

Aplicación minuciosa de los criterios de evaluación



Aplicación minuciosa de los criterios de evaluación

El presente trabajo se publica bajo la responsabilidad del Secretario General de la OCDE. Las opiniones expresadas y los argumentos utilizados en el mismo no reflejan necesariamente el punto de vista oficial de los Países miembros de la OCDE.

Tanto este documento, así como cualquier dato y cualquier mapa que se incluya en él, se entenderán sin perjuicio respecto al estatus o la soberanía de cualquier territorio, a la delimitación de fronteras y límites internacionales, ni al nombre de cualquier territorio, ciudad o área.

Los datos estadísticos para Israel son suministrados por y bajo la responsabilidad de las autoridades israelíes competentes. El uso de estos datos por la OCDE es sin perjuicio del estatuto de los Altos del Golán, Jerusalén Este y los asentamientos israelíes en Cisjordania bajo los términos del derecho internacional.

Nota de Turquía

La información del presente documento en relación con “Chipre” se refiere a la parte sur de la Isla. No existe una sola autoridad que represente en conjunto a las comunidades turcochipriota y grecochipriota de la Isla. Turquía reconoce a la República Turca del Norte de Chipre (RTNC). Mientras no haya una solución duradera y equitativa en el marco de las Naciones Unidas, Turquía mantendrá su postura frente al “tema de Chipre”.

Nota de todos los Estados Miembros de la Unión Europea que pertenecen a la OCDE y de la Unión Europea
Todos los miembros de las Naciones Unidas, con excepción de Turquía, reconocen a la República de Chipre. La información contenida en el presente documento se refiere a la zona sobre la cual el Gobierno de la República de Chipre tiene control efectivo.

Por favor, cite esta publicación de la siguiente manera:

OECD (2025), *Aplicación minuciosa de los criterios de evaluación*, OECD Publishing, Paris, <https://doi.org/10.1787/d1ee36bb-es>.

ISBN 978-92-64-43340-3 (PDF)

Imágenes: Portada © TFK.

Las erratas de las publicaciones se encuentran en línea en: <https://www.oecd.org/en/publications/support/corrigenda.html>.

© OCDE 2025



Atribución/Reconocimiento 4.0 Internacional (CC BY 4.0)

Este trabajo está disponible bajo la licencia Creative Commons Attribution/Reconocimiento 4.0 Internacional. Al utilizar este trabajo, acepta estar sujeto a los términos de esta licencia (<https://creativecommons.org/licenses/by/4.0/>).

Atribución – debe citar el trabajo.

Traducciones – debe citar la obra original, identificar cambios al original y agregar el siguiente texto: *En caso de discrepancia entre la obra original y la traducción, solo se considerará válido el texto de la obra original.*

Adaptaciones – debe citar el trabajo original y agregar el siguiente texto: *Esta es una adaptación de un trabajo original de la OCDE. Las opiniones expresadas y los argumentos utilizados en esta adaptación no deben considerarse representativos de los puntos de vista oficiales de la OCDE o de sus Países miembros.*

Material de terceros – la licencia no se aplica al material de terceros en la obra. Si utiliza dicho material, usted es responsable de obtener el permiso del tercero y de cualquier reclamación por infracción.

No debe utilizar el logotipo, la identidad visual o la imagen de portada de la OCDE sin permiso expreso ni sugerir que la OCDE respalda su uso del trabajo.

Cualquier disputa que surja bajo esta licencia se resolverá mediante arbitraje de conformidad con el Reglamento de Arbitraje de la Corte Permanente de Arbitraje (PCA) de 2012. El lugar del arbitraje será París (Francia). El número de árbitros será de uno.

Prólogo

La evaluación desempeña un papel fundamental en dar forma al diseño y la ejecución de políticas y programas que conducen a mejores efectos directos sobre el desarrollo, es decir, más justos y sostenibles. La evidencia de la evaluación y la evaluación del pensamiento crítico pueden contribuir y jugar un papel crucial a la hora de ayudar a los responsables de la toma de decisiones y a las comunidades a garantizar que las políticas y los programas brinden resultados positivos y duraderos para las personas y el planeta.

Para apoyar la práctica de la evaluación, en 1991 el Comité de Ayuda al Desarrollo (CAD) de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) publicó las definiciones de cinco criterios de evaluación. Estos criterios han dado forma al diseño y a la evolución de las evaluaciones internacionales del desarrollo en los últimos 30 años. Es muy conocido el papel central que desempeñan los criterios en la mejora de la calidad de las prácticas mundiales de evaluación y en el apoyo a la colaboración. Esos criterios han permitido a las organizaciones diseñar y realizar evaluaciones pertinentes para las necesidades de los responsables de la toma de decisiones, captar una amplia variedad de resultados intencionales y no intencionales, y producir evidencias y conocimientos valiosos.

Estos cinco criterios se adaptaron aún más entre 2018 y 2019. El proceso de revisión se basó en cerca de tres decenios de aprendizaje por parte de los miembros de la Red de Evaluación del Desarrollo (EvalNet) del CAD, junto con la comunidad evaluadora mundial en general. Se basó en la Agenda 2030 para el Desarrollo Sostenible, incluidos los Objetivos de Desarrollo Sostenible (ODS), y el Acuerdo de París dentro de la Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC). En diciembre de 2019, se publicó un nuevo conjunto de definiciones, el cual incluye dos principios para orientar su uso.

Esta guía complementa las definiciones de los criterios y sus principios de uso. Su objetivo es apoyar la aplicación minuciosa y contextualizada de los criterios. A través de la inclusión de preguntas y ejemplos, la guía fomenta la reflexión crítica y un análisis que aprecie los matices. La guía ayudará a los lectores a considerar cómo interpretar y aplicar los criterios para mejorar la calidad de sus evaluaciones y apoyar mejor el aprendizaje y la rendición de cuentas.

El aprendizaje sobre la mejor manera de utilizar estos criterios adaptados, tanto en la cooperación internacional para el desarrollo como en otras esferas, solo acaba de comenzar. La intención es continuar recopilando lecciones y perspectivas de su uso, y volver a revisar esta guía sobre la base de la experiencia colectiva y la retroalimentación de la comunidad evaluadora de todo el mundo. En particular, serán valiosas las lecciones derivadas de las nuevas formas de utilizar los criterios que reflejen mejor los principios de la Agenda 2030, como las evaluaciones participativas o que aplican un enfoque basado en los derechos humanos, las evaluaciones de procesos de cambios complejos y las evaluaciones que engloban sinergias y compensaciones de manera integral.

Reconocimientos

La Secretaría desea dar las gracias a la comunidad evaluadora mundial por las amplias conversaciones y debates sobre los criterios que tuvieron lugar entre 2017 y 2019, los cuales contribuyeron a identificar la necesidad de esta guía y ayudaron a dar forma a su contenido. En particular, los participantes de la Conferencia Internacional sobre Capacidades Nacionales de Evaluación (CNE), el Grupo de Evaluación de las Naciones Unidas (UNEG) y el Grupo de Cooperación en Materia de Evaluación (ECG) fueron asociados clave en la adaptación de los criterios, y en este documento se reflejan las aportaciones de muchos de sus miembros. La Secretaría agradece las contribuciones de los asociados de evaluación de todo el mundo y acoge con beneplácito el debate para seguir desarrollando este proyecto de guía de trabajo.

El presente documento refleja las amplias contribuciones de los miembros de la Red de Evaluación del Desarrollo (EvalNet), órgano subsidiario del Comité de Ayuda al Desarrollo (CAD) de la Organización para la Cooperación y el Desarrollo Económicos (OCDE). Los miembros de EvalNet contribuyeron con muchos comentarios y reflexiones útiles sobre el proyecto de definiciones de criterios y participaron en debates en profundidad sobre los conceptos contenidos en este documento. En muchos casos, esto implicó amplias conversaciones dentro de los organismos miembros y ministerios, así como consultas con las partes interesadas en la evaluación del desarrollo. Los miembros de EvalNet proporcionaron información sobre sus propias experiencias y desafíos con los criterios, incluyendo ejemplos concretos, lo que ayudó a la elaboración del documento. El Grupo de Trabajo sobre Igualdad de Género de EvalNet contribuyó a desarrollar elementos pertinentes de esta guía. La Secretaría también reconoce con gratitud el papel esencial del Presidente de EvalNet, Per Øyvind Bastøe, a la hora de orientar el proceso de adaptación y el desarrollo de esta guía.

Megan Grace Kennedy-Chouane dirigió el desarrollo de esta guía para la OCDE, con el apoyo y las aportaciones de Rahul Malhotra, Hans Lundgren, Alison Pollard, Ola Kasneci, Yee King Ho y Nelson Torbay-Holguin. Muchos otros colegas de la Dirección de Cooperación para el Desarrollo proporcionaron retroalimentación sobre proyectos anteriores. Andrew Brenner y Suzanne Parandian brindaron retroalimentación editorial. Stephanie Coic y TFK elaboraron el diseño gráfico de las imágenes y los gráficos.

Ima Bishop, Joe Watkins y Nick York (IOD PARC) redactaron la guía y realizaron contribuciones inestimables a su contenido. Julia Betts realizó la revisión de pares y proporcionó retroalimentación, además de contribuir de forma amplia al proceso inicial de revisión de los criterios, las nuevas definiciones y su conceptualización.

El Banco Interamericano de Desarrollo (BID) y el Instituto Alemán de Evaluación del Desarrollo (DEval) lideraron la traducción al español de esta guía, con el apoyo de la Secretaría de la OCDE, que expresa su agradecimiento a Cinthia Josette Arévalo Gross (BID), Ivory Yong Protzel (BID), María Cristina González Arias (BID), José Claudio Linhares Pires (BID), Nataly Salas (DEval), Mariana Guedes Gamarra (OCDE), Nelson Torbay-Holguín (OCDE) y Megan Kennedy-Chouane (OCDE) por sus valiosas contribuciones a la traducción al español y a la revisión de los capítulos de la guía.

Índice

Prólogo	3
Reconocimientos	4
Abreviaturas y siglas	8
Resumen ejecutivo	10
1 Propósito y uso de la guía: mejores criterios para una mejor evaluación	12
Razones por las que se necesita una guía	13
Propósito de la guía y cómo utilizarla	13
Personas que pueden utilizar esta guía	15
Referencias	16
Notas	16
2 Los seis criterios: propósito y función dentro de la evaluación	17
¿Qué es un criterio?	18
¿Qué tipos de intervenciones podemos evaluar con estos criterios?	19
¿En qué medida los criterios son compatibles con el ámbito de evaluación más amplio?	21
Referencias	22
Notas	24
3 Uso de los criterios de evaluación en la práctica	25
Cómo aplicar los criterios con minuciosidad	26
Cómo adaptar los criterios al propósito de la evaluación	27
Cómo entender la dimensión de temporalidad	30
Cómo utilizar los criterios para mejorar las preguntas de evaluación	31
Relaciones entre los criterios	31
Cómo elegir qué criterios utilizar	32
Cómo operativizar y adaptar los criterios a nivel institucional	33
Cómo responder a la Agenda 2030 y a los Objetivos de Desarrollo Sostenible	34
Cómo aplicar un lente de género a los criterios	36
Cómo usar otros criterios	38
Referencias	38
Notas	40

4 Cómo entender los seis criterios: definiciones, elementos de análisis y principales desafíos	41
Pertinencia	43
Coherencia	51
Eficacia	59
Eficiencia	66
Impacto	74
Sostenibilidad	82
Referencias	89
Notas	93

Cuadros

Cuadro 3.1. Cómo aplicar un lente de género a los criterios	37
Cuadro 4.1. Desafíos de evaluar la pertinencia	47
Cuadro 4.2. Desafíos de evaluar la coherencia	55
Cuadro 4.3. Desafíos de evaluar la eficacia	63
Cuadro 4.4. Desafíos de evaluar la eficiencia	69
Cuadro 4.5. Desafíos de evaluar el impacto	78
Cuadro 4.6. Desafíos de la evaluación de sostenibilidad	86

Gráficos

Gráfico 2.1. En qué medida son compatibles los criterios con otras normas y estándares	22
Gráfico 3.1. Cómo cubrir los criterios para una evaluación específica	33
Gráfico 4.1. Los seis criterios y sus preguntas relacionadas	42

Recuadros

Recuadro 1.1. Otras investigaciones y próximas publicaciones de la OCDE sobre evaluación	14
Recuadro 2.1. Cómo adaptar los criterios a la intervención: evaluación del plan de acción para la inclusión de las personas con discapacidad del BMZ	20
Recuadro 3.1. Uso minucioso de los criterios: evaluación de la donación de la sociedad civil noruega para países en desarrollo	27
Recuadro 3.2. Definición del propósito de la evaluación: ejemplos de Uganda, Kenia y Suecia	29
Recuadro 3.3. Uso de los ODS para informar las agendas nacionales de evaluación	35
Recuadro 3.4. Mapeo de los criterios en función de la Agenda 2030 y los ODS	35
Recuadro 4.1. Evaluación de pertinencia de una iniciativa de apoyo al presupuesto general	49
Recuadro 4.2. Evaluación del proyecto “Mejora de medios de subsistencia de los agricultores de la Zona Seca de Myanmar”	50
Recuadro 4.3. Evaluación de la pertinencia de iniciativas de salud en Bolivia	51
Recuadro 4.4. Evaluación de los esfuerzos noruegos para garantizar la coherencia de las políticas de desarrollo	57
Recuadro 4.5. Evaluación de la coherencia de la asistencia para la reconstrucción tras un desastre natural en las Filipinas	58
Recuadro 4.6. Evaluación de la coherencia de la cooperación para el desarrollo de Eslovenia con Montenegro	59
Recuadro 4.7. Evaluación de la eficacia y la inclusión de la asistencia electoral australiana	64
Recuadro 4.8. Evaluación de la eficacia de la estrategia y el programa de país del Banco Africano de Desarrollo en Cabo Verde	65
Recuadro 4.9. Evaluación de la eficiencia del apoyo del Banco Mundial para electrificación rural	71
Recuadro 4.10. Evaluación de la eficiencia de los subsidios para insumos agrícolas en el África subsahariana	72
Recuadro 4.11. Evaluación de la eficiencia a nivel de políticas: la política de asistencia al desarrollo neerlandesa para mejorar la gestión del agua	73

Recuadro 4.12. Evaluación de la eficacia y el impacto del programa Más Familias en Acción de Colombia	79
Recuadro 4.13. Evaluación del impacto del apoyo de Asdi a la construcción de la paz en países afectados por conflictos y en situaciones posconflicto	79
Recuadro 4.14. Valoración del impacto en la evaluación de una política estratégica	80
Recuadro 4.15. Evaluación de impacto, difusión y aumento del impacto de un enfoque integral de planificación del uso de la tierra en las Filipinas	80
Recuadro 4.16. Evaluación del impacto del programa Empoderamiento de la Niña (GE, por sus siglas en inglés) en Liberia	81
Recuadro 4.17. Examen de la sostenibilidad en una evaluación <i>ex post</i> del Proyecto de Desarrollo Rural Intercomunitario del Japón en Afganistán	87
Recuadro 4.18. Lecciones de sostenibilidad de la evaluación del cese del apoyo al presupuesto general en Malaui, Ruanda, Uganda y Zambia	88
Recuadro 4.19. Evaluación del impacto y la sostenibilidad del programa de salud materna del DFID	89

Abreviaturas y siglas

3ie	Iniciativa Internacional para la Evaluación de Impacto
ACNUR	Oficina del Alto Comisionado de las Naciones Unidas para los Refugiados
ALNAP	Red para un Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria
AOD	Asistencia oficial para el desarrollo
Asdi	Agencia Sueca de Cooperación Internacional para el Desarrollo
BAfD	Banco Africano de Desarrollo
BMZ	Ministerio Federal de Cooperación Económica y Desarrollo de Alemania
CAD	Comité de Ayuda al Desarrollo
CEDAW	Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer
CMNUCC	Convención Marco de las Naciones Unidas sobre el Cambio Climático
CRPD	Convención sobre los Derechos de las Personas con Discapacidad
DANIDA	Agencia Danesa de Desarrollo Internacional
DEval	Instituto Alemán de Evaluación del Desarrollo
DFID	Departamento de Desarrollo Internacional del Reino Unido
ECG	Grupo de Cooperación en Materia de Evaluación
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura

GEI	Grupo de Evaluación Independiente, Banco Mundial
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
ICAI	Comisión Independiente para el Impacto de la Ayuda del Reino Unido
IFD	Institución de Financiación del Desarrollo
JICA	Agencia de Cooperación Internacional del Japón
OCDE	Organización para la Cooperación y el Desarrollo Económicos
ODS	Objetivos de Desarrollo Sostenible
ONG	Organización no gubernamental
ONU	Organización de las Naciones Unidas
PMA	Programa Mundial de Alimentos
SMART	Específicos, mensurables, alcanzables, pertinentes y oportunos
UE	Unión Europea
UNEG	Grupo de Evaluación de las Naciones Unidas

Resumen ejecutivo

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) ha establecido definiciones comunes para seis criterios de evaluación (pertinencia, coherencia, eficacia, eficiencia, impacto y sostenibilidad) para respaldar una evaluación consistente y de alta calidad. Estos criterios proporcionan el marco normativo utilizado para determinar el mérito o el valor de una intervención (políticas, estrategias, programas, proyectos o actividades). Sirven como base sobre la cual se formulan los juicios evaluativos.

