (ver Capítulo 8) teniendo como base un análisis de pregunta por pregunta.

5.3 Políticas complementarias

Se trata de identificar las principales intervenciones complementarias y situar la intervención evaluada con respecto a ellas.

El informe de evaluación expone de manera sucinta los siguientes elementos:

- Principales organismos nacionales e internacionales que intervienen en el mismo territorio, en el mismo sector y / u orientados al mismo público
- Principales políticas aplicadas por dichos organismos en el mismo territorio, en el mismo sector y / u orientadas al mismo público.

La identificación de las políticas complementarias es útil para:

- Ayudar a definir los criterios de juicio pertenecientes a las familias de la coherencia / complementariedad externa y de la pertinencia (ver sección 6.4)
- Si fuera necesario, examinar la calidad del diseño de la intervención, y, en especial, el hecho de que los objetivos se complementen con los de las otras políticas complementarias.

Algunas preguntas de evaluación pueden referirse a las políticas complementarias a fin de permitir una mayor profundización, por ejemplo:

- La intervención, ¿es coherente y complementaria con las políticas y las prioridades del país asociado? ... ¿con las otras políticas comunitarias?
- La ejecución, ¿estuvo sujeta a una coordinación con las acciones de otros donantes?, y la complementariedad, ¿fue mejorada gracias a esta coordinación?
- ¿En qué medida hubo un valor añadido por el hecho de conducir la intervención a nivel de la CE en vez de los Estados Miembros?

6 Preguntas de evaluación

RESUMEN

Utilidad de las preguntas

Las preguntas sirven para concentrar el trabajo de evaluación en una limitada cantidad de puntos clave a fin de permitir una recolección de información más acotada, un análisis más profundo y un informe más útil.

Origen de las preguntas

Las preguntas de evaluación pueden derivar directamente o indirectamente de la lógica de intervención. También pueden referirse a impactos inesperados, no figurando en la lógica de intervención, o sobre aspectos transversales (género, medio ambiente, etc.)

Las preguntas

Las preguntas (una decena como máximo) son seleccionadas según la utilidad potencial de las respuestas y su factibilidad. La elección de las preguntas es objeto de un debate al punto de la reunión de inicio.

Preguntas y criterios de evaluación

Las preguntas se clasifican en diferentes familias: los criterios de la OCDE - DAC (pertinencia, eficacia, eficiencia, durabilidad, impacto) y dos criterios propios de la CE (coherencia / complementariedad, valor añadido comunitario).

Preparación de una pregunta de evaluación

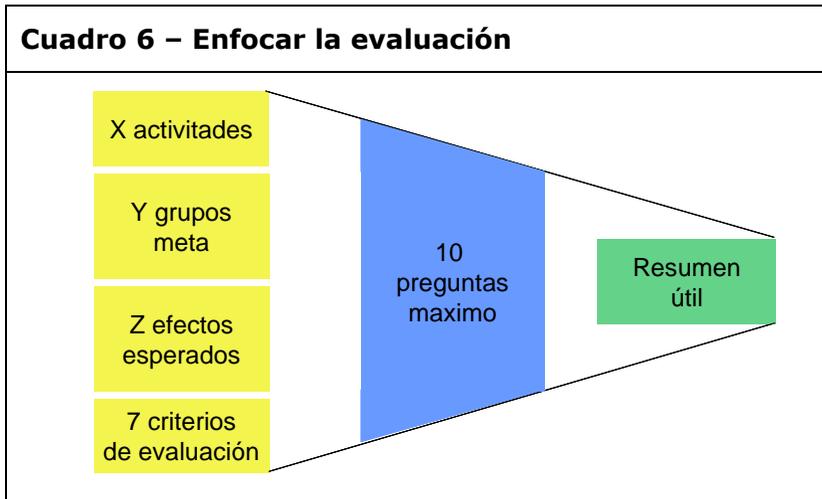
Es importante verificar que la respuesta a la pregunta sea útil, precisar la naturaleza de la utilización esperada, verificar que la pregunta dependa de una evaluación, precisar el campo de la pregunta, conectarlo otra vez a la lógica de intervención y al criterio de intervención.

6.1 Utilidad de las preguntas

Las preguntas sirven para concentrar el trabajo de evaluación en una cantidad limitada de puntos clave a fin de permitir una recolección de información más acotada, un análisis más profundo y un informe más útil.

Concentrar la evaluación en algunas preguntas clave es particularmente necesario cuando la intervención evaluada tiene múltiples dimensiones (ver cuadro 1). En dicho caso, si se desearan tratar todas las dimensiones de la ayuda y todas las dimensiones de la evaluación, el trabajo sería extremadamente largo y costoso o bien extremadamente superficial. Por lo tanto es necesario elegir.

Como lo muestra el siguiente cuadro, la elección de las preguntas ayuda a optimizar la utilidad de la evaluación.



6.2 Origen de las preguntas

6.2.1 Preguntas deducidas directamente de la lógica de intervención

La lógica de intervención es el conjunto de supuestos hechos para explicar de qué manera la intervención alcanzará sus objetivos y producirá los efectos esperados (ver Capítulo 5).

Una vez que el equipo de evaluación ha identificado los efectos esperados y los supuestos de causalidad que los vinculan, se puede entonces plantear todo tipo de preguntas del tipo:

- ¿En qué medida [*la actividad X*] ha contribuido a [*generar el efecto Z*] ?

- ¿En qué medida [las actividades X1, X2, X3...] han contribuido a [generar el efecto Z] ?
- ¿En qué medida [las actividades X1, X2, X3...] han contribuido a [generar los efectos Z1, Z2, Z3, ...] ?

Estas preguntas derivan directamente de la lógica de intervención en un formato estandarizado y pertenecen a la familia de la eficacia (obtención de uno o varios efectos esperados).

6.2.2 Preguntas indirectamente deducidas de la lógica de intervención

Cada pregunta de formato estandarizado puede ser declinada en múltiples preguntas:

... precisando el campo, por ejemplo:

- ¿En qué medida la intervención [*y más específicamente la que se puso en marcha a través del instrumento A o del procedimiento B*] ha contribuido a generar el efecto Z?
- ¿En qué medida [*la coordinación con los otros actores del desarrollo*] ha contribuido a generar el efecto Z?

... o precisando el efecto correspondiente, por ejemplo:

- ¿En qué medida la intervención ha contribuido a generar el efecto Z [*para las categorías más pobres de la población*]?

... o cambiando el criterio de evaluación (ver Capítulo 6.4), por ejemplo:

- ¿En qué medida la intervención ha generado el efecto Z [*con una gran probabilidad de que los efectos perduren después de que la intervención misma se termine*]? (sostenibilidad).
- ¿En qué medida la intervención ha contribuido a generar el efecto Z [*con un costo limitado comparado con...*]?(eficiencia o costo-eficacia).
- Cuando la intervención tiene por objetivo obtener el efecto Z, ¿en qué medida ello [*corresponde a las necesidades de la población en cuestión*]?(pertinencia).
- Cuando la intervención tiene por objetivo obtener el efecto Z, ¿en qué medida ello [*es compatible o contradictorio con los objetivos de otras políticas de la Comisión Europea*] ? (coherencia / complementariedad).

- Cuando la intervención tiene por objetivo obtener el efecto Z, ¿en qué medida ello [*aporta un valor añadido con respecto a una intervención similar pero conducida por los estados miembros*]? (valor añadido comunitario).

6.2.3 Preguntas no deducidas de la lógica de intervención

Las preguntas que siguen a continuación no requieren un examen previo de la lógica de intervención porque se refieren a efectos que no figuran.

- ¿En qué medida [la actividad X, el instrumento A, el procedimiento B] han generado efectos inesperados? y, si es así, ¿en beneficio o detrimento de quién?
- ¿En qué medida la C.E. ha integrado [el aspecto transversal T] en el diseño y la ejecución de sus intervenciones?

6.3 Seleccionar las preguntas

La selección de las preguntas de evaluación se hace en varias etapas:

Hacer una primera lista de preguntas

Se propone una primera versión de las preguntas de evaluación sobre la base de:

- El análisis de la lógica de intervención
- El análisis de la razón de ser de la intervención
- Temas que justificaron la decisión de lanzar la evaluación
- Temas que hay que estudiar, como se establece en los términos de referencia
- Eventuales preguntas puestas en la evaluación *ex ante*.

En una segunda versión, la lista y la formulación de las preguntas además tienen en cuenta:

- Las cuestiones planteadas por los informadores clave al inicio de la evaluación
- Las expectativas de los miembros del grupo de referencia
- Los temas descubiertos por el evaluador.

Evaluar la potencial utilidad de las respuestas

Más tarde, conviene evaluar la potencial utilidad de las respuestas, bajo el supuesto de que la evaluación aportará respuestas de buena calidad. Para cada pregunta prevista, se trata de saber:

- ¿Quién debería utilizar la respuesta?
- ¿Cuál es la utilización esperada: conocimiento, negociación, decisión, comunicación?
- ¿No se conoce ya la respuesta?
- ¿No existe otro estudio (auditoría, balance) en vías de desarrollo y susceptible de aportar respuestas?

Si se somete la elección de las preguntas a un debate en reunión, tal vez sea interesante clasificarlas en tres categorías de potencial utilidad: más elevada, media y menos elevada.

Verificar que no haya un olvido grave

La experiencia muestra que los olvidos más perjudiciales para la calidad de una evaluación tienen que ver con:

- Las preguntas de eficiencia y de sostenibilidad
- Las preguntas que se refieren a efectos negativos, en particular si dichos efectos incumben a los grupos desfavorecidos
- Las preguntas correspondientes a los efectos a muy largo plazo.

Evaluar la factibilidad de las preguntas

Se debe examinar la factibilidad (evaluabilidad) de una pregunta, siempre que antes se haya examinado su utilidad. A tal efecto, conviene consultar:

- El servicio que gestione la intervención
- Uno o varios expertos en el tema
- Uno o varios profesionales de la evaluación.

Los límites más frecuentes de la factibilidad

- **Debilidades de gestión**

Los datos y los informes de control son inadecuados o poco fiables.

Las dificultades de gestión generaron conflictos que les limitan el acceso a ciertos informadores o distorcionan sus respuestas.

En el caso de una fuerte sospecha de prácticas ilícitas o ilegales, es preferible postergar la evaluación hasta más tarde y realizar una auditoría.

- **Un campo demasiado complejo**

Demasiados datos que recabar, personas con quienes hay que reunirse, y análisis a efectuar respecto del tiempo y el presupuesto disponibles.

- **Una actividad demasiado reciente**

La actividad todavía no ha producido efectos observables.

Las personas interrogadas todavía no han estabilizado sus opiniones.

- **Una actividad demasiado innovadora**

Es difícil definir los términos de la pregunta sin ambigüedad.