Los criterios pueden considerarse como un conjunto de lentes que proporcionan perspectivas que se complementan y que, en conjunto, brindan una imagen integral de una intervención y sus resultados. Los criterios animan a los evaluadores, así como a las personas que participan en el diseño o la gestión de las intervenciones, a pensar con mayor profundidad en la naturaleza de una intervención, su proceso de ejecución y sus resultados.

La Red de Evaluación del Desarrollo del CAD de la OCDE estableció dos principios rectores junto con las definiciones de los seis criterios:

1. Primer principio: Los criterios deben aplicarse con minuciosidad para conseguir evaluaciones útiles y de calidad.
2. Segundo principio: El uso de los criterios depende del propósito de la evaluación.

Los criterios deben aplicarse con minuciosidad y adaptarse al contexto de la intervención y la evaluación. Aunque, en principio, se desarrollaron para su uso en la cooperación internacional para el desarrollo, los criterios pueden aplicarse en cualquier sector y para las evaluaciones de las intervenciones públicas o privadas. Se pueden utilizar para evaluar muchos temas y tipos diferentes de intervenciones, incluidos asuntos temáticos o cuestiones estratégicas, políticas y proyectos. Se pueden utilizar diferentes metodologías al evaluar con base en los criterios, habida cuenta de que los criterios forman solo una parte de todo el espectro de las normas y los estándares de evaluación.

Los evaluadores deben trabajar de forma que consideren minuciosamente las experiencias diferenciadas y reflexionen sobre cómo la dinámica de poder basada en el género y otras formas de discriminación (por ejemplo, edad, raza o etnia, estatus social o capacidad) afectan a la ejecución y los resultados de la intervención.

Los conceptos clave para cada uno de los criterios se definen de la siguiente manera:

- La pertinencia implica examinar hasta qué punto los objetivos y el diseño de la intervención responden a las necesidades y prioridades de los beneficiarios, así como la alineación con las políticas y prioridades nacionales, mundiales y de los asociados y las instituciones. Entender la dinámica de poder con perspectiva de género y reflexionar sobre el compromiso de “no dejar a nadie atrás” es fundamental para comprender la pertinencia. Si las circunstancias cambian, las evaluaciones también deben considerar si las intervenciones siguen siendo pertinentes.
- La coherencia (los nuevos criterios) examina el grado en el que otras intervenciones (sobre todo las políticas) apoyan o socavan la intervención y viceversa. Esto incluye la coherencia interna (dentro de una institución o un gobierno), como la compatibilidad con las normas y los estándares

internacionales, y la coherencia externa (con las intervenciones de otros actores en el mismo contexto). La coherencia incluye conceptos de complementariedad, armonización y coordinación, así como el grado en que la intervención ofrece un valor añadido y, a su vez, evita la duplicación de esfuerzos. De conformidad con la Agenda 2030, se debe prestar mayor atención a la coherencia, con un mayor énfasis en las sinergias (o compensaciones) entre las áreas de políticas. Este nuevo criterio fomenta un enfoque integrado para comprender las intervenciones complejas y sus resultados.

- La eficacia examina la medida en que la intervención ha logrado, o se espera que logre, sus objetivos y sus resultados, teniendo en cuenta al mismo tiempo la importancia relativa de los objetivos. La nueva definición fomenta el análisis de los resultados diferenciados entre los grupos y la medida en que la intervención contribuye o exacerba las brechas de equidad. La eficacia es el criterio evaluado con más frecuencia y, a menudo, se utiliza como una medida general del éxito.
- La eficiencia ayuda a los evaluadores a formular preguntas sobre el grado en el que la intervención produce, o es probable que produzca, resultados de manera económica y oportuna. Por “de manera económica” se entiende la transformación de insumos (fondos, conocimientos especializados, recursos naturales, tiempo, etc.) en resultados de la manera más eficaz en función de los costos, en comparación con otras alternativas viables según el contexto. La nueva definición incluye la dimensión de “entrega oportuna”. Este criterio es una oportunidad para comprobar si los recursos de una intervención pueden justificarse por sus resultados, que son de gran importancia práctica y política. Muchas partes interesadas, incluidos los beneficiarios, se preocupan por la eficiencia, ya que puede respaldar un mejor uso de los recursos limitados para obtener mayores logros.
- El impacto es la medida en que la intervención ha tenido o se prevé que tenga importantes efectos de carácter general, positivos o negativos, intencionales o no. El impacto tiene en cuenta la trascendencia última y los efectos potencialmente transformadores de la intervención: cambios integrales y duraderos en los sistemas o las normas. El criterio de impacto va más allá de la eficacia y fomenta la consideración de la gran pregunta: “¿Por qué importa?”. Aquí es donde los evaluadores determinan si la intervención creó, o no, un cambio que realmente importa a las personas.
- La sostenibilidad es la medida en la que se mantienen o pueden mantenerse los beneficios netos de la intervención. Dependiendo de la cronología de la evaluación, puede implicar el análisis del flujo real de beneficios netos o la estimación de la probabilidad de que los beneficios netos se mantengan en el mediano y largo plazo. Si bien el concepto subyacente de los beneficios continuos permanece sin cambios, la nueva definición abarca varios elementos para el análisis (financiero, económico, social y ambiental) y se debe prestar atención a la interacción entre ellos.

Pueden presentarse una variedad de retos al aplicar cada uno de los criterios, incluidos los problemas comunes relacionados con las deficiencias en el diseño de la intervención y la disponibilidad de datos. Se proporcionan sugerencias prácticas sobre cómo los evaluadores y los gestores de evaluaciones pueden superar estos desafíos, incluso mediante la evaluación temprana del proceso, el trabajo con las partes interesadas para documentar las decisiones tomadas o recrear las bases de referencia, y la transparencia respecto de cualquier limitación para manejar las expectativas. Estos cuadros, así como el resto de la guía, se actualizarán a medida que se recopilen más lecciones de la experiencia.

1 Propósito y uso de la guía: mejores criterios para una mejor evaluación

Este capítulo explica el propósito del documento y cómo puede ser de utilidad para que el lector o lectora entienda las definiciones adaptadas de los criterios y las utilice en su trabajo. Explica también cómo se elaboró la guía, entre otras cosas el rol de las partes interesadas de la evaluación a nivel mundial a la hora de fundamentar su diseño y su contenido. Por último, el capítulo explica cómo aplicar los criterios con minuciosidad para mejorar tanto la ejecución como el diseño de las evaluaciones.

Razones por las que se necesita una guía

El Comité de Ayuda al Desarrollo (CAD) de la OCDE estableció por primera vez cinco criterios de evaluación en 1991 (OCDE, 1991^[1]) y proporcionó definiciones para ellos en 2002 (OCDE, 2002^[2]). Estos cinco criterios —pertinencia, eficacia, eficiencia, impacto y sostenibilidad— constituyen ahora un referente común para los evaluadores de la cooperación para el desarrollo y muchas otras partes interesadas.

En respuesta al deseo de cumplir con la Agenda 2030 y a la oportunidad de aprovechar la considerable experiencia y los extensos aprendizajes desde que se definieron los criterios por primera vez, entre 2017 y 2019 se llevó a cabo un proceso de consulta y adaptación de gran alcance, y el CAD aprobó las definiciones revisadas de los criterios en diciembre de 2019 (OCDE, 2019^[3])¹. Durante el proceso de consulta, los usuarios manifestaron que encontraban útiles muchos aspectos de los criterios de evaluación y que deseaban continuar utilizándolos en su evaluación de futuras intervenciones. La retroalimentación mostró que los criterios de evaluación se aceptaban y aplicaban ampliamente, lo que dio lugar a consistencia y cierto nivel de comparabilidad². Han creado un lenguaje común, simple y neutro, así como un marco normativo para el sector que es adaptable, enfatiza los resultados y cubre adecuadamente las principales áreas requeridas para la rendición de cuentas y el aprendizaje. En consecuencia, los criterios revisados reflejan el apoyo a su continuidad y amplían los aspectos que funcionaron adecuadamente en las definiciones originales.

La guía debe leerse en conjunción con los Estándares de Calidad para la Evaluación del Desarrollo (OCDE, 2010^[4]) y los Principios de Evaluación de la Ayuda al Desarrollo (OCDE, 1991^[1]) de la OCDE. Los Principios se centran en la gestión y la configuración institucional de los sistemas de evaluación, mientras que los Estándares fundamentan principalmente los procesos y productos de evaluación. La presente guía ayuda a los usuarios a operativizar las definiciones de los criterios de evaluación, ya que los analiza en mayor detalle, ofrece ejemplos y examina los desafíos prácticos que se suelen encontrar durante la aplicación de los criterios en el transcurso de las evaluaciones. El Recuadro 1.1 muestra futuras publicaciones que servirán de ayuda para la aplicación y la comprensión de los criterios.

En los últimos 25 años, se ha aprendido mucho sobre cómo utilizar los criterios establecidos por el CAD. La consulta realizada entre 2017 y 2019 señaló como principal desafío la aplicación práctica de los criterios. Los debates recalcaron la importancia de cómo deberían utilizarse los criterios, y no solo cómo deberían definirse. Por ejemplo, en ocasiones los criterios se terminan aplicando mecánicamente, sin adaptarse de manera adecuada al contexto. La idea de la guía es, por lo tanto, apoyar a los diferentes grupos de usuarios en la aplicación minuciosa de los criterios en la práctica, lo que contribuirá a producir evaluaciones de alta calidad y utilidad.

En consonancia con lo anterior, el documento incluye terminología y orientaciones para ayudar a los lectores a comprender y aplicar las definiciones de los criterios, incluidos los dos principios desarrollados para orientar su uso (OCDE, 2019^[3]). La presente guía desarrolla estos dos principios y explica la definición de cada criterio, lo que permitirá a los usuarios interpretar y aplicar los criterios en su trabajo.

Propósito de la guía y cómo utilizarla

El presente documento es una guía sucinta para ayudar a los usuarios a interpretar los criterios y aplicarlos en su trabajo con minuciosidad. Si bien los criterios se utilizan en primer lugar en la evaluación, tienen una amplia aplicabilidad en el seguimiento y la gestión basada en los resultados, el diseño de intervenciones o la planificación estratégica. La guía está pensada para ayudar a los usuarios a analizar detalladamente algunos de los desafíos que pueden surgir y dónde buscar las soluciones. Ayuda a los lectores a entender todos los componentes de las definiciones, cómo se relacionan los criterios entre sí,

Recuadro 1.1. Otras investigaciones y próximas publicaciones de la OCDE sobre evaluación

Además de revisar esta guía a partir de la experiencia con las nuevas definiciones, EvalNet y los asociados también están desarrollando guías específicas para tratar determinados temas en mayor profundidad, entre ellos el uso de los criterios desde la perspectiva de la igualdad de género y la aplicación de los criterios en contextos humanitarios. En colaboración con EvalNet, ALNAP está igualmente en proceso de actualizar su guía de amplia utilización para la evaluación de la acción humanitaria y el uso de los criterios (Beck, 2016^[1]).

Traducciones de las definiciones de los criterios y los principios

Trabajando con asociados clave y en concordancia con el mandato de apoyar el aprendizaje y el desarrollo de capacidades de evaluación, la OCDE está traduciendo las definiciones de los criterios a varios idiomas, además de los idiomas oficiales de la OCDE, el inglés y el francés. Ya existen traducciones al español y al chino y, para 2021, se espera contar con nuevas versiones en árabe, persa darí, pastún, portugués, ruso, tailandés y urdu.

Segunda edición del Glosario de los principales términos sobre evaluación y gestión basada en resultados

En el Glosario de los principales términos sobre evaluación y gestión basada en resultados de la OCDE (OCDE, 2002^[2]), se publicaron definiciones por primera vez para los cinco criterios originales. Actualmente está en elaboración una segunda edición del Glosario, que incorpora las actualizaciones de los criterios y otros términos. La segunda edición (de próxima aparición) será un valioso acompañante de esta guía, dado que ofrece definiciones sucintas de varios términos utilizados a lo largo del texto, entre ellos “resultados”, “efectos directos”, “impactos”, “teoría del cambio” y “beneficiario”.

los desafíos que surgen durante su aplicación y cómo afrontarlos. Hemos incluido una selección de ejemplos prácticos que ilustrarán cómo aplicar los criterios con minuciosidad en escenarios de la vida real.

La guía no se ha diseñado para usarse como un manual, una plantilla o una lista de verificación de una evaluación. Es una fuente a la cual recurrir en la etapa de diseño de los planes de evaluación y puede ser útil asimismo para complementar las políticas, normas y orientaciones institucionales por las que se rigen los evaluadores.

De la misma manera, la guía no sustituye las directrices y los estándares éticos ya existentes que han sido elaborados por otras instituciones o propuestos por el CAD de la OCDE (OCDE, 1991^[1]); (OCDE, 2010^[4]). La guía complementa los estándares internacionales utilizados por la Organización de las Naciones Unidas y otros organismos multilaterales, entre ellos las normas y los estándares de evaluación del Grupo de Evaluación de las Naciones Unidas (UNEG), los estándares de buenas prácticas del Grupo de Cooperación en Materia de Evaluación (ECG) y la guía de la Red para un Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria (ALNAP). Otros recursos valiosos son los sitios web de las unidades de evaluación bilaterales y de los organismos de las Naciones Unidas, así como otras unidades de evaluación multilaterales, el sitio web de Better Evaluation y los sitios web de EvalPartners y sus organismos conexos, que inciden en áreas específicas, como la Iniciativa Internacional para Evaluación de Impacto (3ie).