La habilidad técnica es insuficiente para entender los mecanismos de causalidad.

- **Una intervención demasiado marginal**

La actividad evaluada no alcanza la masa crítica que permitiría el análisis de su contribución.

- **Un impacto demasiado lejano**

Hay muchos factores externos y son tan influyentes que se vuelve imposible analizar la contribución de la intervención.

Si una pregunta es potencialmente muy útil pero difícil de tratar, conviene investigar si no hay otra pregunta similar que sea a la vez más fácil e igualmente útil. Por ejemplo, si una pregunta se refiere al impacto relativamente lejano o global, probablemente se podrá mejorar su factibilidad basándola en el impacto

inmediatamente antes en la reconstrucción de la lógica de intervención.

Debatir sobre la elección de las preguntas

La elección de las preguntas clave está sujeta al debate en la etapa de la reunión inicial.

La selección es incluso más exitosa cuando se han consultado los potenciales usuarios, quienes se apropiaron de las preguntas seleccionadas y no se ha censurado ningún punto de vista legítimo.

Razones para seleccionar una pregunta

- **Porque alguien la ha planteado**

La persona que solicita el tratamiento de una pregunta tiende a cooperar en la construcción de la respuesta y a apropiarse de las conclusiones.

Un actor plantea una pregunta con la intención principal de influir o incluso de contrariar la acción de otro actor. Se debe examinar con atención la utilidad potencial de este tipo de preguntas.

- **Porque es útil, por ejemplo:**

La intervención o uno de sus aspectos es innovador y varios actores esperan una validación.

Se ha de tomar una decisión y las conclusiones pueden llegar a tiempo para ayudar a tomar dicha decisión.

Se prevé un debate público y las conclusiones pueden llegar a tiempo para alimentarlo.

- **Porque no se conoce la respuesta,** una pregunta es inútil si:

Otra evaluación, una auditoría u otro estudio acaba de dar respuesta a la misma.

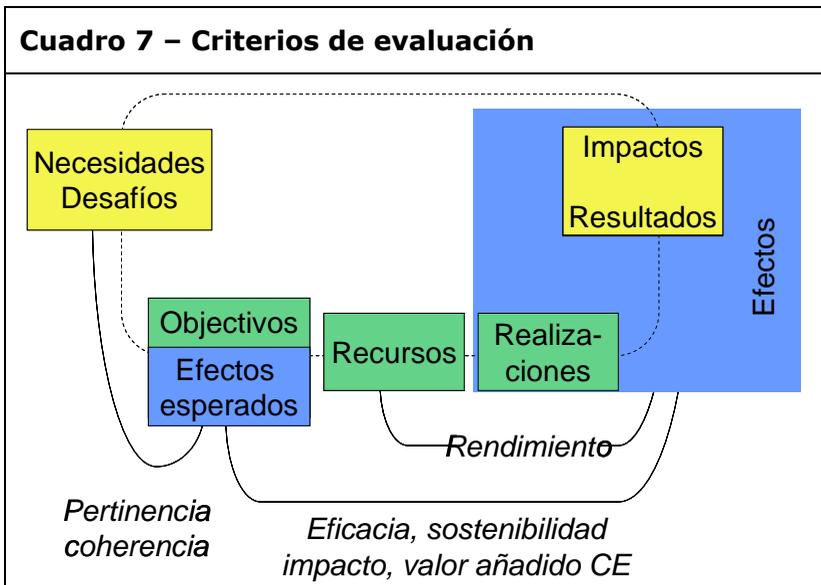
Ya ha sido planteada en muchas otras evaluaciones y siempre recibió la misma respuesta.

En el sitio web: ejemplos de preguntas a nivel país, a nivel regional, global, sectorial o temática

6.4 Preguntas y criterios de evaluación

Las preguntas se clasifican en diferentes familias que corresponden a diferentes "puntos de vista" sobre lo que se evalúa. Se han de considerar siete de estos puntos de vista, también llamados criterios de evaluación: pertinencia, eficacia, eficiencia, sustentabilidad, impacto, coherencia / complementariedad y valor añadido comunitario.

Los cinco primeros corresponden a la práctica tradicional de la evaluación de la ayuda al desarrollo formalizada por la OCDE (CAD). Los dos siguientes se aplican al conjunto de políticas de la CE.



Pertinencia

¿En qué medida los objetivos de la acción de desarrollo corresponden a las expectativas de los beneficiarios, a las necesidades del país, a las prioridades globales, a las políticas de los países socios y a las de los donantes?

- Ejemplos: ¿En qué medida la concentración de la ayuda en la educación básica corresponde a las necesidades del país socio?

Eficacia

¿En qué medida se alcanzaron los objetivos de la acción de desarrollo, o están por ser alcanzados, habida cuenta de su importancia relativa?

- Ejemplo: ¿En qué medida la ayuda ha contribuido a que haya un acceso equitativo a una educación básica de buena calidad?

Eficiencia

La medida en la que se obtuvieron los efectos deseados con el mínimo posible de recursos (fondos, habilidad técnica, tiempo, costos administrativos).

- Ejemplo: La puesta en práctica a través de la ayuda presupuestaria sectorial, ¿ha permitido obtener los mismos efectos que antes con costos mínimos de transacción para la CE y el país socio?

Sostenibilidad

¿En qué medida se perpetúan los beneficios resultantes de la acción de desarrollo una vez finalizada la intervención de la CE?, ¿o probablemente se perpetúen a largo plazo superando los riesgos?

- Ejemplo: ¿En qué medida la ayuda ha contribuido a una recuperación sostenible del retraso en el mantenimiento de la red vial?

Impacto

Efectos a largo plazo, positivos y negativos, primarios y secundarios, inducidos por una acción de desarrollo, directamente o no, intencionalmente o no.

- Ejemplo: Del punto de vista de los públicos afectados, las molestias ambientales ¿son aceptables en comparación con los efectos positivos de la intervención?

Coherencia/complementariedad

Este criterio puede tener varias dimensiones:

Coherencia interna del programa de desarrollo de la Comisión.

- Ejemplo: ¿Se puede decir que las actividades y los productos permiten alcanzar los objetivos de manera lógica? ¿Existen contradicciones entre los diferentes niveles de objetivos? ¿Existen duplicaciones entre las actividades?

Coherencia / complementariedad con las políticas del país asociado y con las intervenciones de otros donantes.

- Ejemplo: ¿Se puede decir que no existe superposición entre la intervención considerada y otras intervenciones del país asociado y / o otros donantes y Estados miembros en particular?

Coherencia / complementariedad con las otras políticas comunitarias.

- Ejemplo: ¿Existe convergencia entre los objetivos de la acción de desarrollo y los de las otras políticas comunitarias (comercio, agricultura, pesca, etc.)?

Valor añadido comunitario

¿En qué medida se suman las ventajas aportadas por la acción de desarrollo a las que habrían resultado exclusivamente de la intervención de los Estados miembros en el país socio?

- Ejemplo: En qué medida el hecho de compartir funciones entre la CE y los Estados miembros ha contribuido a optimizar el impacto de la ayuda ?

6.5 Formular una pregunta de evaluación

6.5.1 Verificar que la pregunta se refiera a la evaluación

Antes de redactar una pregunta, hay que asegurarse de que no incumba a la auditoría o al seguimiento.

Si la pregunta trata únicamente acerca del respeto de las reglas o de los procedimientos, es una auditoría más que una pregunta de evaluación.

Si la pregunta trata únicamente sobre el avance de los productos, es una pregunta de seguimiento.

6.5.2 Precisar el campo de la pregunta

El campo de la pregunta puede ser la intervención en su conjunto u otra dimensión más particular de su concepción o de su ejecución, por ejemplo:

- La intervención, ¿ha contribuido a generar el efecto Z como se lo esperaba? (el conjunto de la intervención).
- La participación de los actores en el diseño de la intervención, ¿ha permitido generar mejor el efecto Z? (modalidad de diseño).
- Las medidas tomadas para asegurar la coordinación con los otros donantes, ¿han permitido generar mejor el efecto Z? (modalidad de puesta en práctica).
- La elección de la ayuda presupuestaria sectorial, ¿ha contribuido a generar el efecto Z de manera más satisfactoria? (modalidad de financiación).

6.5.3 Vincular la pregunta a la lógica de intervención

¿Sobre qué efecto trata la pregunta? Por ejemplo:

- La intervención, ¿ha contribuido a generar el efecto Z como se esperaba?
- La intervención, ¿ha contribuido a generar el efecto Z para los públicos más pobres?
- La intervención, ¿ha aumentado las posibilidades de los públicos más pobres de obtener el efecto Z?

Si la pregunta trata sobre un resultado a corto plazo o un impacto específico, será evaluable con mayor facilidad que si trata sobre un impacto global. En el sentido contrario, una pregunta que trata sobre un impacto más global tendrá una utilidad mucho mayor para los responsables de la toma de decisiones a nivel estratégico.

6.5.4 Vincular la pregunta con un criterio de evaluación

La pregunta está redactada teniendo en cuenta un criterio de evaluación (Ej.: pertinencia, eficacia, eficiencia, sostenibilidad, impacto, valor añadido comunitario y coherencia / complementariedad, etc.), excepto si apunta únicamente a conocer o comprender.

6.5.5 Finalizar la redacción de la pregunta

Es preferible elegir una redacción abierta apelando a una respuesta matizada más que una pregunta cerrada que apele a una respuesta por un "sí" o por un "no". Además, la redacción de la pregunta debe ser simple y concisa. Por último, se aconseja agregar a título indicativo comentarios sobre todos o parte de los siguientes puntos:

- Precisiones sobre el campo de la pregunta
- Definición de los principales términos utilizados

En el sitio web: [ejemplos de preguntas para diferentes categorías de evaluación](#)

7 Referencial de juicio

RESUMEN

Criterios de juicio

Un criterio de juicio (también denominado "criterio de evaluación razonada") precisa un aspecto de los méritos o del éxito de la intervención evaluada. Permite responder de manera positiva o negativa a una pregunta de evaluación.

Nivel meta

Para formular un juicio explícito, se recomienda hacer referencia a un criterio de juicio y a un nivel meta a alcanzar, es decir, a una norma de desempeño, un umbral de éxito o una buena práctica que sirva de punto de comparación.

Indicador

Los indicadores especifican de antemano la información que será recabada. Ayudan así a concentrar el proceso de colecta.

La evaluación se concentra principalmente en los indicadores de resultado e impacto utilizados en el marco de las encuestas, bases de datos de gestión o fuentes estadísticas.

De la pregunta al criterio y al indicador

Este capítulo propone un enfoque para pasar de las preguntas de evaluación a los criterios de juicio y a los indicadores.

7.1 Criterios de juicio

Un criterio de juicio especifica un aspecto de la intervención evaluada que va a permitir apreciar su mérito o éxito. Si bien "criterio de juicio" es la palabra adecuada, una alternativa aceptable es "criterio de evaluación razonada".

¡Atención! No confundir los conceptos

Los criterios de juicio presentados en este capítulo.

Los criterios de evaluación: pertinencia, eficacia, eficiencia, sostenibilidad, impacto, valor añadido comunitario y coherencia / complementariedad (ver Sección 6.4)

Los criterios de apreciación de la calidad de los informes de evaluación (ver Sección 12.1).

Un ejemplo de pregunta es:

- ¿En qué medida la ayuda de la CE ha mejorado la capacidad del sistema educativo primario para recibir alumnos provenientes de grupos desfavorecidos sin discriminación?

Esta pregunta pertenece a la familia de la eficacia.

Como la mayoría de las preguntas de evaluación, ésta consta de dos partes:

- Lo que se juzga, "la ayuda de la CE ..."
- La manera de juzgar, " ... ha mejorado la capacidad del sistema educativo primario para recibir los alumnos provenientes de grupos desfavorecidos sin discriminación?"

Los criterios de juicio desarrollan y especifican la segunda parte de la pregunta, por ejemplo:

- Criterio 1: Capacidad del sistema de enseñanza primaria para recibir a los alumnos de la minoría étnica E de manera satisfactoria.
- Criterio 2: Capacidad del sistema de enseñanza primaria para recibir a los alumnos provenientes de zonas urbanas desfavorecidas de manera satisfactoria.

Para ser utilizado en la práctica, cada criterio de juicio debe estar acompañado de un nivel meta (*target*) a alcanzar y uno o varios indicador(es).

Recomendaciones

Siempre definir el criterio de juicio antes de seleccionar un indicador existente o otro nuevo. Esto es indispensable para aclarar los conceptos. Si se presta atención a los indicadores de manera precipitada, se corre el riesgo de quedar prisioneros de la información existente, incluso si la misma no es suficiente para responder a la pregunta planteada.

Hacer que el grupo de referencia debata la definición de los criterios de juicio de modo que se tenga en cuenta la diversidad de los puntos de vista con respecto a la intervención.

Pueden existir discrepancias sobre los criterios de juicio, por ejemplo, un mismo efecto puede tener una dimensión juzgada de manera positiva por determinados miembros del grupo de referencia y otra dimensión juzgada negativamente por otros. En este caso, hay dos opciones posibles: (1) elegir un solo criterio de juicio, aunque en ese caso hay que evitar sesgar la evaluación, o bien (2) elegir varios criterios, aunque traiga como consecuencia el aumento y la complicación del trabajo de recolección y de análisis de información.

Para optimizar la recolección y el análisis de información, lo ideal es definir una cantidad limitada de criterios de juicio por cada pregunta. Además, esta recomendación tiene en cuenta la capacidad que tienen los usuarios para absorber la información.

Justificar, eventualmente las diferencias que podrían existir entre los criterios utilizados para formular el juicio al final del proceso de evaluación y los identificados en la primera fase de la evaluación (*desk*).

7.2 Nivel meta (*target*)

El concepto de "nivel meta" es ampliamente utilizado en el contexto de la gestión pública para establecer un objetivo verificable o un nivel de desempeño a alcanzar. En el contexto de la evaluación, se lo utiliza en un sentido aún más amplio ya que es posible que la intervención evaluada sea juzgada respecto de

niveles meta no establecidos de antemano pero sí identificados de manera *ad hoc* en el momento de la evaluación.

7.2.1 Objetivo definido de manera verificable

El nivel meta a alcanzar puede aparecer en los objetivos de la intervención, siempre y cuando se hayan establecido de manera verificable. En ese caso en particular, el mismo indicador sirve para definir el objetivo, para hacer que el criterio de juicio sea operativo y para determinar el nivel meta.

- Ejemplo: la cantidad de docentes calificados y con experiencia por cada 1.000 niños en edad escolar primaria corresponde por lo menos 20.

7.2.2 Buenas prácticas comparables

El nivel meta se establece durante el curso de la evaluación con referencia a una buena práctica comparable a la intervención evaluada.

- Ejemplo: el acceso a la educación primaria con docentes calificados y experimentados es por lo menos tan satisfactorio como en el caso de W (buena práctica reconocida a nivel regional).

El procedimiento es el siguiente:

- Identificar una práctica comparable y reconocida por su calidad (intervención similar de la CE en otro país, intervención de otro donante, intervención en otro sector pero que utiliza los mismos instrumentos)
- Informarse sobre la práctica a comparar (es más fácil si ya ha sido evaluada)
- Garantizar que las condiciones del contexto sean lo suficientemente parecidas como para permitir la comparación
- Proceder a la comparación (esencialmente cualitativa)
- Discutir y validar la comparación con el grupo de referencia.

7.2.3 Mejores prácticas identificadas dentro de la intervención

Se establece el nivel meta al final de la evaluación con referencia a una buena práctica identificada dentro de la intervención evaluada.

- Ejemplo: en las zonas en las que se concentra la minoría étnica X, la cantidad de docentes calificados y experimentados por cada 1.000 niños en edad escolar primaria es comparable a la de la zona Z, considerada como una de las que posee un mejor desempeño entre las otras zonas rurales del país.

7.3 Indicadores

El recurso a los indicadores es necesario para:

- Recabar y tratar los datos bajo un formato que sea directamente utilizable en el momento de responder a las preguntas
- Ir a lo esencial en la colecta de información.

Un indicador cuantitativo está fundado en un recuento (Ej.: cantidad de docentes calificados y experimentados). El indicador básico es el recuento propiamente dicho. Puede ser utilizado para calcular indicadores derivados (razones, índices) tales como el costo por alumno escolarizado o la cantidad de docentes calificados y experimentados por cada 1.000 niños en edad de asistir a la escuela primaria.

Un indicador cualitativo adopta la forma de una declaración que deberá ser verificada durante la recolección de información (Ej.: la opinión de los padres es que sus hijos tengan la posibilidad de asistir a clases en una escuela primaria con un docente calificado y experimentado).

iCuidado!

Durante la evaluación se recogen también datos no estructurados, ya sea incidentalmente, o porque se emplearon herramientas tales como estudios de caso. Se trata de elementos de prueba que pueden tener una fuerza suficiente para fundar conclusiones, pero que no son indicadores.

7.3.1 Indicadores de evaluación

Los principales indicadores de evaluación son aquellos que están vinculados a los criterios de juicio y especifican qué información se debe recabar para emitir un juicio según dichos criterios.

Un indicador puede ser construido específicamente para una evaluación (indicador *ad hoc*) y medido, por ejemplo, en el marco de una encuesta. También puede ser extraído de bases de datos de seguimiento, de un marco de evaluación o de fuentes estadísticas.

Indicadores de evaluación y otros

- Cuando una pregunta de evaluación se refiere a un resultado o a un impacto esperado, conviene verificar si dicho resultado o impacto ha sido objeto de un monitoreo del desempeño. En este caso, el equipo de evaluación utiliza los indicadores y los datos correspondientes, lo que representa una ayuda considerable, sobre todo si se registraron datos de línea de base.
- El monitoreo del desempeño, sin embargo, puede resultar de poca o nula utilidad cuando una pregunta de evaluación se refiere a los aspectos transversales, los factores de sostenibilidad, los efectos no esperados, la evolución de las necesidades o de los problemas, la coherencia, etc.

7.3.2 Categorías de indicadores

Indicadores y el ciclo de intervención

Se utilizan indicadores en primer lugar para el análisis del contexto y la formulación de la estrategia de intervención, luego para el seguimiento de los productos y resultados y por último para la evaluación.

Indicadores de seguimiento y de medición del desempeño

Los sistemas de seguimiento y de medición de desempeño hacen uso de los indicadores derivados del diagrama de los efectos esperados (también llamados cadenas de resultados).

Los indicadores de seguimiento están relacionados sobre todo con los recursos y los productos. Los indicadores de desempeño se concentran principalmente en los resultados.

Indicadores y evaluación

Los indicadores de evaluación sirven para ayudar a responder las preguntas de evaluación específicas. Dependiendo de la pregunta, pueden guardar relación con las necesidades, los problemas y los desafíos que han justificado la intervención, o con el logro de productos, resultados e impactos esperados.

Indicadores globales y específicos

Los indicadores globales o contextuales se aplican a la totalidad de un territorio, población o grupo, sin hacer distinción entre los que están aludidos por la intervención y los que no. Se los extrae principalmente de los datos estadísticos. Se establece su definición fuera de la evaluación. (Un capítulo del volumen 4 trata esta cuestión).

Los indicadores específicos conciernen sólo al grupo o territorio realmente aludido. Los indicadores específicos permiten el monitoreo de los cambios que ocurren entre los afectados por la intervención. Algunas veces el equipo de evaluación los define y los prueba, por ejemplo, bajo la forma de una escala de calificación *ad hoc*. La mayor parte de estos indicadores provienen de las encuestas y de las bases de datos de gestión.

En el sitio web: cómo construir una plantilla de calificación

Indicadores y lógica de intervención

Los indicadores pueden servir a producir una información sobre los recursos, los productos, los resultados o los impactos.

Cuadro 8– Indicadores y lógica de intervención

Categoría	... dan información sobre:
Indicadores de recursos	Los medios financieros, los recursos humanos, materiales, organizativos o reglamentarios movilizados para poner en marcha la intervención.
Indicadores de actividades	Los procesos de ejecución y la gestión.
Indicadores de productos	Los productos y servicios entregados bajo la responsabilidad de los que gestionan la intervención. Para simplificar, se puede decir que los productos corresponden a lo que se compra con el dinero público.
Indicadores de resultados	<p>Los efectos inmediatos de la intervención para sus destinatarios directos. Un efecto es inmediato si el operador puede percibirlo fácilmente mientras está en contacto con el destinatario.</p> <p>Como los operadores los reconocen con facilidad, los indicadores de resultados directos a menudo pueden ser cuantificados de manera exhaustiva por el sistema de seguimiento.</p>
Indicadores de impactos	Las consecuencias más remotas y/o más indirectas de la intervención.

7.4 De la pregunta al indicador

El cuadro 9 da un ejemplo de paso de la pregunta al indicador. Hay que guardar al espíritu que una pregunta es generalmente declinada en varios criterios de juicio, ellos mismos declinados en varios indicadores.

En el ejemplo siguiente, la pregunta se refiere a una familia de criterios de evaluación: la eficacia

El criterio de juicio es derivado de la pregunta del modo siguiente:

- Muestra el tipo de éxito que se procura valorar planteando la pregunta
- Precisa varios conceptos tales como " sistemas educativos", " grupos desfavorecido" y "discriminación".