Del mismo modo, al igual que los propios criterios, la guía está diseñada para estimular la autorreflexión y el pensamiento crítico, y no para utilizarse mecánicamente. Un enfoque mecánico de esta naturaleza no es congruente con las prácticas idóneas y los principios de la Agenda 2030 (que se describen en mayor detalle en el capítulo 3). La guía no prescribe herramientas o enfoques metodológicos específicos, ya que

los criterios no constituyen una metodología. En el documento se han incluido algunas referencias a otras fuentes que sí contienen explicaciones detalladas de las metodologías como punto de partida.

La guía puede interpretarse como un puente entre los documentos formales centrales, como las definiciones de los criterios (OCDE, 2019^[3]) y los Estándares de Calidad de la OCDE (OCDE, 2010^[4]), que contienen definiciones y términos consistentes y claros, y la aplicación cotidiana de los criterios en instituciones específicas de manera minuciosa y apropiada, con las muchas decisiones requeridas que se toman en función del contexto.

Además de proporcionar ejemplos de cómo se han aplicado los criterios en evaluaciones reales, la guía incluye elementos de análisis que explican los conceptos contenidos en cada definición. Dichos elementos no pretenden ser definitivos ni están pensados para su uso de manera discrecional ni como subcriterios, sino que ilustran cada definición y ayudan a desmenuzar sus componentes. Ayudarán a los usuarios a encontrar la interpretación contextual más apropiada de cada criterio para una evaluación individual. Durante la fase de diseño, el diálogo entre los evaluadores, los comisionados y los usuarios previstos de los hallazgos de la evaluación puede explorar asimismo cómo deberían valorarse los diferentes elementos de cada criterio y cuáles son los más importantes para extraer conclusiones generales, ya que esto variará en función de la intervención, del contexto y del uso previsto de la evaluación.

El capítulo 2 introduce la función de los criterios de evaluación y la medida en que estos son compatibles con las normas y los estándares más amplios. El capítulo 3 ofrece orientaciones para ayudar a los evaluadores a interpretar los criterios y aplicarlos con minuciosidad, y también a utilizar los criterios en su contexto de políticas. El capítulo 4 contiene una descripción más detallada de cada definición, llama la atención sobre los desafíos comunes y proporciona ejemplos prácticos. Reflejando la prioridad de política de la Agenda 2030 de “no dejar a nadie atrás”, el capítulo 4 ofrece un resumen de los aspectos clave relacionados con la inclusión que se deben considerar para cada criterio.

Dado que la mayoría de los ejemplos han sido tomados de evaluaciones sustentadas en las definiciones originales y que las definiciones y los principios nuevos solo se están empezando a aplicar, esta guía debe verse como un proyecto de trabajo. Se utilizará durante una fase de validación de su aplicación inicial y se revisará posteriormente, al cabo de varios años, a efectos de incorporar el aprendizaje adicional y actualizar los ejemplos. La Secretaría está especialmente interesada en recabar ejemplos más allá de la membresía de EvalNet, así como lecciones sobre el uso de los criterios en evaluaciones locales y nacionales.

Personas que pueden utilizar esta guía

La guía es aplicable para su uso en una amplia variedad de contextos: locales, nacionales e internacionales. Los criterios son deliberadamente amplios y han sido diseñados para su aplicación a toda una gama de intervenciones y sectores. Su pertinencia trasciende las intervenciones tradicionales de asistencia oficial soberana al desarrollo. Aplicados con minuciosidad, los criterios pueden ser igualmente útiles para intervenciones vinculadas al sector privado o de índole público-privada, o para áreas de política similares como el comercio, entre otras.

Esta guía está pensada en primer lugar para evaluadores y gestores y comisionados de evaluaciones. Es apropiada tanto para estudiantes o profesionales jóvenes que utilizan los criterios por primera vez como para profesionales más curtidos que trabajan en evaluaciones complejas. Será de interés asimismo para orientar el diseño, la planificación estratégica y la gestión de los resultados de una intervención. Una mejor evaluación depende en parte de la integración eficaz de sistemas de seguimiento, evaluación y aprendizaje en el ciclo de los programas o las políticas. La guía será de utilidad para ayudar a los evaluadores y a los asociados operativos a desarrollar una visión compartida y un lenguaje común en torno a en qué consistirán los logros desde un inicio (es decir, en la etapa de diseño de la intervención),

aumentando con ello la evaluabilidad de las intervenciones, al tiempo que se mejoran los resultados generales.

Referencias

- Beck, T. (2016), *Evaluación de la acción humanitaria utilizando los criterios del CAD-OCDE*, [5]
<https://alnap.org/help-library/resources/evaluating-humanitarian-action-using-the-oecd-dac-criteria-spanish/> (fecha de consulta de la publicación original: 11 de enero de 2021).
- OCDE (2010), *Estándares de Calidad para la Evaluación del Desarrollo*, Serie: Directrices y [4]
 Referencias del CAD, OECD Publishing, París,
https://www.oecd.org/es/publications/estandares-de-calidad-para-la-evaluacion-del-desarrollo_9789264094949-es.html.
- OCDE (2002), *Evaluation and Aid Effectiveness No. 6, Glosario de los principales términos sobre evaluación y gestión basada en resultados*, OECD Publishing, París, [2]
<https://doi.org/10.1787/9789264034921-en-fr>.
- OCDE (2019), *Mejores criterios para una mejor evaluación: Definiciones revisadas de los criterios de evaluación y principios para su utilización*, Red de Evaluación del Desarrollo del [3]
 CAD, OECD Publishing, París, https://www.oecd.org/es/publications/2019/12/better-criteria-for-better-evaluation_f7a307eb.html (fecha de consulta de la publicación original: 11 de enero de 2021).
- OCDE (1991), *Principles for Evaluation of Development Assistance*, Comité de Ayuda al [1]
 Desarrollo, OECD Publishing, París,
<https://www.oecd.org/development/evaluation/2755284.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021).

Notas

¹ El documento “Mejores criterios para una mejor evaluación” (OCDE, 2019^[3]) contiene una explicación detallada de las modificaciones hechas a las definiciones originales y la justificación de dichos cambios.

² Ver un resumen de los resultados de la consulta en el portal web de EvalNet: oe.cd/criteria

2 Los seis criterios: propósito y función dentro de la evaluación

El presente capítulo incide en el propósito de los seis criterios de evaluación y explora el papel que cumplen dentro de la evaluación. Explica cómo los criterios pueden verse como una serie de lentes que se complementan y la forma en que cada criterio puede aportar una perspectiva distinta para entender la intervención y sus resultados. El capítulo considera además diferentes tipos de intervenciones susceptibles de evaluación aplicando los criterios como base para juzgar su mérito o valor. La sección final explora cómo se relacionan los criterios con otras normas y estándares de evaluación.

¿Qué es un criterio?

Un criterio es un estándar o un principio utilizado en la evaluación como base para el juicio evaluativo.

Cada uno de los seis criterios se encuentra resumido en una pregunta general, que ilustra su significado más amplio. Cada uno representa un importante elemento que se debe considerar:

- Pertinencia: ¿La intervención¹ está haciendo lo correcto?
- Coherencia: ¿En qué medida es compatible la intervención?
- Eficacia: ¿La intervención está logrando sus objetivos?
- Eficiencia: ¿Se están utilizando adecuadamente los recursos?
- Impacto: ¿Qué diferencia genera la intervención?
- Sostenibilidad: ¿Serán duraderos los beneficios?

El propósito de los criterios de evaluación es fomentar una evaluación consistente y de alta calidad dentro de un marco de referencia común. Los criterios configuran un marco normativo con el que evaluar una determinada intervención. Pueden utilizarse asimismo en procesos que trascienden la evaluación, por ejemplo en la definición de los marcos e indicadores para los mecanismos de seguimiento y gestión basada en los resultados, en la aprobación de financiamiento, en la planificación estratégica y en el diseño de la intervención, en especial para mejorar intervenciones futuras. En términos colectivos, resulta de gran valor contar con criterios que gozan de una definición común y se aplican de manera similar a todas las intervenciones. Los criterios ofrecen también un lenguaje consistente en todo el ámbito del desarrollo, lo que ofrece un alto grado de estandarización y permite la comparación y el aprendizaje entre intervenciones.

Los criterios se deben considerar como un conjunto de lentes a través de los cuales se puede comprender y analizar una intervención. Los criterios aportan perspectivas que se complementan y ofrecen una visión panorámica de la intervención. Fomentan una reflexión más profunda en torno a la naturaleza de la intervención, su ejecución, su proceso y sus resultados. En conjunto, describen los atributos deseados de las intervenciones, plantean supuestos explícitos y establecen normas: que las intervenciones deben ser pertinentes en función del contexto, guardar coherencia con otras intervenciones, producir resultados de manera eficiente y generar impactos positivos y duraderos para el desarrollo sostenible. Además, los criterios de evaluación suministran un marco de amplia aceptación para desarrollar un enfoque de evaluación, un método sistemático y completo, y una terminología común que se utiliza desde el inicio mismo del proceso de evaluación.

Los criterios están relacionados entre sí y pueden verse como un conjunto que se complementa al que cada criterio aporta una perspectiva diferente y única para la comprensión de la intervención. La definición de cada criterio presenta un concepto distintivo (recogido en las preguntas centrales para cada criterio) y, sin embargo, dichos conceptos se encuentran interconectados en muchos sentidos. Las intervenciones para el desarrollo suelen presentar múltiples facetas, por lo que utilizar los criterios de manera relacional puede ayudar al evaluador a ver la intervención como un todo.

Los criterios de evaluación no constituyen una metodología ni son los objetivos que la intervención está intentando alcanzar, sino que proporcionan claves para hacer las preguntas correctas durante la evaluación de una intervención. Cada intervención (el objeto o sujeto de la evaluación²) es diferente, razón por la cual el proceso de evaluación debe ser flexible y el uso de los criterios de evaluación se debe analizar y adaptar minuciosamente en función del propósito y de los usuarios. De conformidad con los estándares de calidad de la evaluación, las conclusiones y recomendaciones respecto del avance y los resultados deben sustentarse en evidencia apropiada y creíble. La lógica, la credibilidad y la interpretación de la evidencia no deben generar dudas y deben guiarse por la teoría del cambio o la lógica de la intervención.

¿Qué tipos de intervenciones podemos evaluar con estos criterios?

El término "intervención" se utiliza a lo largo de esta guía para denominar el tema u objeto de evaluación. Comprende todos los diferentes tipos de esfuerzos que se pueden evaluar en función de estos criterios. Puede tratarse de esfuerzos internacionales o nacionales de apoyo a objetivos de desarrollo sostenible o humanitarios.

Las intervenciones pueden ser proyectos, programas, políticas, estrategias, áreas temáticas, asistencia técnica, asesoría sobre políticas, instituciones, mecanismos de financiamiento, instrumentos u otras actividades. El término abarca intervenciones de desarrollo, asistencia humanitaria, apoyo a la construcción de la paz, trabajo normativo y otras actividades de cooperación internacional, así como iniciativas de actores del sector privado y de gobiernos nacionales en contextos de política interna. Los criterios han sido utilizados, por ejemplo, para evaluar temas que van desde la eficiencia de los programas nacionales de alimentación escolar hasta la coherencia del apoyo internacional de una serie de actores a una región en conflicto. También se utilizan para evaluar políticas o estrategias (ver ejemplo en el Recuadro 2.1). Los criterios se utilizan adicionalmente en la evaluación de un grupo de intervenciones (por ejemplo, una evaluación de varios proyectos de apoyo al sector educación) o de una única "intervención" de gran magnitud que se extiende a lo largo de muchos años con participación de múltiples asociados (por ejemplo, apoyo al presupuesto general).

El término "intervención" se utiliza en este documento porque es más fácil de entender, pero los evaluadores deberán esforzarse por definir claramente el tema de la evaluación (la intervención que se está analizando), así como el ámbito de la evaluación, en una etapa temprana del proceso. El uso de otros términos, más específicos, para referirnos a la intervención en cuestión —por ejemplo, proyecto o política— probablemente será más útil y fácil de entender para los asociados.

Recuadro 2.1. Cómo adaptar los criterios a la intervención: evaluación del plan de acción para la inclusión de las personas con discapacidad del BMZ

La evaluación del plan de acción para la inclusión de las personas con discapacidad del Ministerio Federal de Cooperación Económica y Desarrollo de Alemania (BMZ) es un interesante ejemplo de cómo los criterios se pueden adaptar a las particularidades de la intervención objeto de evaluación (en este caso, un plan de acción estratégico) y cómo se conectan con las preguntas de evaluación. La evaluación se diseñó para examinar en qué medida el plan de acción había logrado transversalizar la inclusión de manera sistemática en la cooperación para el desarrollo alemana. La evaluación se basó en las preguntas de evaluación generales que se señalan a continuación, que se tomaron del propio plan de acción. Cada una de estas preguntas de carácter general se subdividió en preguntas más específicas para cada criterio:

1. ¿En qué medida el BMZ ha demostrado ser un buen ejemplo en su propia organización en términos de inclusión de las personas con discapacidad?

- Pertinencia: ¿En qué medida son congruentes las actividades y los ámbitos de acción seleccionados con las disposiciones de la Convención Internacional de las Naciones Unidas sobre los Derechos de las Personas con Discapacidad (CDPD)?
- Eficacia: ¿En qué medida está teniendo éxito el BMZ en establecer estructuras y prácticas inclusivas?

2. ¿En qué medida contribuye el plan de acción a impulsar la inclusión de las personas con discapacidad en los países asociados de la cooperación para el desarrollo alemana?

- Pertinencia: ¿En qué medida son congruentes las actividades y los ámbitos de acción seleccionados con las disposiciones de la CDPD?
- Eficacia: ¿En qué medida se ha incluido a las personas con discapacidad con mayor eficacia en las actividades de la cooperación para el desarrollo alemana en los países asociados?

3. ¿En qué medida actúa el BMZ a nivel nacional, regional e internacional como activista y asociado de los derechos de las personas con discapacidad en la cooperación para el desarrollo?

- Pertinencia: ¿En qué medida son congruentes las actividades y los ámbitos de acción seleccionados con las disposiciones de la CDPD?
- Eficacia: ¿En qué medida ha logrado sumar el BMZ a otros actores bilaterales y multilaterales a la causa de la inclusión?

4. ¿Cómo se gestionaron el diseño y la ejecución del plan de acción para la inclusión?

- Eficacia: ¿Qué mecanismos y estructuras de gestión (incluidos los roles de los diferentes grupos de partes interesadas en el proceso de gestión) tomaron las decisiones respecto del diseño y la ejecución del plan de acción?
- Eficiencia: ¿En qué medida la ejecución de las actividades se ciñó al plan de acción?

5. ¿Cómo deberíamos calificar los beneficios del plan de acción para la inclusión en términos de su amplitud de impacto e impulso como instrumento de gobernanza?

- Impacto: ¿En qué medida el plan de acción generó un amplio impacto?
- Impacto: ¿En qué medida el plan de acción generó impulso en términos generales?