Cuadro 9 – De la pregunta al indicador: ejemplo

Pregunta	¿En qué medida la ayuda de la CE ha mejorado la capacidad del sistema educativo para acoger alumnos provenientes de grupos desfavorecidos sin discriminación?
Criterio(s) de juicio	Capacidad del sistema de educación primaria para acoger alumnos de la minoría étnica E con calidad satisfactoria.
Indicador(es)	Cantidad de docentes calificados y con experiencia por cada 1.000 niños en edad de asistir a la escuela primaria en áreas donde se concentra la minoría étnica E.

El indicador es derivado del criterio de juicio del siguiente modo:

- Describe de manera precisa la información que ha de ser buscada para juzgar según el criterio de juicio escogido
- El indicador es fiel al criterio de juicio ("capacidad del sistema de enseñanza primaria que acoge a los alumnos con una calidad satisfactoria ")

- Es cuantitativo, pero habríamos podido definir un indicador cualitativo, por ejemplo: "los padres interrogados confirman que sus niños tienen la posibilidad de frecuentar una clase primaria que beneficia de un profesor cualificado y experimentado".

7.4.1 Calidad de un indicador

Un indicador mide o califica con precisión el criterio de juicio o la variable que se busca observar. En su defecto, se puede recurrir a varios indicadores similares (*proxies*).

El indicador proporciona una información simple, fácilmente comunicable y comprendida de la misma manera tanto por el proveedor como por el usuario de la información. Es preciso, es decir que está asociado a una definición que no posee ningún término ambiguo. Es sensible, es decir que genera datos que varían significativamente si hay un cambio en lo que se observa.

A menudo se exige que los indicadores del desempeño y los niveles meta sean SMART (en inglés: Specific, Measurable, Attainable, Realistic and Timely): específicos, mensurables, alcanzables, realistas y oportunos. Lo más importante es que el indicador sea fiel al criterio de juicio.

Indicadores y efectos: ¡Cuidado!

El indicador empleado para la evaluación de un efecto no es una medida o una prueba de dicho efecto en sí mismo. El indicador sólo da información acerca de los cambios suscitados, los cuales pueden resultar de la intervención (efecto) o de otras causas.

El equipo de evaluación siempre debe analizar o interpretar el indicador a fin de estimar un efecto (ver Capítulo 10).

8 Diseño metodológico

RESUMEN

Se trata de establecer el método que permitirá al equipo de evaluación externo responder a las preguntas y llegar a un juicio global. El diseño de evaluación incluye una estrategia de recopilación y análisis de los datos, los campos seleccionados para los estudios, una serie de herramientas específicamente desarrolladas y un plan de trabajo.

Tabla de diseño por pregunta

La tabla de diseño explica como se va a responder a las preguntas de evaluación, inclusive la secuencia de razonamiento que va a vincular los datos, los hallazgos y las conclusiones. Declina : el campo de la pregunta, los criterios de juicio, los indicadores, los niveles meta, la secuencia de razonamiento lógico, la estrategia de análisis, los terrenos de investigación y las fuentes de información y de herramientas.

Secuencia de razonamiento

La secuencia de razonamiento lógico muestra como el equipo de evaluación proyecta responder a la pregunta. No se trata de hacer que la pregunta sea más compleja sino de ayudar a responderla.

Optimización del método

Se trata de finalizar la totalidad del método de evaluación para cubrir el conjunto de las preguntas de evaluación y disponer de una buena combinación de herramientas de evaluación, considerando el tiempo disponible y los recursos.

8.1 Introducción

El diseño metodológico consiste en establecer el método que permitirá al equipo de evaluación externo responder a las preguntas y llegar a un juicio global. El diseño de evaluación incluye:

- La cadena de razonamiento lógico a seguir
- Una estrategia de recopilación y análisis de los datos
- Los campos seleccionados para realizar las investigaciones

- Una serie de herramientas específicamente desarrolladas
- Un plan de trabajo.

El equipo de evaluación comienza a diseñar un método provisional a partir del momento en que redacta su propuesta para efectuar una estimación de los costes. En esta fase hay una hipótesis clave, que consiste en estimar hasta qué punto la evaluación podrá apoyarse en datos secundarios y en qué medida requerirá trabajos específicos de recopilación de información sobre el terreno.

Las líneas principales del método se trazan durante la etapa inicial, en función de las preguntas de evaluación, de los criterios de juicio, de los indicadores, de las herramientas de recopilación de datos y de la estrategia de análisis.

El método se perfecciona y completa antes de la fase de terreno, y se describe en su totalidad en el informe de la primera fase (*desk*).

El informe final incluye una presentación breve y precisa del método de evaluación, así como sus límites, si los hay. El método se describe de forma detallada en el anexo, incluyendo el diseño inicial, las soluciones a los problemas encontrados, el método realmente aplicado y sus límites.

El método de evaluación se diseña según un proceso iterativo a tres niveles:

- Un enfoque pregunta por pregunta que permite que el equipo de evaluación prepare las tablas de diseño con vistas a aportar una respuesta adecuada a cada una de las preguntas
- Un enfoque global que abarca todas las preguntas y que le permite al equipo de evaluación optimizar su método en conjunto, respetando las limitaciones de tiempo y de recursos
- Una serie de herramientas desarrolladas para recoger y analizar la información.

8.2 Tabla de diseño para cada pregunta

Una tabla de diseño explica cómo se va a responder a una pregunta de evaluación, incluso la cadena de razonamiento que va a unir los datos, los hallazgos y las conclusiones.

Se elabora una tabla de diseño para cada una de las preguntas.

Cuadro 10 – Estructura de una tabla de diseño

Pregunta	Texto de la pregunta
Comentarios	Motivos por los que se ha escogido la pregunta
Campo	¿Qué cubre la pregunta?
Criterio(s) de juicio	¿Cómo serán apreciados los méritos o los éxitos?
Indicador(es)	¿Cuáles son los datos que ayudarán a apreciar los méritos o los éxitos?
Niveles meta	¿Qué nivel o umbral puede ser considerado como un éxito?
Secuencia de razonamiento	Secuencia de razonamiento para responder a la pregunta: Informando los indicadores Analizando la información Formulando el juicio
Estrategia de análisis	Tipos de análisis que hay que emprender (ver Sección 10.1.1)
Áreas de investigación	Nivel en el que los datos serán recogidos y tratados (ver Sección 8.3.2)
Fuentes de información y herramientas	De dónde vendrán las informaciones (ver Sección 8.3.3)

En el sitio web: ejemplos de tablas de diseño

Esta tabla se va refinando progresivamente en versiones sucesivas:

- Versión preliminar adjunta al informe de inicio
- Versión(es) provisoria(s) preparada(s) durante la fase de documentación de la evaluación hasta optimizar progresivamente el diseño metodológico
- Versión final adjuntada al informe de la fase de documentación (*desk*).

8.2.1 Secuencia de razonamiento

La cadena de razonamiento describe todas las etapas que se desplegarán para responder a la pregunta. No se trata en absoluto de que la pregunta se vuelva más compleja. El razonamiento previsto es indicativo pero más vale aclararlo de antemano ya que:

- Así los miembros del grupo de referencia tienen la oportunidad de aportar opiniones y contribuciones específicas
- Todos los miembros del equipo de evaluación entienden por qué recogen y tratan los datos lo que implica trabajar de manera más eficaz, homogénea y coherente
- Los miembros del equipo que no están familiarizados con la evaluación reciben consejos útiles sobre los datos que han de ser recabados y tratados.

Etapas relativas a los indicadores

Pueden referirse a:

(1) el nivel o el estado actual de los indicadores, potencialmente desglosados por país, zona, grupo social, etc. Por ejemplo:

- ¿Cuál es el valor actual del indicador cuantitativo I en el ámbito nacional? ¿Y para las zonas /grupos objetivo?

(2) los cambios en los indicadores. Por ejemplo:

- ¿Las partes interesadas ¿han notado un cambio positivo en el indicador cualitativo J durante el período evaluado?

Etapas relativas al análisis

Están previstas para:

(3) confirmar las hipótesis sobre el éxito de la intervención y justificar una respuesta positiva a una pregunta. Por ejemplo:

- ¿El plan de ayuda de la CE incluye el compromiso de hacer el seguimiento de indicadores de desempeño relativos al efecto Z?
- ¿Se ha realizado dicho seguimiento?
- ¿Se han comentado los datos relativos al seguimiento con los socios donantes de forma periódica?
- ¿Han adoptado medidas los socios donantes a raíz de estas conversaciones?
- ¿El objetivo de estas medidas era la obtención del efecto Z?
- Etc.

(4) poner en tela de juicio las hipótesis sobre el éxito de la intervención y justificar una respuesta negativa a la pregunta. Por ejemplo:

- ¿Hay otros socios que hayan promocionado un seguimiento de los indicadores relativos al efecto Z?
- ¿Los actores no gubernamentales han contribuido a priorizar el tema de la consecución del efecto Z en el debate público?
- ¿Hasta qué punto los socios han influido las acciones en favor de los grupos desfavorecidos?

Etapas relativas a la aplicación de los criterios de juicio

Están previstas para:

(5) aplicar y ajustar los criterios de juicio en el contexto específico de la intervención. Por ejemplo:

- ¿Se remiten de forma espontánea las partes interesadas a los mismos criterios de juicio que los elegidos para la evaluación? De no ser así, ¿por qué motivo?

(6) aplicar o desarrollar niveles objetivo en el contexto específico de la intervención. Por ejemplo:

- ¿Qué territorios /grupos presentan un mejor desempeño en el país respecto de los criterios de juicio seleccionados? Entre ellos, ¿cuáles pueden ser legítimamente comparados con los territorios /grupos objetivo?

8.2.2 Áreas de investigación

El equipo de evaluación puede desear recopilar y tratar los datos relativos a la intervención en su conjunto o estudiar de forma más específica determinadas áreas, por ejemplo:

- Cada etapa se estudia por países e incluye una visita a cada uno de ellos, así como una nota por país
- Además de emplear estadísticas del ámbito nacional, el equipo de evaluación estudia una selección de distritos representativos respectivamente (a) del grupo/territorio objetivo y (b) de los grupos/territorios con mejor desempeño
- El equipo de evaluación elige cuidadosamente X proyectos que serán objeto de una investigación detallada para poder trabajar algunas de las subpreguntas.

8.2.3 Fuentes de información y herramientas

En la tabla de diseño se identifican las fuentes de información que habrá que utilizar, como por ejemplo:

- Estadísticas que incluyan indicadores de contexto disponibles en las bases de datos internacionales
- Bases de datos de gestión o de seguimiento
- Informes, revisiones, auditorías, evaluaciones, artículos, etc
- Declaraciones de las partes interesadas
- Declaraciones de expertos.

La mayoría de las veces, cada tipo de fuente de información se asocia a una herramienta determinada de evaluación, por ejemplo, para determinado grupo de partes interesadas se recurrirá a las entrevistas o a los *focus groups*, se obtendrá la opinión de los expertos mediante intercambios de correos electrónicos o un panel, etc.

8.3 Optimización del diseño global

Se trata de completar totalmente el método de evaluación para responder a las preguntas de evaluación y disponer de una buena combinación de herramientas de evaluación, teniendo en cuenta el tiempo y los recursos disponibles.

Puede ser necesario realizar varias iteraciones para que el equipo de evaluación optimice el diseño global, garantizando la máxima calidad para cada una de las preguntas.

Las iteraciones sucesivas merecen una reflexión pero no requieren mucho tiempo y se realizan esencialmente en el curso de la fase inicial (*inception*) y de documentación (*desk*).

8.3.1 Combinar herramientas y preguntas

El equipo de evaluación establece la lista de herramientas de evaluación previstas en las tablas de diseño. Cada herramienta es considerada entonces en función de su capacidad para ayudar a responder a varias preguntas.

El equipo de evaluación explica y justifica las principales decisiones técnicas que ha adoptado, así como las opciones alternativas, sus ventajas e inconvenientes y los riesgos asociados.

El equipo de evaluación aprecia si las herramientas escogidas pueden reforzarse mutuamente, por ejemplo:

- Una serie de entrevistas puede ayudar a identificar los documentos apropiados con vistas a un análisis documental
- Una serie de entrevistas puede ayudar a preparar un *focus group*
- Una serie de entrevistas puede ayudar a perfeccionar las preguntas que se van a plantear a una muestra de beneficiarios
- Se puede someter a examen una serie de monografías de estudio de caso ante un panel de expertos con miras a reforzar el análisis y obtener hallazgos más precisos.

8.3.2 Preparación del juicio global

El equipo de evaluación valora de forma transversal todas las tablas de diseño con vistas a preparar una síntesis final, es decir, un juicio global derivado de las respuestas a todas las preguntas de evaluación.

8.3.3 Distribución de los recursos

Al diseñar su método, el equipo de evaluación trata de distribuir sus recursos limitados de forma adecuada entre las preguntas.

Algunas preguntas merecen ser tratadas con herramientas costosas, tales como una encuesta con cuestionarios a los usuarios finales, varios estudios de caso, un *focus group*, etc., mientras que para otras bastará un simple análisis documental y algunas entrevistas en la CE y el gobierno del país socio.

También resulta razonable destinar recursos sustanciales a una cuestión que plantea problemas de factibilidad.

Es poco frecuente que una pregunta no se pueda evaluar de ningún modo. Los problemas de factibilidad provienen más bien de una acumulación de dificultades y de limitaciones. Al principio de la evaluación, es frecuente que se pueda enmendar una pregunta difícil para hacerla más evaluable. Esto puede hacerse, por ejemplo, limitando el ámbito de la pregunta o aplicándolo a un efecto menos alejado o a un efecto probable si el efecto real todavía no resulta observable. Una vez validada la pregunta, el equipo de evaluación debe diseñar un método apropiado y destinarle los recursos adecuados.

8.3.4 Coste y limitaciones de tiempo

El equipo de evaluación diseña sucesivas versiones del método que hay que seguir, hasta cumplir con los siguientes condicionantes:

- Que la implementación de las herramientas de evaluación se ajuste al calendario del proceso de evaluación
- Que el coste de las herramientas de evaluación (recursos humanos, costes técnicos, gastos de desplazamiento y de estancia) se ajuste al presupuesto global de la evaluación
- Que se cuente con la fuerza laboral cualificada suficiente para implementar las herramientas de evaluación de forma profesional.

8.4 Desarrollo de las herramientas

El Volumen 4 presenta una docena de herramientas frecuentemente empleadas, con ejemplos de utilización en el contexto de la evaluación de la ayuda al desarrollo.

Cada herramienta implica una fase preparatoria como la redacción de una guía de entrevistas, la elaboración y prueba de un cuestionario, etc. El desarrollo de las herramientas se hace respetando las siguientes etapas:

- Lista de preguntas y etapas de razonamiento que la herramienta debe ayudar a tratar
- Especificaciones técnicas para la utilización de la herramienta
- Riesgos previsible que pueden comprometer o debilitar la puesta en ejecución de la herramienta y la manera de administrarla
- Modalidades de informes dentro del equipo de evaluación
- Modalidades de informes en el informe final
- Responsabilidades dentro del equipo de evaluación en la puesta en ejecución de la herramienta
- Criterios de calidad y proceso de garantía calidad
- Calendario de ejecución
- Recursos asignados.

9 Recolección de información

RESUMEN

El plan de colecta de información

El equipo de evaluación reúne la información que ya se encuentra disponible (datos secundarios) con el objetivo de responder a las preguntas y aplica las herramientas de recolección a fin de obtener información nueva (datos primarios).

Las dificultades más frecuentes

Aun cuando se haya preparado el programa de recolección de información del terreno con cuidado, el equipo de evaluación puede llevarse sorpresas en el transcurso de su trabajo en el terreno como: el acceso a los informadores, la falta de cooperación de estos últimos, la diferencia cultural, la ausencia o la debilidad de una información buscada.

Fiabilidad de los datos recabados

Las encuestas de evaluación están sujetas a varios tipos de riesgos. Sesgos tales como los que se describen a continuación pueden poner en peligro la fiabilidad de los datos recabados como: sesgos de confirmación, sesgos de empatía, autocensura, etc.

9.1 Plan de colecta de información

El plan de colecta de información finaliza el trabajo de diseño metodológico (ver Sección 8.5).

Los datos (primarios) de encuesta son recabados con la ayuda de herramientas tales como: entrevistas, cuestionarios, focus groups, visitas, observaciones directas, etc.

Para evitar duplicaciones, costos inútiles y falta de interés de los informadores, conviene apoyarse tanto como sea posible en la información existente (datos secundarios), a saber: datos de gestión y de seguimiento, estudios e investigaciones relativas al área en cuestión, fuentes estadísticas publicadas, informes de evaluación y sus anexos, cualquier otro informe, etc...

Dicha información puede ser obtenida con un costo mínimo. La misma puede ayudar a aportar elementos de respuesta a una parte de las preguntas planteadas.

Si el equipo de evaluación olvida mencionar datos existentes y pertinentes, su trabajo será proclive a sufrir las críticas de las partes interesadas, que consideran que esta información es importante, y la credibilidad de la evaluación disminuirá.

Principales canales para identificar y recabar datos secundarios

- Los responsables, organismos de ejecución, operadores y socios
- Expertos en el área en cuestión
- Internet
- Oficinas de estadísticas y órganos de monitoreo
- Los científicos y la bibliografía profesional.

9.2 Dificultades frecuentes y soluciones

Incluso si el programa de recolección ha sido preparado con prudencia, el equipo de evaluación puede encontrarse con sorpresas al realizar su trabajo en el terreno. Las secciones sucesivas exponen los problemas que surgen con más frecuencia y el modo de resolverlos.

9.2.1 El acceso a los informadores

El proceso de muestreo resulta difícil:

- Decidir si un tamaño de muestra reducido puede arrojar resultados estadísticamente válidos o no.

Un informador no se expresa con libertad:

- Concentrar las entrevistas en los hechos más que en las opiniones
- Proponer que la información recabada permanezca anónima y explicar cómo se garantizará el anonimato.

Un informador se expresa de tal manera que parece sesgar sus palabras a propósito:

- Concentrar las preguntas en los hechos y no en las opiniones.

9.2.2 Brecha cultural

El informador o la fuente de información es accesible únicamente en el idioma local:

- El equipo de evaluación debería contar por lo menos con un miembro que hable el idioma local de manera fluida (tanto la traducción como la interpretación siempre generan pérdidas de información significativas).

Existe una gran distancia cultural entre el equipo de evaluación y la población encuestada:

- El equipo de evaluación debería contar con uno o varios miembros capaces de mitigar la brecha entre ambas culturas.

9.2.3 Carencia de datos o debilidad de los mismos

Una fuente de información resulta ser incompleta:

- En la medida de lo posible, extrapolar los datos que faltan y hacer verificaciones cruzadas de otras fuentes.

Una fuente de información resulta no ser fiable:

- En la medida de lo posible, comprender los sesgos, adaptar los datos y hacer verificaciones cruzadas de otras fuentes.

Recomendaciones

Una evaluación genera una sensación de incertidumbre, lo que hace que algunas partes interesadas se muestren renuentes a cooperar e incluso hostiles. En ese caso, hay que mantener una actitud positiva, insistir en el uso previsto de la evaluación, asegurar la imparcialidad, y concentrarse en los hechos más que en las opiniones.

Si una fuente de información no es accesible o si una técnica de encuesta no es manejable, entonces hay que cambiar el plan de trabajo de recolección de datos para recabar la misma información a través de otras fuentes.

Prestar la debida atención a los sesgos y a los riesgos de falta de fiabilidad. Esforzarse por entenderlos. Informarlos.

Evitar confiar en una única fuente de información para facilitar la verificación cruzada en la etapa del análisis.

9.3 Fiabilidad de los datos

Al recabar la información, el equipo de evaluación se enfrenta a varios riesgos de sesgos que pueden socavar la fiabilidad de los datos recogidos.

¿Por qué se deben considerar con cuidado los sesgos?

- Para mejorar la fiabilidad de la recolección de datos
- Para evaluar la calidad de la evaluación
- Para entender las limitaciones de las conclusiones extraídas de datos poco fiables.

Sesgo de confirmación

Este riesgo es una amenaza para todos los enfoques de la recolección de datos. Es el resultado de la tendencia a buscar pruebas congruentes con la lógica de la intervención en vez de buscar pruebas que puedan refutarla.

Al estar sujetos a este sesgo, tanto el equipo de evaluación como los informadores tienden a centrarse en los efectos esperados y sistemáticamente pierden de vista los factores externos, los efectos no esperados, los efectos negativos, las interacciones con otras políticas, las partes interesadas externas, las opciones de instrumentación alternativas, etc.

Autocensura

En determinadas instancias, los informadores pueden mostrarse reticentes a responder libremente a las preguntas, por la simple razón de que sienten que corren un riesgo. Tienden a expresar con rigidez los puntos de vista de su institución o de su jerarquía.

Tradicionalmente se combate este sesgo garantizando el tratamiento confidencial de las respuestas. El entrevistador debería

además hacer hincapié en las preguntas basadas sobre hechos y evitar recoger opiniones.

Estrategia de los informadores

Los que tienen un interés en la intervención pueden distorsionar los elementos de información que proporcionan con el objetivo de acercar las conclusiones de la evaluación a sus puntos de vista.