Fuente: Schwedersky, Ahrens y Steckhan (2017^[2]), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, <https://d-nb.info/1186644206/34>

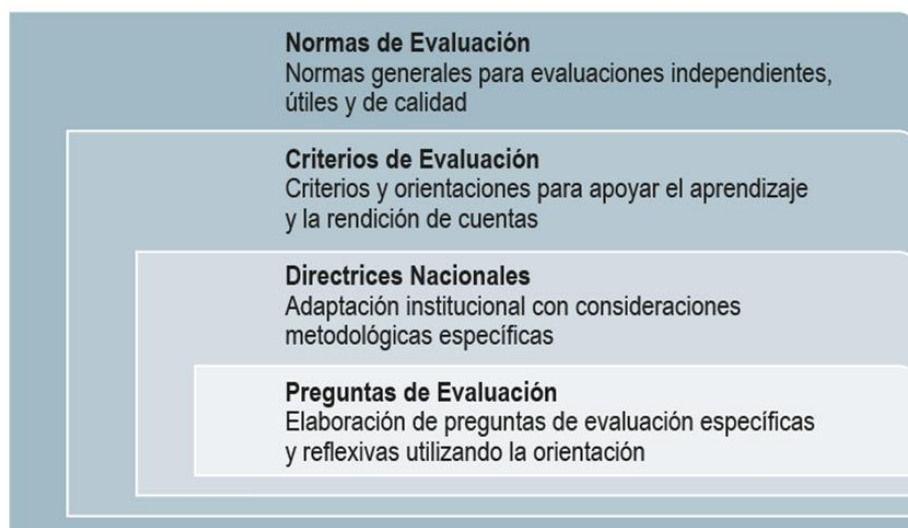
¿En qué medida los criterios son compatibles con el ámbito de evaluación más amplio?

Dado que cada intervención (la entidad que está siendo evaluada) es diferente, el proceso de evaluación debe ser flexible y los criterios de evaluación se deben usar con minuciosidad y se deben adaptar al propósito y a los usuarios. El propósito de los criterios no es, por tanto, ofrecer un conjunto de normas de cumplimiento obligatorio.

Los criterios no existen de manera aislada. Para darles el mejor uso posible, es importante establecer en qué medida son compatibles con otras normas y estándares, metodologías y directrices institucionales.

- El nivel macro es el nivel en que se sitúan los principios de evaluación básicos: imparcialidad e independencia, credibilidad, utilidad y participación (OCDE, 1991^[2]), así como los Estándares de Calidad para la Evaluación del Desarrollo (OCDE, 2010^[3]). En este nivel se ubican igualmente los estándares éticos que rigen la investigación y la recolección de datos. Estos establecen los principales estándares de buenas prácticas que toda evaluación debería seguir a lo largo de todo el proceso.
- Los criterios de evaluación se ubican en el nivel meso, de modo que proporcionan los lentes a través de los cuales la intervención se analiza y se comprende.
- El nivel institucional es el nivel en que cada organización adapta los criterios y los traduce a sus propias directrices y políticas institucionales (por ejemplo, enfatizando ciertos criterios, ordenando poner calificaciones o añadiendo criterios adicionales), de modo tal que reflejen sus propios mandatos y prioridades.
- El nivel micro, es decir, el contexto de cada evaluación individual, es donde se ubican las decisiones en torno a las preguntas y las metodologías de evaluación que se emplean en cada evaluación, y el modo en que se aplican los criterios en ese contexto determinado y a ese sujeto de la evaluación en particular.

Gráfico 2.1. En qué medida son compatibles los criterios con otras normas y estándares



Referencias

- ADE (2019), *Evaluación conjunta del modelo de soluciones integrales en Kalobeyei y alrededores, Turkana, Kenia*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d. [5]
- Bamberger, M., J. Vaessen and E. Raimondo (2015), “Dealing With Complexity in Development Evaluation - A Practical Approach”, https://www.betterevaluation.org/en/resources/dealing_with_complexity_in_development_eval. [23]
- Beck, T. (2016), *Evaluating Humanitarian Action using the OECD-DAC Criteria*, <https://www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria>. [1]
- BMZ (2020), *Evaluation Criteria for German Bilateral Development Co-operation*. [7]
- Bryld (2019), *Evaluación del apoyo de Sida a la construcción de la paz en situaciones de conflicto y posconflicto: informe de país de Somalia*, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5. [6]
- D’Errico, S. (n.d.), “Evaluation to connect national priorities with the SDGs | Publications Library”, <https://pubs.iied.org/17739IIED>. [20]
- Danida (2019), *¿De donantes a socios? Evaluación del apoyo noruego al fortalecimiento de la sociedad civil en países en desarrollo a través de organizaciones de la sociedad civil noruegas*, <http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf>. [4]

- Davis, R. (2013), “Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations”, Vol. No. 20, <https://www.gov.uk/government/publications/planning-evaluability-assessments>. [22]
- Eurecna Spa (2020), *Bolivia - Evaluation of Health Initiatives (2009-2020)*, http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf. [9]
- Helle, E. (2018), *¿De donantes a socios? Evaluación del apoyo noruego al fortalecimiento de la sociedad civil en países en desarrollo a través de organizaciones de la sociedad civil noruegas*, https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf. [3]
- IDD and Associates (2006), *A Joint Evaluation of General Budget Support Evaluation of General Budget Support: Synthesis Report*, <https://www.oecd.org/development/evaluation/dcdndep/37426676.pdf> (accessed on 11 January 2021). [10]
- Molund, S. and G. Schill (2004), *Looking Back, Moving Forward: Sida Evaluation Manual*, <https://www.oecd.org/derec/sweden/35141712.pdf>. [18]
- OCDE (2019), *Mejores criterios para una mejor evaluación: Definiciones revisadas de los criterios de evaluación y principios para su utilización, Red de Evaluación del Desarrollo del CAD*, <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>. [13]
- OCDE (2010), *Estándares de Calidad para la Evaluación del Desarrollo, Serie: Directrices y Referencias del CAD*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264094949-es>. [15]
- OCDE (2002), *Evaluation and Aid Effectiveness No. 6, Glosario de los principales términos sobre evaluación y gestión basada en resultados*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [12]
- OCDE (1991), *Principles for Evaluation of Development Assistance*, Comité de Ayuda al Desarrollo, OECD Publishing, Paris, <https://www.oecd.org/development/evaluation/2755284.pdf> (accessed on 11 January 2021). [11]
- OECD (2019), *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*, <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>. [17]
- OECD (2010), *Directrices y Referencias del CAD (series) Estándares de Calidad para la Evaluación del Desarrollo*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264094949-es>. [16]
- OECD (2002), *Evaluation and Aid Effectiveness No. 6 - Glossary of Key Terms in Evaluation and Results Based Management (in English, French and Spanish)*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [21]
- OECD (1991), *Principles for Evaluation of Development Assistance*, <https://www.oecd.org/development/evaluation/2755284.pdf>. [14]
- Ofir, E. (2016), “Resumen informativo: cinco consideraciones para agendas nacionales de evaluación informadas por los ODS”, <https://pubs.iied.org/sites/default/files/pdfs/migrate/17374IIED.pdf>. [8]

Schwedersky, T., L. Ahrens and H. Steckhan (2017), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, German Institute for Development Evaluation (DEval), Bonn, https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2017_API/DEval_Bericht_APIInklusion_2017_EN.pdf. [2]

WFP (2016), *Technical Note: Evaluation Methodology, DEQAS, World Food Programme*, <https://docs.wfp.org/api/documents/704ec01f137d43378a445c7e52dcf324/download/>. [19]

Notas

¹ El término “intervención” se utiliza aquí y a lo largo de todo el documento para denominar el objeto de evaluación (lo que está siendo evaluado). Ver el capítulo 3.2, Cómo adaptar los criterios al propósito de la evaluación, para un análisis más detallado.

² El término “sujeto de la evaluación” se refiere al objeto de evaluación: lo que está siendo evaluado

3

Uso de los criterios de evaluación en la práctica

El presente capítulo considera los aspectos prácticos relacionados con el uso de los criterios en el diseño y la ejecución de una evaluación. Revisa los dos principales principios que deberían guiar la aplicación de los criterios. El capítulo empieza explorando diferentes enfoques que pueden ayudar a una aplicación minuciosa de los criterios. A continuación, describe cómo se pueden aplicar los criterios en diferentes contextos institucionales, con distintas prioridades estratégicas, modalidades de trabajo y culturas. Describe el modo en que los criterios pueden incentivar a los evaluadores a considerar las experiencias y los impactos diferenciados aplicando un lente de género. Finalmente, examina cómo los criterios pueden ayudar a los evaluadores y gestores de evaluaciones a trabajar para fomentar el logro de los Objetivos de Desarrollo Sostenible y de la Agenda 2030 en general. El capítulo incluye una serie de ejemplos prácticos que ilustran cómo se han aplicado los criterios en diversas evaluaciones.

Como ayuda para los evaluadores y las personas que participan en el diseño o la gestión de intervenciones a desarrollar evaluaciones que sean de utilidad y apropiadas para diferentes contextos y partes interesadas, se han elaborado los dos principios que presentamos a continuación para orientar el uso de los criterios. Con el fin de evitar que se apliquen de manera mecánica, lo que resultaría desfavorable para el pensamiento crítico, la creatividad y la apropiación por los participantes, dichos principios se deben aplicar junto con los criterios cada vez que se utilicen (OCDE, 2029^[1]):

- **Primer principio:** Los criterios deben aplicarse minuciosamente para fomentar evaluaciones de alta calidad y utilidad.
- **Segundo principio:** El uso de los criterios depende del propósito de la evaluación.

La sección a continuación profundiza en estos dos principios y describe importantes conceptos adicionales para trabajar con los criterios, entre ellos cómo adecuar los criterios a contextos específicos, cómo examinar los criterios en diferentes coyunturas y cómo se relacionan los criterios entre sí.

Cómo aplicar los criterios con minuciosidad

El primer principio hace hincapié en que los criterios deben aplicarse minuciosamente. En la práctica, esto significa considerar de forma crítica qué criterios son más útiles para fomentar evaluaciones de alta calidad y utilidad que sean de valor para los usuarios previstos. El Recuadro 3.1 ofrece un ejemplo de cómo aplicar los criterios con minuciosidad en la evaluación de una intervención.

Considerar los siguientes seis aspectos y sus preguntas relacionadas ayudará a los evaluadores a aplicar los criterios con minuciosidad:

- **Contexto:** ¿Cuál es el contexto de la intervención en sí y cómo se entienden los criterios en el contexto de la evaluación individual, la intervención y las partes interesadas?
- **Propósito:** ¿Qué está tratando de conseguir la evaluación y qué preguntas son más útiles para trabajar en pos de ese objetivo y alcanzarlo?
- **Funciones y dinámicas de poder:** ¿Quiénes son las partes interesadas y cuáles son sus respectivas necesidades e intereses? ¿Cuáles son las dinámicas de poder entre ellas? ¿Quiénes deben participar en las decisiones respecto de qué criterios aplicar y cómo entenderlos en función del contexto local? Aquí podrían incluirse preguntas acerca de la apropiación y sobre quién decide qué se evalúa y prioriza.
- **Intervención (sujeto de la evaluación):** ¿Qué tipo de intervención está siendo evaluado (un proyecto, una política, una estrategia o un sector)? ¿Cuál es su ámbito y su naturaleza? ¿En qué medida son directos o indirectos sus resultados previstos? ¿Qué pensamiento en sistemas complejos entra en juego aquí?
- **Evaluabilidad:** ¿Existen restricciones en términos de acceso, recursos y datos (incluidos los datos desagregados) que puedan causar un impacto en la evaluación y cómo afecta este hecho a los criterios?
- **Cronología:** ¿En qué etapa del ciclo de vida de la intervención se llevará a cabo la evaluación? ¿Ha cambiado con el tiempo el contexto en que está operando la intervención y, de ser así, cómo? ¿Se deben tener en cuenta estos cambios durante la evaluación? La cronología influirá en el uso de los criterios, así como en la fuente de evidencia.

Recuadro 3.1. Uso minucioso de los criterios: evaluación de la donación de la sociedad civil noruega para países en desarrollo

Esta evaluación examina la donación de la sociedad civil noruega para fortalecer la sociedad civil de una serie de países en desarrollo, contribuyendo con ello a una sociedad civil más robusta con mayores capacidades y aptitudes para promover la democratización, mejorar la situación de los derechos humanos y reducir la pobreza.

Esta evaluación aplicó los criterios de manera “minuciosa” y flexible, interpretándolos de modo tal que reflejaran la naturaleza específica de la intervención (asociación y desarrollo de capacidades con organizaciones de la sociedad civil) y teniendo en cuenta las prioridades noruegas y el contexto. La evaluación utiliza los siguientes criterios, adaptados en algunos aspectos:

- Impacto: definido en relación con los efectos directos clave, como la democracia y las oportunidades de generación de ingresos y, por lo tanto, de reducción de la pobreza.
- Pertinencia: ningún cambio.
- Eficacia: definida aquí en función de los objetivos específicos de la asociación (prestación de servicios, incidencia política y fortalecimiento de capacidades). Incluye algunos efectos directos de término medio, como la creación de espacios para la sociedad civil.
- Sostenibilidad: ningún cambio.

La evaluación utiliza un criterio adicional, “valor agregado”, definido como competencia profesional, competencia organizacional y financiera, y competencia para crear redes con organizaciones afines (*networking*). Tiene que ver con aspectos que, en otros contextos, podrían ser englobados por los criterios de eficacia y eficiencia.

Fuente: Helle *et al.* (2018⁹³), *From Donors to Partners? Evaluation of Norwegian Support to Strengthen Civil Society in Developing Countries through Norwegian Civil Society Organisations*, https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf

Cómo adaptar los criterios al propósito de la evaluación

El aspecto más importante de la decisión de cómo utilizar los criterios es vincularlos al objetivo superior de la evaluación y su contexto, y luego desarrollar los criterios y las preguntas de evaluación alrededor de este propósito. El Recuadro 3.2 contiene dos ejemplos de cómo definir el propósito de la evaluación.

Los criterios no están pensados para aplicarse de una manera estándar o fija para cada intervención ni para marcarse como los casilleros de una lista de verificación. De hecho, los criterios deben interpretarse o entenderse cuidadosamente en relación con la intervención objeto de evaluación. Esto fomenta la flexibilidad y la adaptación de los criterios en función de cada evaluación individual. Es necesario aclarar de qué conceptos específicos de los criterios se valdrá la evaluación y por qué.

El propósito de la evaluación debe definirse con sumo cuidado y claridad. Es necesario incluir a las partes interesadas de la evaluación en esta etapa, a fin de cerciorarse de que comprendan el objetivo de la evaluación y el uso que se le dará.

Las preguntas clave en las que hay que concentrarse para definir el propósito de la evaluación son:

1. ¿Cuáles son las exigencias de una evaluación, quiénes son los destinatarios previstos y cómo utilizarán los hallazgos?

2. ¿Qué cosas resultan viables a la luz de las características y el contexto de la intervención?
3. ¿Qué grado de certidumbre se requiere para responder las preguntas centrales?
4. ¿Cuándo se necesita la información?
5. ¿Qué se conoce ya de la intervención y sus resultados? ¿Quiénes poseen ese conocimiento y cómo lo están utilizando?

Las decisiones subsiguientes acerca de los enfoques metodológicos, el diseño, la ejecución y la gestión del proceso de evaluación se deben fundamentar en los estándares éticos y de calidad.

Durante el proceso de adecuación de los criterios a una evaluación específica, es importante asimismo ahondar en la causalidad y considerar la medida en que la evaluación podrá atribuir los efectos a la intervención objeto de evaluación. Estas consideraciones pueden ayudar a gestionar las expectativas de las partes interesadas y a decidir en qué criterios profundizar. Esto es especialmente importante en el caso de la eficacia y el impacto, que se analizan en las dos secciones siguientes, pero puede aplicarse también de manera indirecta a otros criterios y serles beneficiosos. Por ejemplo, la eficiencia y la sostenibilidad podrían valerse de los beneficios reales (o previstos) que se atribuyen a la intervención como si hubieran sido evaluados bajo los criterios de eficacia e impacto. El análisis de la evaluabilidad es una herramienta de utilidad para verificar que el propósito sea comprendido en profundidad¹. El capítulo 4 ofrece una explicación más detallada de cómo interpretar cada criterio.

Recuadro 3.2. Definición del propósito de la evaluación: ejemplos de Uganda, Kenia y Suecia

Evaluación de la Agencia Danesa de Desarrollo Internacional (DANIDA) de los programas daneses de agua potable, saneamiento y medio ambiente en Uganda

El Departamento de Evaluación del Ministerio de Relaciones Exteriores de Dinamarca ordenó una evaluación independiente de las iniciativas danesas de agua potable, saneamiento y medio ambiente en Uganda entre 1990 y 2017.

Los objetivos específicos de la evaluación fueron:

- documentar los resultados y los logros de los subsectores;
- analizar el “valor agregado” del apoyo de DANIDA a los subsectores;
- extraer las lecciones aprendidas.

Con este fin, la evaluación se enfocó en evaluar la eficacia en dos etapas: 1) a nivel sectorial y transversal, y 2) a nivel subsectorial. La evaluación analizó asimismo el valor de las iniciativas y evaluó la sostenibilidad de los resultados.

Evaluación conjunta del Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) y Dinamarca

La evaluación conjunta llevada a cabo por el ACNUR y Dinamarca de la respuesta a la situación de la población refugiada de Kalobeyi y sus alrededores, en Kenia, define el propósito como sigue:

“El propósito principal de la evaluación es contribuir al aprendizaje acerca del modelo de asentamiento integrado en Kalobeyi y sus alrededores. Al documentar las lecciones aprendidas de un esfuerzo concreto para combinar la acción humanitaria y la asistencia para el desarrollo a largo plazo, la intención es generar evidencia de las potencialidades y los desafíos del diseño y de la ejecución de un modelo de soluciones integradas.”

Evaluación del programa de la Agencia Sueca de Cooperación Internacional para el Desarrollo (Asdi) de apoyo a la construcción de la paz en situaciones de conflicto y posconflicto

El informe de evaluación es un buen ejemplo de cómo la evaluación refleja el propósito. Ofrece la siguiente explicación:

El propósito o el uso previsto de la evaluación es sistematizar las lecciones aprendidas a partir de la práctica de construcción de la paz. La evaluación servirá de insumo para el proceso de conceptualización y desarrollo del enfoque de construcción de la paz utilizado por la Agencia Sueca de Cooperación Internacional para el Desarrollo (Asdi). Estos enfoques, a su vez, influirán en la planificación estratégica y el diseño del apoyo futuro de Asdi en contextos afectados por situaciones de conflicto. Además, se espera que la evaluación contribuya a fortalecer la comprensión de la construcción de la paz como un concepto y una práctica.

El objetivo específico de la evaluación fue valorar la metodología de Asdi frente a la construcción de la paz a nivel estratégico en distintos contextos. Con este fin, la evaluación puso especial énfasis en cuatro de los criterios:

- pertinencia del trabajo de construcción de la paz de Asdi en relación con las necesidades del contexto y de los beneficiarios, y con las prioridades de política de Asdi;
- eficacia de su trabajo de construcción de la paz en términos de su contribución a los objetivos generales de construcción de la paz y de su capacidad para proporcionar un marco propicio para el trabajo de construcción de la paz de sus asociados;

- impacto de su trabajo de construcción de la paz;
- sostenibilidad de su trabajo de construcción de la paz.

En los cuatro criterios, se puso especial énfasis en los resultados en las áreas de igualdad de género, empoderamiento y derechos de la mujer, e inclusión de las minorías étnicas y los grupos marginados. En los cuatro países comprendidos en el estudio de caso, las evaluaciones se enfocaron en los aspectos de marginación que estaban vinculados con la construcción de la paz y que aparecen en la bibliografía y las entrevistas. Se consideró principalmente a los grupos marginados en razón de su origen étnico o pertenencia a una minoría. Esto se debe a que, en los cuatro casos, los factores étnicos formaban parte importante de las causas primordiales del conflicto.

Fuentes: DANIDA (2019^[4]), *Evaluation of Water, Sanitation and Environment Programmes in Uganda (1990-2017)*,

<http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf>;

ADE (2019^[5]), *Joint Evaluation of the Integrated Solutions Model in and around Kalobeyei, Turkana, Kenya*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d;

Bryld (2019^[6]), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf

Cómo entender la dimensión de temporalidad

Los criterios pueden aplicarse a diferentes momentos en el tiempo. Cada uno de los criterios se puede utilizar para evaluar una intervención antes, después y a lo largo de su ejecución. Del mismo modo, los criterios pueden evaluarse en distintos momentos del ciclo de vida de una intervención. Sin embargo, la interpretación de los criterios y de las fuentes de evidencia puede variar según el momento. Por ejemplo, antes de que una intervención haya tenido lugar, la eficacia y la sostenibilidad serían previsiones; mientras que, después de la intervención, se dispondrá de mayor información para extraer conclusiones más firmes.

Los criterios (y las preguntas de evaluación asociadas) deben reflejar por lo tanto los dos aspectos clave siguientes en relación con la cronología: 1) el momento en el ciclo de vida de la intervención en que tendrá lugar la evaluación; y 2) la etapa de la intervención o el punto en la cadena de resultados en que se enfocará la evaluación.

El marco conceptual de cada criterio no varía si una evaluación se está llevando a cabo antes, después o a lo largo de una intervención. Sin embargo, los datos y la evidencia disponibles para *valorar* los criterios y los métodos utilizados sí cambian. Los evaluadores deben tener presentes estas diferencias en el momento de analizar los (posibles) hallazgos con las partes interesadas de la evaluación, ya que pueden influir en la utilidad y la credibilidad percibidas de la evaluación. Por ejemplo, una evaluación de la sostenibilidad *ex ante* (es decir, que se lleva a cabo antes de que empiece la intervención) podría examinar las probabilidades de continuación de los beneficios de una intervención enfocándose en el diseño de la intervención y en la evidencia disponible sobre la validez de los supuestos acerca de la continuación de los beneficios previstos. Una vez concluida la intervención, la evaluación de la sostenibilidad examinaría si los beneficios realmente continuaron o no, esta vez sobre la base de datos y evidencia de los beneficios reales generados por la intervención.

Al examinar las cosas en retrospectiva, las evaluaciones deben tener en cuenta el contexto y los datos disponibles en ese momento para emitir juicios basados en expectativas razonables de lo que pudo o debió haberse hecho. Sería injusto juzgar las acciones de las personas que diseñaron programas pasados sobre la base de la información disponible hoy, pero que no les era conocida en aquel momento. Sin embargo, muchas evaluaciones han constatado que la información disponible no se aprovechó en toda su magnitud (cuando pudo haberse aprovechado), lo cual hubiese incrementado la pertinencia de la intervención. Un ejemplo sería que la población local no hubiera sido consultada o

involucrada en grado suficiente en el diseño de la intervención. Cabría esperar desde un punto de vista razonable que un descuido de esa naturaleza se hubiese evitado y así deberá anotarse en la evaluación.

Cómo utilizar los criterios para mejorar las preguntas de evaluación

Formular buenas preguntas de evaluación es un elemento indispensable del proceso de evaluación, y el uso de los seis criterios interactúa con el proceso de decidir las preguntas de evaluación y lo respalda². El proceso empieza con una reflexión acerca del propósito de la evaluación y de cómo será utilizada y por quién. Una colaboración efectiva con las partes interesadas como parte de un proceso participativo bien diseñado puede ayudar a los evaluadores y gestores a entender cómo se utilizará la evaluación. Una comprensión profunda de la intervención y su contexto, sus objetivos específicos y su teoría del cambio debería complementar este diálogo con las partes interesadas. Se puede programar un período inicial para explorar estos asuntos y generar (un pequeño número de) preguntas clave que la evaluación necesita responder.

Después de este paso inicial, los criterios de evaluación sirven como herramienta para revisar desde distintas perspectivas si se ha pasado algo por alto, lo cual permitirá perfeccionar y refinar las preguntas que son esenciales para el diseño de evaluación. Esto hace que el proceso devenga en sistemático y garantiza que la evaluación sea exhaustiva. Es un componente fundamental de lo que constituye una “mejor evaluación”. Se recomienda encarecidamente que esta etapa de desarrollo de las preguntas se acometa con un enfoque general claro y coherente.

La institución que encarga la evaluación puede haber tomado la decisión de evaluar las intervenciones en función de ciertos criterios. En ese caso, se deberán seguir los parámetros institucionales paralelamente a una aplicación minuciosa de los criterios con el fin de asegurar la consistencia.

Un ejemplo de lo anterior es el manual de evaluación de Asdi, que contiene ejemplos de preguntas estándar para cada uno de los criterios (Molund y Schill, 2004^[6]). Hay otros ejemplos de cómo las instituciones interpretan y analizan los criterios, entre otros la guía y las notas técnicas elaboradas por el Programa Mundial de Alimentos para los gestores de evaluaciones y evaluadores (PMA, 2016^[7]).

Relaciones entre los criterios

Los criterios constan de múltiples lentes a través de los cuales se puede apreciar una intervención y sus resultados. Están conectados entre sí, en el sentido de que los conceptos sobre los que se sostiene cada criterio contribuyen a analizar dimensiones complementarias del proceso de obtención de resultados. Por ejemplo, la eficacia y el impacto inciden en diferentes niveles de la cadena de resultados, dependiendo de cómo se hayan definido los objetivos específicos de la intervención. Los dos criterios están, por lo tanto, interconectados a lo largo de la cadena de causalidades.

Los criterios a menudo dependen unos de otros. Por ejemplo, una intervención que no era pertinente para las prioridades de los beneficiarios tiene pocas probabilidades de tener el impacto deseado (a no ser que existan algunos canales de impacto separados, lo cual es más bien poco común). Una intervención que se lleva a cabo de manera deficiente (menos eficaz) probablemente será también menos sostenible. Por otro lado, también es posible que una intervención sea altamente pertinente, pero ineficaz; o muy coherente, pero ineficiente en razón de los altos costos de transacción. Los evaluadores deben explorar y reflexionar acerca de las relaciones y las sinergias entre los diferentes criterios, considerando, entre otras cosas, si están causalmente relacionados entre sí y de qué manera.

La mayoría de las evaluaciones extraen conclusiones con base en los hallazgos para cada criterio, así como una conclusión general basada en todos los criterios, en ocasiones utilizando un puntaje numérico

para calificar el desempeño. Al extraer conclusiones acerca de la intervención, y dependiendo del propósito de la evaluación, los evaluadores deben mirar todo el panorama completo y considerar cómo ponderar apropiadamente todos los criterios aplicados. Algunas instituciones ponderan los criterios definiendo un criterio dominante (también conocido como criterio *knock out*). Si el desempeño no es satisfactorio en lo que se refiere a ese criterio, no importa si los otros criterios obtuvieron puntajes altos: se considerará fallida la intervención (o, en el caso de una evaluación *ex ante*, se decidirá no financiar la intervención).

En el capítulo 4, se analizan vínculos más específicos entre los distintos criterios.

Cómo elegir qué criterios utilizar

Tal como se ha descrito anteriormente, el propósito, las prioridades, el ámbito y el contexto de la intervención y de la evaluación determinarán el énfasis relativo en diferentes criterios por encima de otros. Los evaluadores deben considerar el valor relativo que añadirá cada criterio. Esto implica dos decisiones básicas: ¿Constituye este criterio una consideración importante para esta evaluación? ¿Es factible responder preguntas acerca de este criterio?

Aunque los usuarios pueden sentirse tentados a simplemente aplicar los seis criterios independientemente del contexto, el mejor enfoque (es decir, un enfoque acorde a la intención original de los criterios y que genera la evaluación de más alta calidad) consiste en seleccionar y utilizar deliberadamente los criterios de formas que sean apropiadas para la evaluación y para las preguntas que la evaluación está buscando responder.

Para ello, es conveniente hacer las siguientes preguntas:

- Si pudiéramos hacer una sola pregunta en relación con esta intervención, ¿cuál sería?
- ¿Qué preguntas es mejor responder por medio de una evaluación y cuáles se podrían contestar por otros medios (por ejemplo, un proyecto de investigación, una síntesis de evidencia, un ejercicio de seguimiento o un proceso de aprendizaje facilitado)?
- ¿Los datos disponibles son suficientes para dar una respuesta satisfactoria a esta pregunta? De no ser así, ¿habrá mejores datos o mayor disponibilidad de datos más adelante?
- ¿Quién ha aportado insumos a la lista de preguntas? ¿Faltan algunas perspectivas importantes?
- ¿Tenemos el tiempo y los recursos suficientes para tratar todos los criterios de interés o, si circunscribimos el análisis a solo algunos de los criterios, obtendremos información más valiosa?

Es importante identificar un término medio entre flexibilidad (evitar una aplicación mecánica de todos los criterios) y “selección sesgada” (limitarse a elegir solo los criterios más fáciles o los más susceptibles de generar resultados positivos) en el momento de utilizar los criterios. Por ejemplo, no debemos inhibirnos de responder preguntas cruciales acerca del impacto y de la coherencia, aunque estas, en ocasiones, puedan representar un mayor desafío. El gráfico 3.1 a continuación muestra algunos aspectos que debemos considerar, recordando que se trata solo de ejemplos, y no de una lista de verificación exhaustiva.

Gráfico 3.1. Cómo cubrir los criterios para una evaluación específica



Un buen conocimiento de las partes interesadas implicadas (tanto en la intervención como en la evaluación) puede contribuir a identificar las potenciales tensiones entre sus diferentes intereses y prioridades en lo que atañe al diseño y a la ejecución de la evaluación. La población beneficiaria puede estar más interesada en determinar la eficacia (por ejemplo, si la salud de sus niños y niñas está mostrando mejoras a raíz de su participación en un programa de tratamiento de la desnutrición), mientras que la prioridad de los entes de ejecución puede ser determinar la eficiencia, con la mirada puesta en aumentar el número de familias beneficiarias del programa. En la mayoría de los casos, no es posible responder todas las preguntas potencialmente interesantes en una sola evaluación, y se tendrán que tomar decisiones al respecto. Para aumentar las probabilidades de que las preguntas que se han omitido sean abordadas en otro documento, es una práctica recomendable documentar el proceso y los efectos directos de los debates en torno a la priorización de las necesidades de las diferentes partes interesadas y las decisiones acerca de las preguntas de evaluación.

Cómo operativizar y adaptar los criterios a nivel institucional

Las definiciones de los criterios y esta guía proporcionan una plataforma común y un conjunto de definiciones establecidas de común acuerdo en las que basarse. No obstante, es preciso adecuarlas al contexto institucional. Los evaluadores y los gestores de evaluaciones a menudo utilizan los criterios en nombre de una organización de desarrollo, un ministerio o alguna otra institución que tiene su propio mandato, sus propias prioridades de política y sus propios estándares, guías y políticas de evaluación, todo ello según lo acordado por sus órganos rectores.

Al considerar cómo operativizar y aplicar los criterios, es importante que los evaluadores y los comisionados tengan muy en cuenta las prioridades estratégicas, la cultura y las oportunidades de la

organización como antecedente para la toma de decisiones. La forma en que se utilizan ciertos términos (por ejemplo, el impacto) varía, y es importante prestar mucha atención a las potenciales confusiones o malas interpretaciones de los criterios y su énfasis previsto. Esto ayudará a los evaluadores y comisionados a medida que aplican los criterios, maximizan el uso de los hallazgos de la evaluación y mejoran la pertinencia para las necesidades de los usuarios previstos. Los gestores de la evaluación deben fomentar un debate entre los evaluadores, los comisionados y los destinatarios de la evaluación, en el marco del cual considerar cómo deberían aplicarse e interpretarse los criterios. Un proceso como el descrito puede fomentar el diseño de evaluaciones creíbles y oportunas que respondan mejor a las necesidades de los usuarios.