Se reducirá este riesgo en la medida en la que el plan de trabajo para la recolección de datos tenga en consideración a todas las partes interesadas y se hagan verificaciones cruzadas de varias fuentes de información.

Respuestas inducidas por las preguntas

Este sesgo al igual que los siguientes son frecuentes en las entrevistas y cuestionarios.

La manera y el orden en el que se plantean las preguntas puede generar un sesgo positivo o negativo.

Este sesgo será limitado al confiar el diseño y la comprobación de los cuestionarios a profesionales experimentados.

Sesgo de empatía

Al realizar las entrevistas, el equipo de evaluación debe crear una atmósfera cordial (empática), a fin de facilitar la colecta de información, conseguir un alto índice de respuestas y concluir rápidamente la encuesta.

Este fenómeno introduce un sesgo positivo sistemático en las respuestas, lo que tiende a sobreestimar los beneficios de la intervención y a subestimar el papel que desempeñan los factores externos.

Los entrevistados pueden no tener una opinión completamente preestablecida acerca de la pregunta que se les formula. Se esfuerzan en decidir en pocos segundos la respuesta que le darán al entrevistador o al llenar el cuestionario. Al hacer este ejercicio en una atmósfera artificialmente positiva, se reforzará el sesgo de empatía.

Muestra no representativa

Este sesgo puede ser preocupante si el equipo de evaluación genera datos cuantitativos a través de una encuesta por cuestionario. También debería ser considerado cuando se utilizan datos secundarios que surgen de una encuesta por cuestionario. En esta instancia, se debería verificar que la muestra de informadores encuestados sea lo suficientemente amplia y representativa de la población en su totalidad.

Sesgo de selección de muestras

Las personas que aceptan ser entrevistadas pueden no ser representativas del público meta en general. Se podría controlar este sesgo emprendiendo objetivamente una encuesta cualitativa especial de algunos "no encuestados", sabiendo que este ejercicio implica costos adicionales.

Recomendaciones

Sistemáticamente mezclar las preguntas positivas y negativas a fin de reducir el sesgo de empatía y el sesgo sobre las preguntas.

Dar una gran credibilidad al prometer confidencialidad a fin de limitar la autocensura del encuestado, y cumplir estrictamente con la promesa.

No depender nunca de una única categoría de partes interesadas (por ejemplo: responsables de programa, beneficiarios) a fin de reducir los sesgos estratégicos.

10 Análisis

RESUMEN

Estrategias de análisis

Se pueden prever cuatro estrategias de análisis: análisis del cambio, meta-análisis, análisis de atribución y análisis de contribución. Los tres últimos permiten responder a las preguntas relativas a las causas y los efectos.

Desarrollo del análisis

Una vez elegida la estrategia y recolectados los datos, el análisis sigue toda o una parte de estas cuatro etapas: tratamiento de datos, exploración, explicación, confirmación.

Validez del análisis

La validez es obtenida cuando las conclusiones y las lecciones derivan de los hallazgos de tal manera que garantizan la capacidad de transferencia (validez externa), los hallazgos derivan de datos sin ningún sesgo (validez interna), y cuando los datos recogidos reflejan los cambios o necesidades que han de ser evaluados sin sesgos (validez de construcción).

10.1 Estrategias de análisis

10.1.1 Análisis del cambio

Esta estrategia apunta a definir la evolución de los indicadores medidos y/o calificados a través de un periodo de tiempo. No busca establecer una relación de causalidad entre la intervención evaluada y el cambio.

Preguntas causales

Las preguntas causales se refieren a los efectos de la intervención evaluada.

Estas preguntas están destinadas a una observación del cambio, a una atribución del cambio observado a la intervención o a un análisis de la contribución de la intervención al cambio observado.

Las preguntas relativas a la eficacia y al impacto tienden a ser preguntas de tipo causal, ya que suelen relacionar la intervención evaluada (la causa) con sus efectos. Las preguntas relativas a la eficiencia y a la sostenibilidad son indirectamente preguntas causales puesto que primero se deben analizar los efectos reales antes de calificarlos como eficientes o sostenibles.

En general, las preguntas relativas a la pertinencia y a la coherencia no son preguntas causales.

10.1.2 Meta-análisis

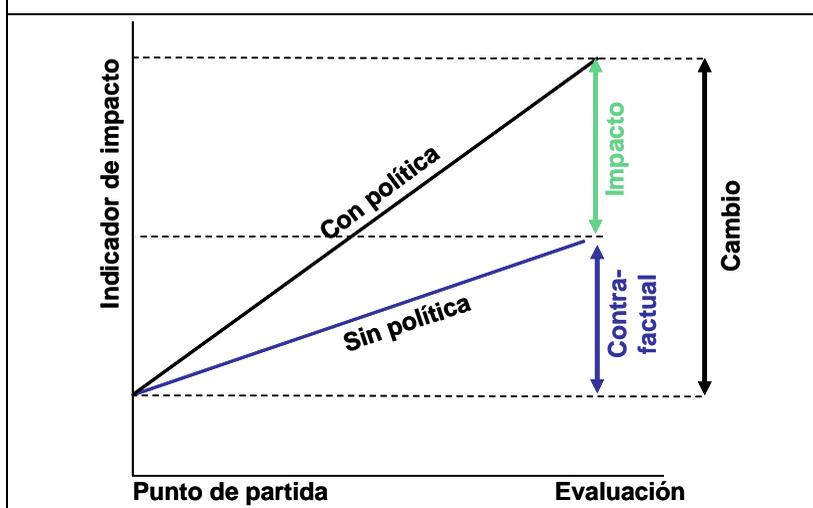
Esta estrategia extrapola los hallazgos de otras evaluaciones y estudios, tras haber verificado con cuidado su validez y transferabilidad.

10.1.3 Análisis de atribución

El objetivo del análisis de atribución es identificar la proporción de los cambios observados que son efectivamente atribuibles a la intervención que se evalúa. Implica la construcción de un escenario contrafactual (ver Cuadro 11).

La línea "con política" muestra los cambios observados, medidos con un indicador de impacto, entre el inicio del período evaluado (punto de partida) y la fecha de la evaluación. Por ejemplo: el empleo local ha aumentado, la tasa de analfabetismo ha disminuido.

La línea "sin política", también denominada contrafactual, es una estimación de lo que habría ocurrido de no haberse realizado la intervención. Puede obtenerse mediante una serie de métodos adecuados como los grupos de comparación o técnicas de modelización. El impacto se estima a partir del indicador "con política" tal y como se haya observado, sustrayendo la estimación de la evolución "sin política".

Cuadro 11 – Escenario contrafactual

Derivada de una estimación del enfoque contrafactual, la apreciación del impacto es en sí misma una estimación. Dicho de otra forma, los impactos no pueden medirse directamente. Sólo pueden deducirse de un análisis de los indicadores de impacto.

Tan sólo un enfoque contrafactual puede proporcionar una estimación cuantitativa del impacto. Cuando funciona, este tipo de enfoque ofrece un gran potencial de aprendizaje y de retroacción. Sin embargo, requiere un considerable número de datos y de recursos humanos.

10.1.4 Análisis de contribución

El objetivo del análisis de contribución es poner de manifiesto si la intervención evaluada es una de las causas de los cambios observados o no. Se basa en una serie de cadenas argumentales lógicas que se comprueban mediante un esmerado análisis confirmatorio.

Comprende las siguientes etapas sucesivas:

- Identificar la secuencia causa-efecto que establece el vínculo entre, por un lado, el diseño y la ejecución de la intervención y, por el otro, el efecto evaluado. Esta etapa se basa en las explicaciones disponibles relativas al ámbito

evaluado, procedentes del diagrama de efectos esperados realizado durante la primera fase de la evaluación, de las habilidades técnicas del equipo de evaluación y de los análisis exploratorios.

- Reunir elementos de prueba relativos a cada uno de los vínculos causa-efecto, como informes de estudios similares, relaciones causales declaradas por los entrevistados y otros elementos de prueba resultantes de encuestas en profundidad.
- Reunir elementos de prueba relativos a información alternativa (otras intervenciones, factores externos).
- Desarrollar una cadena argumental paso a paso, que establezca que la intervención ha aportado (o no) una contribución.
- Someter el razonamiento a una crítica sistemática, hasta que sea sólido.

10.1.5 Factores externos

Los factores externos se denominan también factores contextuales, exógenos o factores concurrentes; obstaculizan o amplifican los cambios esperados, siendo sin embargo independientes respecto de la intervención en sí misma.

En una evaluación determinada puede haber numerosos factores externos, por lo que es fundamental distinguir los más importantes.

Recomendaciones

No intentar identificar todos los posibles factores externos durante la fase de estructuración de la evaluación y aclaración de la lógica de intervención, ya que son sencillamente demasiados. Esta tarea sólo debería llevarse a cabo en el caso de una pregunta de evaluación dada.

10.2 El desarrollo del análisis

Es necesario efectuar un análisis para convertir los datos en hallazgos, que a su vez son objeto de un juicio para poder convertirse en conclusiones. El análisis se realiza principalmente

pregunta por pregunta, pero determinados análisis pueden abarcar varias preguntas.



10.2.1 El tratamiento de los datos



Los datos son procesados a través de operaciones tales como el cruce, la comparación, la agrupación, el listado, etc.

- El cruce es la utilización de varias fuentes o tipos de datos con el objetivo de establecer un hecho
- La comparación se realiza mediante la construcción de tablas, gráficos, mapas y/o clasificaciones. Los datos pueden ser comparados teniendo en cuenta una o varias dimensiones como el tiempo, los territorios, las categorías administrativas, las categorías socio-económicas
- La agrupación funciona asociando los datos en función de tipologías definidas
- El listado funciona identificando las diferentes dimensiones de una necesidad, un efecto, etc.

En esta etapa los hallazgos provisionales emergen del análisis. Las etapas siguientes tienen como finalidad profundizarlos y reforzarlos.

10.2.2 Exploración

El análisis exploratorio profundiza sistemáticamente en los datos recopilados para descubrir nueva información pausable:

- Nuevas categorías/tipologías
- Factores explicativos imprevistos
- Factores que puedan propiciar/obstaculizar la sostenibilidad
- Efectos fortuitos
- Nuevos supuestos causales.

El análisis examina el conjunto de los datos (cuantitativos y cualitativos) con el fin de identificar estructuras, diferencias, contrastes, similitudes y/o correlaciones. Por ejemplo, el equipo de evaluación efectúa:

- El análisis transversal de varios estudios de caso
- Comparaciones estadísticas que recortan bases de datos de gestión, bases de datos estadísticas, y/o los resultados de una encuesta por cuestionario

- Comparaciones entre las entrevistas, los cuestionarios y los documentos.

El enfoque es sistemático y abierto. Las técnicas de *brainstorming* resultan apropiadas.