Los requisitos metodológicos establecidos por una institución también pueden influir en la forma en que se aplican los criterios. Los evaluadores deben dejar constancia de los requisitos y las directrices particulares, ya sean de su propia organización o de la institución que ordena la evaluación. Otras fuentes pertinentes, como la guía del Grupo de Evaluación de las Naciones Unidas (UNEG) o los estándares de buenas prácticas del Grupo de Cooperación en Materia de Evaluación (ECG) y la guía de evaluación de la acción humanitaria de la Red para un Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria (ALNAP) son igualmente de gran utilidad cuando son aplicables.

Cómo responder a la Agenda 2030 y a los Objetivos de Desarrollo Sostenible

Además de adaptar los criterios a las instituciones en las que se están utilizando, la forma en que se entiendan y apliquen los criterios reflejará el contexto de políticas más amplio, influyendo en la forma en que los gestores de las evaluaciones, los evaluadores y las partes interesadas utilizan los criterios. Para el próximo decenio, especialmente para la comunidad evaluadora que trabaja en la cooperación internacional para el desarrollo, los Objetivos de Desarrollo Sostenible (ODS) y la Agenda 2030 serán el marco de políticas y el conjunto de objetivos globales más importante.

Entre los elementos clave de la Agenda 2030 se encuentran los siguientes:

- acceso universal a los beneficios del desarrollo;
- inclusión, en especial para las personas en mayor riesgo de quedarse atrás;
- derechos humanos, igualdad de género y otras consideraciones de equidad;
- sostenibilidad ambiental, cambio climático y gestión de los recursos naturales;
- complejidad del contexto y de las intervenciones de desarrollo;
- sinergias entre los actores involucrados en el proceso de desarrollo.

Este marco influye tanto en las intervenciones como en su evaluación, entre otras cosas en cómo se interpretan los criterios, así como en el proceso de evaluación propiamente dicho (por ejemplo, en quiénes participan en la aplicación de los criterios y en la identificación de las preguntas prioritarias). El Recuadro 3.3 contiene pautas sobre el uso de la Agenda 2030 para fundamentar las agendas de evaluación nacionales. El Recuadro 3.4 presenta un ejemplo de cómo se procedió en el sistema alemán de evaluación del desarrollo (BMZ, 2020^[71]). Varios gobiernos nacionales han emprendido esfuerzos similares (entre ellos Costa Rica, Nigeria y Finlandia), al igual que algunas organizaciones no gubernamentales (ONG), de modo que han proporcionado lecciones de gran valor a los evaluadores y los gestores de evaluaciones (D'Errico, Geoghe y Piergallini, 2020^[9]).

El texto revisado de las definiciones de los criterios refleja también estos elementos en varios sentidos. Por ejemplo, pone un énfasis especial en el contexto y en las perspectivas y las prioridades de la población beneficiaria al evaluar la pertinencia, la eficacia y el impacto; tiene en cuenta la equidad de los resultados al examinar la eficacia y el impacto; y adopta un enfoque integral para examinar la coherencia. Los principios rectores reflejan también la Agenda 2030 al fomentar una perspectiva integral.

Recuadro 3.3. Uso de los ODS para informar las agendas nacionales de evaluación

El documento *Evaluation to connect national priorities with the SDGs. A guide for evaluation commissioners and managers* propone cinco consideraciones para ayudar a los países a desarrollar sus agendas de evaluación y mejorar el valor de la evaluación. Exhorta a los evaluadores a aplicar el “pensamiento en sistemas complejos” y a:

- pensar más allá de las políticas, los programas y los proyectos individuales;
- examinar las fuerzas macro que influyen en los logros o los fracasos;
- desarrollar una comprensión matizada de lo que significan los logros;
- reconocer la importancia de la cultura;
- adoptar el pensamiento evaluativo y la gestión adaptable.

Fuente: Ofir *et al.* (2016^[8]), *Briefing: Five considerations for national evaluation agendas informed by the SDGs*, <https://pubs.iied.org/sites/default/files/pdfs/migrate/17374IIED.pdf>

Recuadro 3.4. Mapeo de los criterios en función de la Agenda 2030 y los ODS

Las preguntas que presentamos a continuación han sido elaboradas por el Ministerio Federal de Cooperación Económica y Desarrollo de Alemania (BMZ) con el apoyo del Instituto Alemán de Evaluación del Desarrollo (DEval), con el objeto de ayudar a los evaluadores a valorar las contribuciones generales a la Agenda 2030 y a los ODS. Cada pregunta es una respuesta a los principios de los ODS y se relaciona también con los criterios.

Universalidad, responsabilidad compartida y rendición de cuentas

- ¿En qué medida contribuye la intervención al logro de los ODS? (Ver el criterio de impacto).
- ¿En qué medida ha sido diseñada la intervención para aprovechar los sistemas y las estructuras existentes (de los asociados, de otros donantes o de las organizaciones internacionales) para llevar a cabo sus actividades y en qué medida se utilizan estos? (Ver el criterio de coherencia).
- ¿Se aplica la división del trabajo con otros donantes y asociados para el desarrollo en lo que respecta a la ejecución de la intervención? De ser así, ¿en qué medida? (Ver el criterio de coherencia).
- ¿En qué medida se utilizan sistemas comunes para el seguimiento, el aprendizaje y la rendición de cuentas? (Ver el criterio de coherencia).

Interacción con el desarrollo económico, ambiental y social

- ¿La intervención utiliza un enfoque integral de desarrollo sostenible (social, ambiental y económico)? (Ver el criterio de pertinencia).
- ¿En qué medida se generaron interacciones (intencionales o no intencionales, positivas o negativas) entre los efectos directos sociales, económicos y ambientales, y cuál fue el impacto general de la intervención? (Ver los criterios de eficacia e impacto).
- ¿Qué contribución realiza la intervención a la promoción de interacciones (intencionales o no intencionales, positivas o negativas) entre los efectos directos sociales, económicos y

ambientales, y cuál fue el impacto general de la intervención? (Ver los criterios de eficacia e impacto).

Inclusión

- ¿En qué medida es consistente la intervención con las normas y los estándares internacionales que regulan la participación y la promoción de los grupos en especial desventaja y vulnerabilidad? (Ver el criterio de coherencia).
- ¿En qué medida se generaron cambios de desarrollo importantes (intencionales o no intencionales, positivos o negativos) a nivel de los grupos en especial desventaja y vulnerabilidad (posible diferenciación por edad, ingresos, género, origen étnico, etc.)? (Ver el criterio de impacto).
- ¿Qué contribución realizó la intervención a los impactos de desarrollo centrales (intencionales o no intencionales, positivos o negativos) a nivel de los grupos en especial desventaja y vulnerabilidad (posible diferenciación por edad, ingresos, género, origen étnico, etc.)? (Ver el criterio de impacto).
- ¿En qué medida contribuyó la intervención a fortalecer la resiliencia de los grupos en especial desventaja y vulnerabilidad (posible diferenciación por edad, ingresos, género, origen étnico, etc.)? (Ver el criterio de sostenibilidad).

Fuente: BMZ (2020^[7]), *Evaluation Criteria for German Bilateral Development Co-operation*

Cómo aplicar un lente de género a los criterios

Los evaluadores deben trabajar de formas que tengan en cuenta minuciosamente las experiencias y los impactos diferenciados por género, así como la forma en que estos interactúan con otras modalidades de discriminación en un contexto específico (por ejemplo, por edad, raza y origen étnico, o nivel socioeconómico). Independientemente de la intervención, los evaluadores deben considerar cómo las dinámicas de poder de género convergen e interactúan con otras formas de discriminación, afectando así la ejecución y los resultados de la intervención. Esto puede incluir una exploración de cómo influye el contexto sociocultural y político-económico de las intervenciones en la ejecución y el logro de los objetivos.

La aplicación de un lente de género puede generar evidencia para incentivar el aprendizaje y la rendición de cuentas, a la vez que respalda el logro de los objetivos de igualdad de género. Algunas medidas prácticas para aplicar un lente de género a los criterios de evaluación son las siguientes:

- los evaluadores, gestores y comisionados deben trabajar de formas que sean inclusivas y que conduzcan a una participación apropiada en la toma de decisiones, la recolección y el análisis de datos, y el intercambio de los hallazgos;
- considerar el grado en que las barreras de género interactúan con otras barreras sociales, poniendo en riesgo así la igualdad de oportunidades de la intervención;
- considerar cómo interactúa una intervención con el entorno legislativo, económico, político, religioso y sociocultural, a efectos de interpretar mejor los diferentes impactos y experiencias de las partes interesadas;
- considerar las definiciones de masculinidad y feminidad socialmente construidas, así como cualquier cambio en las dinámicas y roles de género;
- al seleccionar a los evaluadores, analizar sus habilidades en enfoques de evaluación con perspectiva de género y sus experiencias de trabajo en diferentes contextos.

El siguiente cuadro ha sido elaborado para ayudar a los evaluadores a considerar cómo aplicar un lente de género a los criterios:

Cuadro 3.1. Cómo aplicar un lente de género a los criterios

Criterios	Preguntas rectoras para la aplicación de un lente de género
Pertinencia	<p>¿La intervención se diseñó de formas que tienen en cuenta las necesidades y las prioridades de todos los géneros? De ser así, ¿cómo?</p> <p>¿En qué medida el diseño de la intervención refleja los derechos de las personas de todos los géneros e incluye la retroalimentación de una diversa gama de partes interesadas locales, entre ellas los grupos marginados?</p> <p>¿Responde la intervención a las necesidades prácticas y estratégicas de todos los géneros?</p>
Coherencia	<p>¿En qué medida son coherentes el diseño, la ejecución y los resultados de la intervención con las leyes y los compromisos internacionales de derechos e igualdad de género, entre ellos la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW), la Declaración y la Plataforma de Acción de Beijing, el Programa de Acción de la Conferencia Internacional sobre la Población y el Desarrollo, y la Agenda 2030?</p> <p>¿En qué medida apoya la intervención las leyes e iniciativas nacionales creadas para mejorar la igualdad de género y los derechos humanos? ¿Qué lecciones se pueden extraer?</p>
Eficacia	<p>¿La intervención logró sus objetivos y resultados previstos de formas que contribuyen a la igualdad de género? De ser así, ¿cómo?</p> <p>¿La intervención generó resultados diferenciados para diferentes grupos poblacionales? De ser así, ¿cómo y por qué?</p> <p>¿Hubo que usar enfoques diferentes para llegar a las personas de géneros diferentes? ¿Se realizó el seguimiento y el análisis de los efectos diferenciados en grado suficiente? ¿Se adaptó la intervención para responder a cualquier inquietud y maximizar la eficacia?</p> <p>¿La teoría del cambio y el marco de resultados se fundamentaron en el análisis de la igualdad de género, el análisis de la economía política y los derechos humanos? De ser así, ¿en qué medida?</p> <p>¿En qué medida y por qué es distinta la eficacia para las personas de géneros diferentes?</p>
Eficiencia	<p>¿Se asignaron recursos diferentes de formas que tuvieron en cuenta la igualdad de género? De ser así, ¿cómo se asignaron? ¿Fue apropiada la asignación de recursos diferenciada?</p> <p>¿Los costos de inversión por persona beneficiada satisfacen las necesidades diferenciadas de las personas de diferentes géneros?</p>
Impacto	<p>¿La intervención tuvo los mismos impactos para los diferentes géneros o hubo diferencias relacionadas con los géneros a nivel de participación, experiencia e impactos? De ser así, ¿por qué ocurrieron estos impactos diferenciados?</p> <p>¿En qué medida convergieron los impactos de género con otras barreras sociales como la raza o el origen étnico, la discapacidad, la edad y la orientación sexual para contribuir a las experiencias y los efectos directos diferenciados?</p> <p>¿Cómo impactan en los efectos directos las normas y las barreras de género del entorno político, económico, religioso, legislativo y sociocultural en general?</p> <p>¿En qué medida han contribuido los impactos a la igualdad en las relaciones de poder entre personas de diferentes géneros y a modificaciones en las normas y los sistemas sociales?</p>
Sostenibilidad	<p>¿La intervención ha contribuido a fortalecer la igualdad de género en los sistemas jurídicos, políticos, económicos y sociales más amplios? De ser así, ¿cómo y en qué medida? ¿Terminó este hecho produciendo cambios duraderos en las normas sociales que eran perjudiciales para las personas de un género en particular o de todos los géneros? Si no se logró ese resultado, ¿por qué no?</p> <p>¿Se mantendrán los logros de igualdad de género una vez concluida la intervención? ¿Han contribuido los procesos al sostenimiento de estos beneficios? ¿Se han creado mecanismos para apoyar el logro de la igualdad de género a más largo plazo?</p>

Cómo usar otros criterios

Los seis criterios han sido diseñados para formar un conjunto que refleje plenamente todos los conceptos importantes que se deben cubrir en las evaluaciones. Si se aplican con minuciosidad y de forma pertinente para el contexto, serán adecuados para las evaluaciones de todos los ámbitos del desarrollo sostenible y la acción humanitaria.

Sin embargo, en algunos contextos se utilizan otros criterios. Por ejemplo, muchas instituciones establecen en sus políticas de evaluación la obligatoriedad de analizar un área de focalización en particular. En 2020, una evaluación de los programas de salud de la cooperación italiana en Bolivia se guio por nueve criterios: pertinencia, eficacia, eficiencia, impacto, sostenibilidad, coherencia, valor agregado de la cooperación italiana, visibilidad de la cooperación italiana y apropiación (Eurecna Spa, 2020^[9]). Otro ejemplo es el uso de los criterios en situaciones humanitarias, donde otros criterios, como la idoneidad, la cobertura y la conectividad, resultan altamente pertinentes³. La adicionalidad es un criterio que también se utiliza en ocasiones, a menudo en la evaluación de iniciativas de financiamiento mixto, financiamiento no soberano y financiamiento climático. Se emplean distintas definiciones de adicionalidad, entre ellas diferentes tipos de adicionalidad financiera y no financiera⁴. Dependiendo de la definición utilizada, la adicionalidad puede examinarse bajo el criterio de pertinencia, la eficacia o el impacto. Otros utilizan la adicionalidad como un criterio transversal distintivo.

Los usuarios deberán ser cautelosos a la hora de evaluar la conveniencia de añadir o no criterios adicionales, ya que hacerlo puede provocar confusión y producir evaluaciones demasiado amplias (generando un análisis menos útil). Es conveniente mantener un número limitado de criterios, a fin de asegurar la profundidad del análisis y la claridad conceptual necesarias, un argumento que se esgrimió repetidamente durante el proceso de consulta para la actualización de las definiciones entre 2017 y 2019.

Si se utilizan otros criterios, es importante definirlos. Explicar las razones por las que se añaden puede resultar útil para que otras personas determinen en qué medida los criterios adicionales son compatibles con los seis criterios aquí descritos. Para apoyar el aprendizaje en todas las evaluaciones, es importante que se valoren los mismos conceptos o elementos bajo los mismos criterios.

Independientemente de qué criterios se utilicen, deben aplicarse las directrices y los principios básicos señalados aquí.

Referencias

- ADE (2019), *Evaluación conjunta del modelo de soluciones integrales en Kalobeyei y alrededores, Turkana, Kenia*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d. [5]
- Bamberger, M., J. Vaessen and E. Raimondo (2015), “Dealing With Complexity in Development Evaluation - A Practical Approach”, https://www.betterevaluation.org/en/resources/dealing_with_complexity_in_development_eval. [23]
- Beck, T. (2016), *Evaluating Humanitarian Action using the OECD-DAC Criteria*, <https://www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria>. [1]
- BMZ (2020), *Evaluation Criteria for German Bilateral Development Co-operation*. [7]

- Bryld (2019), *Evaluación del apoyo de Sida a la construcción de la paz en situaciones de conflicto y posconflicto: informe de país de Somalia*, [6]
https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5
- D'Errico, S. (n.d.), "Evaluation to connect national priorities with the SDGs | Publications Library", [20]
<https://pubs.iied.org/17739IIED>.
- Danida (2019), *¿De donantes a socios? Evaluación del apoyo noruego al fortalecimiento de la sociedad civil en países en desarrollo a través de organizaciones de la sociedad civil noruegas*, [4]
<http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf>.
- Davis, R. (2013), "Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations", Vol. No. 20, [22]
<https://www.gov.uk/government/publications/planning-evaluability-assessments>.
- Eurecna Spa (2020), *Bolivia - Evaluation of Health Initiatives (2009-2020)*, [9]
http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf.
- Helle, E. (2018), *¿De donantes a socios? Evaluación del apoyo noruego al fortalecimiento de la sociedad civil en países en desarrollo a través de organizaciones de la sociedad civil noruegas*, [3]
https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf.
- IDD and Associates (2006), *A Joint Evaluation of General Budget Support Evaluation of General Budget Support: Synthesis Report*, [10]
<https://www.oecd.org/development/evaluation/dcdndep/37426676.pdf> (accessed on 11 January 2021).
- Molund, S. and G. Schill (2004), *Looking Back, Moving Forward: Sida Evaluation Manual*, [18]
<https://www.oecd.org/derec/sweden/35141712.pdf>.
- OCDE (2019), *Mejores criterios para una mejor evaluación: Definiciones revisadas de los criterios de evaluación y principios para su utilización, Red de Evaluación del Desarrollo del CAD*, [13]
<https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>.
- OCDE (2010), *Estándares de Calidad para la Evaluación del Desarrollo, Serie: Directrices y Referencias del CAD*, OECD Publishing, Paris, [15]
<https://doi.org/10.1787/9789264094949-es>.
- OCDE (2002), *Evaluation and Aid Effectiveness No. 6, Glosario de los principales términos sobre evaluación y gestión basada en resultados*, OECD Publishing, París, [12]
<https://doi.org/10.1787/9789264034921-en-fr>.
- OCDE (1991), *Principles for Evaluation of Development Assistance*, Comité de Ayuda al Desarrollo, OECD Publishing, París, [11]
<https://www.oecd.org/development/evaluation/2755284.pdf> (accessed on 11 January 2021).
- OECD (2019), *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*, [17]
<https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>.

- OECD (2010), *Directrices y Referencias del CAD (series) Estándares de Calidad para la Evaluación del Desarrollo*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264094949-es>. [16]
- OECD (2002), *Evaluation and Aid Effectiveness No. 6 - Glossary of Key Terms in Evaluation and Results Based Management (in English, French and Spanish)*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [21]
- OECD (1991), *Principles for Evaluation of Development Assistance*, <https://www.oecd.org/development/evaluation/2755284.pdf>. [14]
- Ofir, E. (2016), "Resumen informativo: cinco consideraciones para agendas nacionales de evaluación informadas por los ODS", <https://pubs.iied.org/sites/default/files/pdfs/migrate/17374IIED.pdf>. [8]
- Schwedersky, T., L. Ahrens and H. Steckhan (2017), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, German Institute for Development Evaluation (DEval), Bonn, https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2017_API/DEval_Bericht_APIInklusion_2017_EN.pdf. [2]
- WFP (2016), *Technical Note: Evaluation Methodology, DEQAS, World Food Programme*, <https://docs.wfp.org/api/documents/704ec01f137d43378a445c7e52dcf324/download/>. [19]

Notas

¹ Evaluabilidad es el grado en que una actividad o un programa es apto para ser evaluado de manera confiable y creíble. El análisis de la evaluabilidad requiere de un examen temprano de la actividad propuesta, para cerciorarse de que sus objetivos específicos estén correctamente definidos y sus resultados sean verificables (OCDE, 2002^[14]). Ver también (Davis, 2013^[22]).

² El proceso de formular las preguntas de evaluación e involucrar a las partes interesadas, al mismo tiempo que se tiene en cuenta la complejidad de la intervención, se analiza en mayor detalle, por ejemplo, en Bamberger, Vaessen y Raimondo (2015^[12]).

³ La Red para un Aprendizaje Activo sobre Rendición de Cuentas y Resultados de la Acción Humanitaria (ALNAP) se encuentra actualmente en proceso de actualizar su guía de 2006 sobre el uso de los criterios en situaciones humanitarias, como complemento de la *Guía de evaluación de la acción humanitaria* de ALNAP, de amplio alcance.

⁴ El Grupo de Trabajo de EvalNet sobre evaluación del *blended finance* o financiamiento mixto ha apoyado una investigación sobre las definiciones de adicionalidad y otros conceptos relacionados. Los hallazgos se publicarán a principios de 2021.

4

Cómo entender los seis criterios: definiciones, elementos de análisis y principales desafíos

El presente capítulo profundiza en cada uno de los seis criterios. Explora lo que su definición significa en la práctica, diferentes elementos de análisis y cómo se pueden aplicar los criterios de forma que reflejen la inclusión y la igualdad como un elemento primordial. Para cada criterio, un cuadro muestra los principales desafíos para su aplicación, con recomendaciones prácticas para hacerles frente, dirigidas a los evaluadores y los gestores de evaluaciones. El capítulo incluye una serie de ejemplos que ilustran la aplicación práctica de los criterios e invitan a una reflexión crítica en torno a las dimensiones de cada definición.

El presente capítulo introduce los seis criterios (**Gráfico 4.1**) presentándolos en el orden más lógico: primero pertinencia y coherencia, luego eficacia y eficiencia, para finalmente pasar a impacto y sostenibilidad. Incluimos una definición de cada criterio y describimos su importancia. Después, explicamos la definición en mayor detalle, examinando sus elementos de análisis (los conceptos más importantes contenidos en cada definición). Dichos elementos no constituyen subcriterios, sino que ilustran las diferentes formas en que los criterios se pueden aplicar, dependiendo del contexto y del propósito de la evaluación. Exploramos asimismo las conexiones con otros criterios.

El capítulo incluye adicionalmente reflexiones preliminares en torno a aspectos clave relacionados con la inclusión y el principio de “no dejar a nadie atrás”, en consonancia con el mandato de la Agenda 2030. Actualmente se está trabajando en la aplicación de un lente de equidad a los criterios y enfoques de evaluación, entre otras cosas produciendo una guía sobre igualdad de género, empoderamiento de la mujer y derechos humanos.

Para cada criterio, un cuadro muestra los desafíos comunes, así como algunas ideas para que los evaluadores y los gestores de evaluaciones les hagan frente. Estos cuadros se actualizarán con el tiempo, a medida que se recojan las experiencias con las nuevas definiciones.

Finalmente, se ofrecen ejemplos de la vida real para ilustrar las formas de interpretar cada criterio.

Gráfico 4.1. Los seis criterios y sus preguntas relacionadas



Pertinencia

Definición de pertinencia:

Pertinencia: ¿La intervención está haciendo lo correcto?

Medida en que los objetivos y el diseño de una intervención responden a las necesidades, políticas y prioridades de los beneficiarios¹, los asociados o las instituciones, internacionales y nacionales, y siguen haciéndolo aun cuando las circunstancias cambian.

Nota: “Responden a” se refiere a que los objetivos y el diseño de la intervención son sensibles a las condiciones económicas, ambientales, de equidad, sociales, de economía política y de capacidad en las que se desarrolla. “Asociado o institución” engloba a gobiernos (nacionales, regionales, locales), organizaciones de la sociedad civil, entidades privadas y organismos internacionales que participan en el financiamiento, la ejecución o la supervisión de la intervención. Para determinar la pertinencia, es necesario analizar las diferencias y compensaciones entre distintas prioridades o necesidades. Exige analizar todo cambio que se produzca en el contexto para establecer la medida en que puede adaptarse (o se ha adaptado) la intervención para seguir siendo pertinente.

¿Qué es la pertinencia y por qué es importante?

Evaluar la pertinencia ayuda a los usuarios a establecer si una intervención está haciendo lo correcto. Permite a los evaluadores valorar en qué medida los objetivos y la ejecución de la intervención están claramente alineados con las necesidades de los beneficiarios y las partes interesadas, así como con las prioridades que sostienen la intervención. Establece si las partes interesadas seleccionadas consideran útil y valiosa la intervención.

La pertinencia es una consideración relevante en todas las etapas del ciclo de un programa o política, desde el diseño hasta la ejecución. La pertinencia puede evaluarse asimismo en función de objetivos universales como los Objetivos de Desarrollo Sostenible (ODS). Puede analizarse en función de cuatro posibles elementos de análisis: pertinencia para las necesidades de los beneficiarios y de las partes interesadas, pertinencia para el contexto, pertinencia de la calidad y el diseño, y pertinencia a lo largo del tiempo, las cuales se analizan en mayor detalle en la sección sobre elementos de análisis. Estos elementos deben utilizarse según lo requiera el propósito de la evaluación y no son exhaustivos.

La evaluación de la pertinencia debe empezar por determinar si los objetivos de la intervención se encuentran definidos de manera adecuada y son realistas y viables, y si los resultados son verificables y están alineados con los estándares internacionales que rigen actualmente las intervenciones de desarrollo. Esto debería ser compatible con el concepto de evaluabilidad que se detalla en los Estándares de Calidad para la Evaluación del Desarrollo de la OCDE (OCDE, 2010^[1]). Los enunciados acerca de los resultados o los objetivos pueden estar formulados en términos difusos o vagos, ser difíciles de medir o referirse a actividades o insumos. En algunos casos, la teoría del cambio se debe mejorar o reconstruir para que el evaluador identifique claramente los objetivos. Los evaluadores deben tener el cuidado de evaluar en función de objetivos realistas y de buena calidad. Los indicadores para medir el logro de los objetivos también se deben validar en función de criterios ampliamente aceptados como los indicadores SMART (específicos, mensurables, alcanzables, pertinentes y oportunos) (IDD and Associates, 2006^[10]). Las evaluaciones que examinan la pertinencia deben considerar si se han operativizado los objetivos apropiados y cómo de una forma que refleje unas buenas prácticas. Además, deben reflexionar acerca de la capacidad y las aptitudes organizacionales de los asociados en la ejecución y sus respuestas a los cambios en el contexto.

Los evaluadores deben identificar claramente asimismo a las partes interesadas cuyas necesidades y prioridades se deben tener en cuenta en la evaluación de la pertinencia. Esto incluye a la población beneficiaria, así como a las instituciones y los asociados de financiamiento, supervisión o ejecución. Se debe poner un énfasis especial en los beneficiarios. La apropiación de una intervención es importante, y los beneficiarios figuran en primer lugar como las principales partes interesadas en la definición de las prioridades y necesidades. Dependiendo de la intervención, puede ser pertinente asimismo considerar las necesidades nacionales y subnacionales (donde proceda), las estrategias locales para suplir las necesidades y el grado en que la intervención está alineada con dichas necesidades. Las necesidades institucionales pueden incluir las necesidades de los donantes, entre otras, lo cual significa que la pertinencia se puede evaluar en diferentes contextos de políticas, incluso aquellos donde no hay un donante claro, sino más bien asociados en una intervención (por ejemplo, en el caso de la política comercial).

Cómo entender la pertinencia: elementos de análisis

La definición de la pertinencia engloba cuatro grandes dimensiones: la respuesta a las necesidades, políticas y prioridades; la sensibilidad y la capacidad de reacción al contexto; la calidad del diseño; y la capacidad de respuesta a lo largo del tiempo.

Respuesta a las necesidades, políticas y prioridades

Quizás el elemento más importante para el análisis de la pertinencia es la valoración del grado en que una intervención responde a las necesidades y las prioridades de la población beneficiaria. Un análisis de esta naturaleza arrojará luz sobre los temas en los que incide una intervención y por qué. Los beneficiarios son las partes interesadas centrales de una intervención y se les debe tener en cuenta en todo momento. Los beneficiarios no necesariamente son personas que reciben servicios directos. Dependiendo del tipo de intervención, los beneficiarios pueden ubicarse a un nivel (mucho) más macro de la cadena de resultados. Por ejemplo, el objetivo de una intervención puede ser fortalecer las capacidades de la oficina nacional de auditoría. Estas capacidades mejoradas redundarán en favor de la gestión de las finanzas públicas y, en última instancia, contribuirán al logro de los Objetivos de Desarrollo Sostenible en las áreas de, por ejemplo, la educación y la salud. No obstante, los beneficiarios primarios, para los fines de evaluar el apoyo al desarrollo de capacidades, serían los miembros del personal de la oficina. Definir con toda claridad a los beneficiarios (primarios y secundarios) es el primer paso necesario para evaluar la pertinencia.

El análisis de las necesidades de los beneficiarios y si se satisfacen o no arroja luz no solamente sobre la capacidad de respuesta de la intervención, sino también sobre la apropiación y la participación en relación con el diseño y la ejecución de la intervención (lo cual puede afectar a otros criterios). Ayuda a entender quiénes participaron en el diseño y quiénes no y, a su vez, cómo afecta este hecho al diseño y la ejecución de la intervención.

Este criterio significa que los esfuerzos deberían centrarse en las áreas de mayor necesidad o, para usar los términos de la Agenda 2030, en llegar primero a los más rezagados. De hecho, la pertinencia es especialmente valiosa para dilucidar quién está involucrado en una intervención y quién es beneficiario de ella. La pertinencia ofrece a los evaluadores la oportunidad de considerar si los grupos marginados están incorporados en las prioridades tanto de la política como de la intervención y en qué medida. Incluso cuando una intervención se encuentra en perfecta sintonía con la política oficial, puede estar desconectada de las prioridades reales de los participantes, que pueden no haber participado en la definición de las prioridades y los planes oficiales.

Una evaluación pertinente debe tener en cuenta asimismo la medida en que la intervención responde a las prioridades de las instituciones o los asociados involucrados. Estos pueden ser: gobiernos (nacionales,

regionales o locales), organizaciones de la sociedad civil, entidades privadas y organismos internacionales implicados en el financiamiento, la ejecución o la supervisión de la intervención. La pertinencia examinará la alineación de la intervención con las estrategias y las políticas de esas instituciones.

A fin de valorar la pertinencia de una intervención para las necesidades, políticas y prioridades globales, una evaluación debería revisar la contribución de la intervención a los objetivos mundiales más amplios (es decir, la importancia relativa de la intervención frente al esfuerzo general). Esto a menudo supondrá que los evaluadores comparen los (posibles) resultados en el contexto o país en cuestión con alternativas en otros lugares. Estas preguntas sobre la pertinencia global no siempre se contemplan durante el diseño de una intervención. Los evaluadores pueden, por lo tanto, aportar un análisis valioso de estas cuestiones, ayudando con ello a las partes interesadas a entender la trascendencia estratégica de la intervención más allá de su contexto local.