10.2.3 La explicación

En esta etapa, se garantiza haber alcanzado una comprensión suficiente en cuanto a:

- Conceptos, categorías y tipologías definidas con precisión
- Explicaciones de causalidad plausibles.

Identificación de los factores externos clave y de las explicaciones alternativas.

10.2.4 Confirmación

El papel del análisis confirmatorio es verificar los hallazgos provisionales mediante una cadena de argumentos válidos y creíbles.

La finalidad del análisis es:

- Garantizar que los resultados sean de buena calidad y que puedan resistir cualquier crítica cuando se edite el informe
- Garantizar la credibilidad de los resultados desde el punto de vista de los usuarios finales
- Ver, en el caso específico de las preguntas causales, los efectos reales del cambio observado.

Para confirmar un hallazgo, se lo puede someter a la crítica por todos los medios posibles, por ejemplo:

- Si el hallazgo resulta de un análisis estadístico, ¿son concluyentes las pruebas de validez?
- Si la constatación ha sido sugerida por un estudio de caso, ¿existe otro estudio de caso que la contradiga?
- Si la constatación resulta de una encuesta, ¿puede responder a algún sesgo en dicha encuesta?

Fuerza de la prueba

Los hallazgos deben resistir a la crítica. Por ello, deben contar con el apoyo de las pruebas, es decir, los hechos convergentes, los documentos o las declaraciones. Se deben considerar cuatro niveles de fuerza:

- **Hecho observado**

Las pruebas basadas en hechos son las más fuertes. Los hechos observados pueden adoptar la forma de informes de visitas, fotografías, archivos de gestión o cualquier tipo de vestigio material.

- **Declaración de testigo**

Sigue siendo un tipo de prueba muy fuerte, por ejemplo, los beneficiarios directos declaran que han cambiado su actitud después de haber participado en el programa.

- **Sustitutivo (Proxy)**

Las pruebas fácticas indirectas también se denominan pruebas circunstanciales o sustitutivas, por ejemplo, durante los últimos meses, varios competidores de una empresa subsidiada quebraron, lo que indica que el nivel de ayuda fue excesivo y distorsionó la competencia.

La fuerza de este tipo de prueba depende de la fuerza del razonamiento lógico que respalda la inferencia.

- **Declaración informal**

El enunciado indirecto es el tipo de prueba más débil, por ejemplo, responsables de programas declaran que empresas beneficiarias han mejorado considerablemente su competitividad.

La fuerza de este tipo de prueba depende del grado de autoridad del experto de quien proviene dicho enunciado.

- Si el hallazgo se basa en una fuente de información, ¿se contradice con otra fuente?
- ¿Se refiere el hallazgo a un cambio que podría explicarse mediante factores externos que habrían sido olvidados por parte del equipo de evaluación?

- ¿Contradice el hallazgo las opiniones de los expertos o las conclusiones obtenidas por otro medio? En caso afirmativo, ¿puede explicarse este hecho?
- ¿Disponen los miembros del grupo de referencia de la evaluación de argumentos que contradigan el hallazgo? En caso afirmativo, ¿están justificados dichos argumentos?

Recomendaciones

Dedicar interacciones relativamente largas al debate del informe final para dar lugar a un análisis confirmatorio meticuloso. Asegurarse de que el equipo de evaluación haya reservado los suficientes recursos a tal efecto.

No todos los hallazgos requieren el mismo nivel de confirmación. Concentrar esfuerzos en hallazgos que apoyen las conclusiones más controvertidas, las lecciones que sean más probables de ser transferidas, o las recomendaciones que sean más difíciles de aceptar.

A fin de mejorar la credibilidad de la evaluación, es valioso presentar las críticas soportadas por los hallazgos durante el análisis confirmatorio.

10.3 Validez del análisis

La validez es una calidad esencial del análisis. Se obtiene cuando:

- Los hallazgos tienen un alcance general que va más allá de la muestra observada o de los casos estudiados (validez externa)
- Los hallazgos derivan de datos sin ningún sesgo (validez interna)
- Los datos recogidos reflejan exactamente lo que se debía describir (validez de construcción).

10.3.1 Validez externa

Corresponde a la calidad de un método de evaluación que posibilita la obtención de hallazgos que pueden generalizarse hacia otros grupos, áreas, períodos, etc. La validez externa se logra de manera íntegra cuando el equipo de evaluación es capaz de dejar en claro que una intervención similar implementada en otro

contexto tendría los mismos efectos bajo determinadas condiciones.

10.3.2 Validez interna

Corresponde a la calidad de un método de evaluación que, en la medida de lo posible, limita los sesgos imputables a la recolección de datos o a las técnicas de análisis. La validez interna se logra cuando el equipo de evaluación aporta argumentos irrefutables mostrando que los hallazgos derivan de hechos.

10.3.3 Validez de construcción

La validez de construcción se logra cuando los conceptos clave son transparentes y cuando los indicadores miden lo que deben medir.

11 Juicio

RESUMEN

Conclusiones y lecciones

Las conclusiones aportan respuestas claras a las preguntas planteadas al inicio de la evaluación. Integran los juicios sobre el valor y los méritos de la intervención.

Las lecciones son conclusiones transferibles a los siguientes ciclos de la misma intervención, a otras intervenciones o a intervenciones similares en otros contextos.

Recomendaciones

Las recomendaciones no implican ningún juicio propiamente dicho, pero sí derivan de las conclusiones. Tienen como finalidad mejorar o reformar la intervención evaluada, o preparar el diseño de una nueva intervención para el próximo ciclo.

11.1 Conclusiones y lecciones

Las conclusiones comprenden un juicio

En la medida de lo posible, el informe de evaluación enuncia por separado los hallazgos (que se desprenden únicamente de los hechos y de los análisis) y luego las conclusiones (que implican un juicio de valor).

Para formular las conclusiones, el equipo de evaluación aplica los criterios de juicio (también llamados "criterios de evaluación razonada") acordados durante la primera fase de la evaluación (*desk*) (ver Sección 7.1). Siempre que sea posible, se comparan los hallazgos con los niveles meta (*targets*).

Estos pasos refuerzan la calidad en la medida en la que obliga al equipo de evaluación a explicitar la manera en la que formula su juicio. El hecho de reconocer que las conclusiones comprenden un juicio de valor conduce además a respetar mejor los principios éticos presentados en el cuadro que sigue a continuación.

Ética del juicio

- **Responsabilidad**

En la fase de síntesis, el equipo de evaluación aplica los criterios de juicio convenidos de la manera más fiel posible, y desarrolla sus propias conclusiones. Se debaten las conclusiones dentro del grupo de referencia aunque siguen estando bajo la plena responsabilidad del equipo de evaluación.

- **Legitimidad**

Las preguntas y criterios tienen en cuenta las necesidades y el punto de vista de la institución pública que ha iniciado la evaluación.

Los miembros del grupo de referencia pueden aportar puntos de vista diferentes, lo que refuerza la legitimidad de la evaluación.

En el transcurso de la fase de documentación, el equipo de evaluación conduce entrevistas que eventualmente permiten identificar otros puntos de vista legítimos pero no expresados en el grupo de referencia. Rinde cuentas de ello en las reuniones del grupo de referencia y, si necesario, lo tiene en cuenta en el criterio de juicio.

- **Imparcialidad**

Los actores proceden de manera tal que todas las opiniones sean escuchadas, incluso las que no provienen de las voces más influyentes o más numerosas. Están conscientes de las relaciones de poder asimétricas y corrigen los sesgos que resulten de las mismas.

Los miembros del equipo de evaluación están familiarizados con las creencias, usos y costumbres de los grupos aludidos y las respetan.

Se rinde cuenta de las divergencias en los informes y en las reuniones del grupo de referencia.

- **Protección de las personas**

Las conclusiones están basadas en los méritos de la intervención evaluada, no en las personas que la aplican o que se benefician de ella.

En la etapa del informe final provisorio, el equipo de evaluación puede tener que afinar sus criterios de juicio y niveles meta. En este caso, se debate la cuestión junto con el grupo de referencia.

Respuestas a las preguntas y conclusiones generales

Las conclusiones aportan respuestas claras a las preguntas de evaluación.

Un capítulo del informe sintetiza estas respuestas bajo la forma de conclusiones a un nivel más general, a fin de dar una visión global de la intervención evaluada. No sigue el orden de las preguntas ni de los criterios de evaluación (eficacia, eficiencia, coherencia, etc.).

Recapitulación de las conclusiones

El informe contiene una lista de conclusiones que hacen referencia a la manera en la que derivan de los hechos, análisis y criterios de juicio. El informe contiene además una auto-evaluación de los límites metodológicos que pueden cercenar el alcance o la utilización de determinadas conclusiones.

El informe identifica las lecciones, que son conclusiones transferibles a los próximos ciclos de la misma intervención, a otras intervenciones o a intervenciones similares en otros contextos.

Un párrafo o un subcapítulo, evitando siempre caer en la repetición, retoma las 3 o 4 conclusiones principales organizadas por orden de importancia. Este ejercicio permite destacar mejor los mensajes de la evaluación dirigidos a los responsables políticos dentro de la Comisión.

i Atención!

El equipo de evaluación y las personas a cargo de la garantía de calidad deben revisar atentamente el informe final para eliminar todos los juicios de valor no indispensables o involuntarios.

El equipo de evaluación expone sus conclusiones de manera equilibrada, sin privilegiar sistemáticamente las conclusiones negativas ni positivas.

11.2 Recomendaciones

11.2.1 ¿Cómo elaborarlas y presentarlas?

Las recomendaciones deben estar vinculadas a las conclusiones sin llegar al extremo de ser solo una repetición. Una recomendación deriva directamente de una conclusión o de varias.

i Atención!

Si una recomendación no deriva de las conclusiones con claridad, es probable que refleje ideas preconcebidas o los intereses de una de las partes interesadas. Su presencia en el informe corre el riesgo de desacreditar la evaluación.

Se debe agrupar y priorizar las recomendaciones. El informe menciona a quién están destinadas, por ejemplo: la Delegación de la C.E., los servicios a cargo del diseño de la próxima intervención, etc.

Las recomendaciones son útiles, operativas y factibles. Se especifican las condiciones de puesta en práctica. Siempre que se pueda y sea pertinente, las principales recomendaciones se presentan en forma de opciones, con condiciones vinculadas a cada opción, así como las consecuencias previsibles de la implementación de cada una de ellas. Se las presentan en un capítulo específico, el cual resalta las recomendaciones derivadas de las tres o cuatro conclusiones principales.

11.2.2 ¿Cómo promoverlas?

Las recomendaciones tienen valor solamente si sus destinatarios las entienden, y en lo posible si se apropian de ellas. Para promover esta apropiación, el responsable prepara una *fiche contradictoire* que permite:

- Enumerar las recomendaciones en forma resumida
- Recabar las respuestas de los servicios destinatarios
- Informar acerca del seguimiento que en realidad se le da a las recomendaciones.

12 Garantía de calidad

RESUMEN

Se trata de establecer puntos de control de calidad en las principales etapas del proceso, de definir las responsabilidades de garantía de calidad y de establecer cuáles serán las reglas del juego en caso de fallas en la calidad.

La garantía de calidad refuerza las conclusiones de la evaluación ante los usuarios mostrando que se trata de un trabajo imparcial y riguroso que respeta estándares profesionales.

La calidad se construye progresivamente para evitar encontrarse con un defecto de calidad en la etapa del informe final.

La reglas de calidad aclaran las relaciones entre el responsable de la evaluación, el grupo de referencia y el equipo de evaluación.

12.1 Reglas del juego

Se especifican las reglas de garantía de calidad en los términos de referencia y se las recuerda en el momento de contratar al equipo de evaluación.

En cada paso se aseguran aspectos sucesivos de calidad a fin de avanzar sobre bases sólidas hacia la siguiente etapa.

Estas reglas están relacionadas con:

- La aprobación de los documentos
- Los criterios de calidad
- La difusión de la evaluación de calidad del informe final.

El proceso de garantía de calidad se beneficia del aporte de todos los actores mientras limita los potenciales conflictos que pueden surgir entre ellos, por ejemplo:

- Los actores que tienen responsabilidades en la intervención evaluada (los evaluados) están en una buena posición para evaluar la pertinencia de los datos recabados. Su

apreciación, sin embargo, puede ser sensible al sesgo de confirmación. Por lo tanto deben desempeñar un papel consultivo, por ejemplo a través de su participación en el grupo de referencia.

- El líder del equipo de evaluación externo cumple una función preponderante en la garantía de la calidad, en particular diseñando un método adecuado y asegurándose de la precisión y la imparcialidad del informe. Sin embargo corre el riesgo de sobrestimar los puntos de vista de las partes interesadas influyentes, en especial los de aquéllas susceptibles de solicitarle otras evaluaciones en el futuro. La evaluación de calidad del responsable de la evaluación contribuye a evitar dicho riesgo.

Criterios de calidad

La calidad del informe de evaluación recibe una calificación basada en nueve criterios con la ayuda de una plantilla que consta de cinco niveles: Excelente, Muy Bueno, Bueno, Pobre e Inaceptable.

Criterio 1: Satisfacción de las necesidades

El informe, ¿trata de manera adecuada los pedidos de información formulados por los comanditarios? y ¿guarda relación con los términos de referencia?

Criterio 2: Pertinencia del campo

¿Se hace un examen completo de la razón de ser de la intervención, de sus hallazgos, resultados e impactos, así como de sus interacciones previstas e imprevistas con otras políticas y sus consecuencias?

Criterio 3: Justificación del método

El diseño de la evaluación, ¿está adaptado y es el adecuado para garantizar que todos los hallazgos, junto con las limitaciones metodológicas, sean accesibles a fin de responder las preguntas de evaluación principales?

Criterio 4: Fiabilidad de los datos

Los datos primarios y secundarios seleccionados, ¿están adaptados?, ¿son lo suficientemente fiables para el uso que se espera de ellos?

Criterio 5: Solidez del análisis

¿Se analiza adecuada y sistemáticamente la información cuantitativa y cualitativa según el estado del arte a fin de que las preguntas de evaluación tengan una respuesta válida?

Criterio 6: Credibilidad de los hallazgos

Los resultados ¿se desprenden de manera razonable y justificada del análisis de los datos y de las interpretaciones basadas en supuestos y en una lógica presentada con esmero?

Criterio 7: Validez de las conclusiones

El informe, ¿brinda conclusiones claras? ¿están basadas en hallazgos creíbles?

Criterio 8: Utilidad de las recomendaciones

Las recomendaciones, ¿son equitativas y sin sesgos que favorezcan los puntos de vista personales o de las partes interesadas?, además, ¿son lo suficientemente detalladas como para ser aplicables concretamente?

Criterio 9: Claridad del informe

El informe, ¿describe claramente la intervención evaluada, inclusive su contexto y finalidad, así como el proceso y los hallazgos de la evaluación de manera tal que la información proporcionada sea comprensible con facilidad?

Los criterios 1 y 2 son verificados en la etapa del informe de inicio.

En la etapa del informe de la fase de documentación, conviene verificar los criterios 1 - 3, y el criterio 4 en cuanto a los datos ya recolectados.

La apreciación de la totalidad de los criterios 1 a 9 se lleva a cabo en la etapa del informe final.

En el sitio web: [consejos para evaluar la calidad según cada uno de los 9 criterios](#)

12.2 Aprobación de los informes

La aprobación de los informes es un reconocimiento formal por parte del responsable de la evaluación del hecho de que un documento entregado posee un contenido adecuado y satisface los criterios de calidad aplicables a la metodología.

12.2.1 ¿Cuáles son las etapas de la aprobación?

Salvo excepción, se procede a la aprobación de un documento de la siguiente manera:

- El responsable verifica que el documento tenga el formato y contenido requeridos, y que no haya defectos significativos de calidad. De no ser así, solicita una nueva versión de inmediato.
- El responsable recoge la opinión de las personas o servicios a ser consultados, ya sea durante el transcurso de una reunión (grupo de referencia), o bien por email, indicando un plazo limitado más allá del cual no se garantiza la toma en consideración de los comentarios.
- El responsable enumera sus propias observaciones de mejora de calidad, adjunta las otras observaciones recibidas, hace una distinción entre observaciones de orden metodológico y observaciones de fondo y las transmite al equipo de evaluación.
- El equipo de evaluación considera todas las observaciones de mejora de calidad, ya sea modificando su documento de manera consecuente, o mencionando en anexo cómo será tratada la observación hecha en una etapa posterior, o bien explicando en anexo por qué no será tratada.
- Sin dilación, se presenta una nueva versión del documento ante el responsable, quien la acepta o la rechaza.
- Desde que se la solicita, la aprobación condiciona el pasaje a la etapa siguiente.

12.2.2 ¿Qué documentos se deben aprobar y quién está a cargo de su aprobación?

Los documentos sometidos a aprobación son los siguientes:

- Informe inicial (*inception*)
- Informe de primera fase (*desk*)

- Informe final provisorio
- Informe final.

Por lo general el responsable de la evaluación efectúa la aprobación y eventualmente su superior jerárquico la confirma.

Glosario

El vocabulario y la terminología desarrollados en las cuatro guías corresponden a las definiciones tanto de la OCDE como de la Comisión Europea.

Definiciones de la OCDE

En 2002, la OCDE/CAD publicó un glosario de términos principales en materia de evaluación y de gestión basada en los resultados (Comité de Ayuda al Desarrollo - Grupo de trabajo sobre la evaluación de la ayuda). Se lo redactó inicialmente en inglés, francés y castellano para luego ser traducido a otros idiomas. Forma parte del esfuerzo conjunto que realiza la comunidad de ayuda al desarrollo, incluida la Comisión Europea, con miras a reducir la confusión terminológica y promover un vocabulario en común.

Definiciones de la Comisión Europea

El Reglamento Financiero establece un marco aplicable a todas las políticas de la Comisión. Dicho marco ha sido plasmado en guías de evaluación así como en un glosario redactado por la Dirección General de Presupuesto

Comparación de las definiciones

Las definiciones de la OCDE y de la Comisión Europea son igualmente legítimas en el contexto de la evaluación de la ayuda europea. Se llevó a cabo una comparación en profundidad obteniendo como resultado que ***casi todos los términos están definidos de manera coherente.***

Sólo 4 pares de definiciones son lo suficientemente diferentes como para generar riesgos de malos entendidos:

<i>Riesgo de malos entendidos (cuatro términos)</i>		
	OCDE	CE
Impacto	... efectos a largo plazo producidos por una intervención para el desarrollo ...	Un término general utilizado para describir los efectos de una intervención en la sociedad...
Outcome <i>(traducido como efecto directo)</i>	Los efectos, a corto o a mediano plazo probables o logrados por los productos de una intervención.	Los impactos a más largo plazo, por lo general expresados como consecuencias socio económicas en el sentido amplio de la palabra.
Resultado	... producto, efecto directo o impacto (intencional o no, positivo y/o negativo)	El impacto inicial de una intervención...
Eficiencia	Medida en que los recursos (fondos, tiempo, etc.) se han convertido económicamente en resultados (según la definición de los resultados de la OCDE, la eficiencia puede estar relacionada con los productos así como con cualquier nivel de efecto).	Medida en la que se logran los efectos producidos a un costo razonable (una definición que no cubre los productos).

A los fines de la redacción de esta guía, se eligieron las siguientes opciones:

<i>Opciones escogidas</i>		
	<i>Glosario de base</i>	<i>Definición</i>
Impacto	OCDE	...efectos a largo plazo producidos por una intervención de desarrollo...
Outcome <i>(efecto directo)</i>	Este término no ha sido ni utilizado ni definido	
Resultado	CE	El impacto inicial de una intervención...
Eficiencia	Mixta (basada en la OCDE)	Medida en la que se alcanzan los efectos deseados (desde los productos hasta los impactos) utilizando la menor cantidad posible de recursos (fondos, habilidades técnicas, tiempo, costos administrativos, etc.)

En otros documentos publicados por la CE, se puede emplear el término resultado en un sentido amplio, tal como lo define la OCDE/CAD, por ejemplo: Seguimiento Orientado hacia los Resultados.

Disponible en el sitio web Europa

<http://ec.europa.eu/europeaid>

Numerosos ejemplos y *listas de verificación*, en especial:

- Ejemplos de preguntas para una evaluación a nivel país, a nivel regional y para una evaluación temática o sectorial.
- Ejemplos de articulación entre preguntas, criterios de juicio e indicadores.
- Ejemplo de tabla de diseño y de desarrollo de una herramienta.
- Plantilla de evaluación de la calidad e indicaciones para completarla.
- Ejemplos de resúmenes y artículos que presentan una evaluación.

Lista de cuadros

Cuadro 1 – Intervenciones simples / complejas.....	15
Cuadro 2 - Distinción de sectores y aspectos transversales	17
Cuadro 3 – Evaluación que concierne a varios ciclos	21
Cuadro 4 – Calendario y papel del grupo de referencia.....	32
Cuadro 5 – Diagrama de los efectos esperados	40
Cuadro 6 – Enfocar la evaluación	43
Cuadro 7 – Criterios de evaluación.....	49
Cuadro 8– Indicadores y lógica de intervención.....	61
Cuadro 9 – De la pregunta al indicador: ejemplo.....	62
Cuadro 10 – Estructura de una tabla de diseño	66
Cuadro 11 – Escenario contrafactual	81