La definición de la pertinencia también llama a los evaluadores a examinar las posibles tensiones o compensaciones al considerar de quién son las necesidades y las prioridades que se satisfacen mediante la intervención. Es posible que varias de las perspectivas de los participantes y otras partes interesadas se encuentren desalineadas entre sí, de modo tal que la evaluación tendrá que esclarecer estas diferencias y explorar las implicaciones de las decisiones tomadas. Para poner un ejemplo, las intervenciones orientadas a erradicar una enfermedad (por ejemplo, la poliomielitis) en todos los países examinarían la carga relativa de la enfermedad para determinar las prioridades de acción mundiales en un país o una región en particular. En ese contexto, puede haber situaciones en que la prioridad global de una intervención de esa naturaleza (inmunización contra la polio en la última región del mundo donde existe la transmisión comunitaria) esté reñida con las prioridades locales (dado que la población beneficiaria prioriza los temas de agua y saneamiento, por ejemplo). Para los evaluadores, puede ser útil desentrañar dichas tensiones realizando un cuidadoso análisis de la pertinencia.

Sensibilidad y capacidad de reacción al contexto

Las necesidades de la población beneficiaria y otras partes interesadas clave no se pueden entender de manera aislada y se ven determinadas por el contexto. El contexto es multifacético e incluye factores económicos, ambientales, de equidad, sociales, culturales, de economía política y de capacidad. Los evaluadores deben establecer qué factores contextuales son más pertinentes para una intervención.

La pertinencia contextual puede analizarse tanto en el diseño como en la ejecución de una intervención. El análisis del contexto dependerá asimismo de si la evaluación se realiza *ex ante* o *ex post*. Por ejemplo, los evaluadores pueden plantear preguntas sobre cómo se entendió y se dio cuenta del contexto cuando se diseñó la intervención. En las evaluaciones *ex post*, los evaluadores deben preguntarse si el contexto cambió entre la fase inicial y el término de la intervención. Las evaluaciones *ex post* dispondrán de más contexto y deben tratar de incorporar este hecho en su análisis. Esto complementa el elemento de la temporalidad para el análisis de la pertinencia, al considerar las posibles fluctuaciones en la pertinencia de una intervención cuando cambian las circunstancias.

Adicionalmente se puede tener en cuenta el contexto histórico. Por ejemplo, ¿hubo intervenciones similares en el pasado? ¿Existen tensiones, leyes o políticas históricas que pudieran haber impactado en la comprensión de las necesidades y la determinación de los objetivos? El contexto histórico puede incluir asimismo los supuestos formulados en el pasado acerca de la pertinencia de una intervención, y habría que corroborar si dichos supuestos se mantienen en el contexto actual. En caso de que existan evaluaciones previas, dichos supuestos pueden ser útiles para reconstruir el contexto histórico y establecer si las intervenciones aprovechan o no las lecciones aprendidas de ejercicios evaluativos previos.

Calidad del diseño

La “calidad del diseño” es la medida en que la intervención fue diseñada para responder a las prioridades y las necesidades pertinentes, y si se han especificado claramente los objetivos. Además, este elemento de análisis evalúa si las prioridades y las necesidades de las partes interesadas se mencionan en los objetivos de la intervención, su teoría del cambio subyacente, su teoría de acción o su *modus operandi*. Esto permite a los evaluadores determinar cuáles son las deficiencias del diseño del programa que pueden haber socavado la pertinencia general de la intervención. Este elemento de análisis influye asimismo en la evaluabilidad de la intervención en general, al enfocarse en la calidad del diseño de la intervención desde el inicio. Asimismo, ofrece información sobre la idoneidad de la intervención para la institución que la ejecuta. Por ejemplo, los evaluadores pueden intentar esclarecer si la intervención se diseñó teniendo en cuenta su viabilidad técnica, organizacional y financiera.

Adaptación a lo largo del tiempo

Los evaluadores deben considerar la medida en que las intervenciones pueden evolucionar a lo largo del tiempo. El estallido de un conflicto o los cambios en el contexto económico y de políticas, por ejemplo, afectarán perceptiblemente la ejecución de una intervención. La pertinencia en relación con las consideraciones de tiempo debe incluir un análisis de la adaptabilidad de la gestión. Los evaluadores no deben limitarse a considerar la pertinencia al inicio y al final de un programa; deben examinar también cómo ha respondido este a los cambios sucedidos a lo largo de su vida útil. Esto les permite examinar las fluctuaciones en la pertinencia de una intervención a medida que cambian las condiciones para su ejecución. Por ejemplo, podría incluir un análisis de la medida en que las adaptaciones fueron apropiadas para seguir satisfaciendo las necesidades y prioridades más importantes, y si las adaptaciones afectaron la calidad del diseño a lo largo del tiempo.

Nuevamente, aquí se debe tener en cuenta la adaptación continua del programa a los contextos externos y los cambios internos (por ejemplo, cuando un cambio en el financiamiento requiere de un cambio en la programación). Además, pueden tenerse en cuenta los riesgos y las oportunidades, por ejemplo, la medida en que el programa mitigó los riesgos que podrían socavar su propósito o en que se adaptó para aprovechar una buena oportunidad o para mejorar la atención de las necesidades. La adaptación puede llevar a compensaciones en torno a las necesidades de quién priorizar, dando lugar a preguntas de rendición de cuentas. Este aspecto se debe explorar en profundidad para determinar cómo puede haber alterado o no la pertinencia de un programa.

Conexiones con otros criterios

- Dado que los criterios están interconectados, la pertinencia se puede vincular a otros criterios en una evaluación. La pertinencia a menudo se considera como un requisito para cumplir los otros criterios.
- La evaluación de la pertinencia sienta las bases para determinar si las necesidades se cubren como parte de la eficacia y el impacto. De hecho, la pertinencia como criterio es un requisito para la eficacia, ya que la identificación de las necesidades y los objetivos debe estar claramente enunciada para poder evaluar la eficacia.
- La pertinencia complementa la coherencia. Ambas requieren de un análisis contextual: en el caso de la pertinencia, para entender la alineación de la intervención con las prioridades y las necesidades de las principales partes interesadas y, en el caso de la coherencia, para entender los vínculos con otras intervenciones. La pertinencia examina la medida en que una intervención responde al contexto. La coherencia se aleja, toma distancia y examina otras intervenciones dentro del mismo contexto y cómo interactúan con la intervención objeto de evaluación. En conjunto, la pertinencia y la coherencia pueden proyectar una visión más clara de la medida en

que la intervención afecta al contexto en que se lleva a cabo (al tiempo que se ve afectada por él).

- El análisis de la pertinencia está relacionado asimismo con el criterio de impacto, que examina la trascendencia última de una intervención, incluyendo su valor para las personas afectadas. Los evaluadores deben tomarse el tiempo necesario para examinar las necesidades, las prioridades y las políticas de todos los actores (incluidas las posibles tensiones entre ellos) para poder evaluar en grado suficiente la pertinencia general de la intervención y profundizar en su trascendencia durante el análisis del impacto.
- Muchos de los elementos de la pertinencia son factores críticos para la eficiencia y la sostenibilidad: una intervención pertinente probablemente concitará un mayor apoyo de las partes interesadas, lo que a su vez puede influir en la puntualidad de la entrega y el uso de los recursos, así como en el nivel de apropiación (y por lo tanto en la sostenibilidad) de los beneficios resultantes.

Cómo integrar la inclusión

Entender las dinámicas de poder desde una perspectiva de género y reflexionar sobre el compromiso de los ODS de “no dejar a nadie atrás” es fundamental para entender la pertinencia. Las dinámicas de poder desde el punto de vista de género y la marginación de ciertos grupos (por ejemplo, raciales o étnicos) son consideraciones centrales para entender la pertinencia en un contexto determinado.

Entender quiénes fueron involucrados en el diseño de la intervención y de qué manera, poniendo especial énfasis en las dinámicas de poder y los grupos marginados, ayudará a los evaluadores a entender la pertinencia de la intervención tal y como fue diseñada, así como el grado en que la intervención demostró capacidad de respuesta frente a los cambios en las necesidades a lo largo del tiempo.

La definición de pertinencia consignada arriba subraya la importancia de considerar las compensaciones entre las necesidades y prioridades de los diferentes grupos, incluyendo un mayor énfasis en la equidad y las dinámicas de poder entre las personas afectadas por la intervención, ya sea directa o indirectamente.

Aquí existe una fuerte conexión con los derechos humanos y la igualdad, en especial cuando se utiliza un lente intersectorial para examinar cómo convergen las múltiples formas de identidad social y política, como el género, la discapacidad, el origen étnico, la sexualidad y la clase social, para crear discriminación y desigualdad. Al identificar las prioridades para el análisis de la pertinencia, es primordial tener en cuenta los grupos subrepresentados y marginados (que pueden tener restricciones de acceso a servicios o derechos) y cómo se han recogido (o no) sus necesidades y prioridades en los documentos y las políticas formales. Además, será importante tener en cuenta si la intervención incorpora los diferentes niveles de acceso, dadas las restricciones que enfrentan ciertos grupos.

Desafíos de evaluar la pertinencia y cómo enfrentarlos

El cuadro a continuación, pensado para los evaluadores y los gestores de evaluaciones, identifica varios desafíos comunes para la evaluación de la pertinencia —la gama de necesidades y prioridades que considerar, objetivos mal definidos y cambios en el contexto— y sugiere opciones para hacerles frente.

Cuadro 4.1. Desafíos de evaluar la pertinencia

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
Existen muchas partes interesadas nacionales e internacionales que considerar, que tienen	Un análisis sólido de la pertinencia debería considerar las compensaciones reales y potenciales en las respuestas a las	Definir de manera clara y sucinta a los beneficiarios de la intervención. Si la cadena de resultados es muy larga, enfocarse en los

múltiples prioridades y necesidades posiblemente contrapuestas.	necesidades. La pertinencia de los beneficiarios debe ser la consideración más importante que tener en cuenta al examinar las necesidades y las prioridades. La apropiación y la participación son también consideraciones que se recomienda tener en cuenta para manejar este desafío y cerciorarse de que se haya dado cuenta de las prioridades contrapuestas y, donde sea posible, que estas se hayan conciliado.	beneficiarios directos para el análisis de la pertinencia y dejar las preguntas sobre los beneficiarios últimos para que se analicen bajo los criterios de eficacia e impacto. Al reducir el ámbito para decidir de quién son las necesidades y prioridades más importantes, considerar el propósito de la evaluación y qué componente de la pertinencia resulta más útil para sustentar este argumento. Analizar cuidadosamente la utilidad de la evaluación en el momento del diseño.
Establecer si una intervención tiene objetivos claros y apropiados, y cómo fueron determinados.	Para hacer una evaluación sólida de la pertinencia, los objetivos específicos de la intervención constituyen un punto de partida esencial. Esto puede representar un desafío si no están claramente formulados. Los evaluadores deben revisar de manera crítica si los objetivos son lo suficientemente estimulantes y reflejan las necesidades tanto de los beneficiarios previstos como en términos de capacidades de las partes interesadas, y explorar cómo se identificaron y verificaron dichos objetivos. La idoneidad de los objetivos podría explorarse más en detalle bajo el criterio de impacto, de cara a establecer si los objetivos contribuyeron a generar cambios transformadores importantes. En esta instancia, los evaluadores deben interactuar estrechamente con las partes interesadas del programa, con el fin de esclarecer los objetivos del programa y cómo fueron definidos, y constatar si existe o existió un plan de contingencia. Deben evaluar si dicho plan se sustenta en un análisis sólido de las necesidades y el contexto, y en qué medida. Este será un importante punto de partida para entender la pertinencia.	Los gestores de la evaluación deben intentar establecer los objetivos específicos de la intervención como parte de la elaboración de los términos de referencia. Los términos de referencia deben permitir a los evaluadores reflexionar de manera crítica sobre si los objetivos fueron apropiados para las necesidades y el contexto externo, y si reflejan la retroalimentación de las partes interesadas.
La intervención carece de una teoría del cambio o de un modelo lógico claramente enunciados, o de alguna otra justificación explícita del diseño.	La valoración de la pertinencia resulta más fácil si existe una teoría del cambio, un modelo lógico o alguna otra justificación explícita del diseño de la intervención, ya que quiere decir que se cuenta con importantes elementos de juicio para evaluar la pertinencia del diseño. Los evaluadores deben identificar si existen dichos modelos y, donde sea necesario, trabajar con las partes interesadas en una reconstrucción o una formulación más clara de la teoría del cambio como punto de partida para la evaluación.	Desde el inicio de la evaluación, los gestores de evaluaciones deben intentar proporcionar a los evaluadores la teoría del cambio, el modelo lógico o cualquier otro documento que describa la lógica explícita detrás del diseño del programa o de la intervención.
El contexto sufrió variaciones a lo largo del tiempo, pero no existe claridad respecto de si la intervención fue adaptada o no, ni cómo.	Por desgracia, las adaptaciones no siempre están adecuadamente documentadas. Es necesario describir primero qué cambios tuvieron lugar y cómo se adaptó la intervención (o no) para poder evaluar la pertinencia. Si no quedan dudas acerca de que el contexto ha sufrido importantes variaciones, afectando la pertinencia, los evaluadores deberán	Los gestores de evaluaciones pueden contribuir a identificar a las instancias de decisión clave y a contactarlas para que colaboren con la “reconstrucción” de los antecedentes de cualquier ajuste al que hubiera sido sometida la intervención. Si las decisiones no hubieran sido adecuadamente documentadas, pueden concertarse entrevistas para averiguar lo que ocurrió.

revisar la documentación de la intervención (por ejemplo, memorias anuales o informes trimestrales, información de seguimiento y, de ser posible, comunicaciones como las actas de las sesiones y anuncios al personal y a los participantes) con el fin de "reconstruir" las modificaciones que se efectuaron e identificar la cronología de las principales decisiones y sus motivaciones.
--

Ejemplos de evaluaciones de pertinencia

Esta sección contiene una muestra transversal de evaluaciones de pertinencia de iniciativas de apoyo al presupuesto general, subsidios de subsistencia para agricultores y programas de apoyo al sector salud.

Recuadro 4.1. Evaluación de pertinencia de una iniciativa de apoyo al presupuesto general

La guía metodológica para la evaluación de intervenciones de apoyo al presupuesto elaborada por la Red de Evaluación del Desarrollo del CAD de la OCDE, EvalNet, sugiere enfocarse en la pertinencia y la idoneidad del diseño de dichas iniciativas y en la combinación de insumos en relación con el contexto político, económico y social del país asociado, el marco de políticas del gobierno y las estrategias de asistencia al desarrollo de los asociados para el desarrollo.

Un ejemplo de una pregunta central sobre pertinencia es: ¿En qué medida fue apropiado y pertinente el diseño del programa de apoyo al presupuesto a la luz del contexto político, económico y social del país, el marco de políticas del gobierno y las estrategias de asistencia al desarrollo de los asociados para el desarrollo externos?

Por ejemplo, la evaluación de la intervención de apoyo al presupuesto en Túnez (citada en la guía) examina el grado en que el apoyo al presupuesto se valió de una combinación adecuada de asistencia técnica, diálogo sobre políticas y fondos apropiados para el contexto institucional del gobierno de Túnez y sus prioridades. En un primer momento, la evaluación examinó los efectos directos para el gobierno (en su calidad de beneficiario). Los efectos directos para los beneficiarios finales —los usuarios de los servicios (por ejemplo, los jóvenes estudiantes de educación superior)— se abordaron en un segundo momento, cuando se analizó la eficacia y el impacto.

Fuente: OCDE OCD (2011^[3]) E, *Evaluating Budget Support: Methodological Approach*,

https://www.oecd.org/dac/evaluation/dcdndep/Methodological%20Approach%20BS%20evaluations%20Sept%202012%20_with%20cover%20Thi.pdf

Recuadro 4.2. Evaluación del proyecto “Mejora de medios de subsistencia de los agricultores de la Zona Seca de Myanmar”

Esta intervención y su evaluación se enfocaron en los medios de subsistencia y la seguridad alimentaria de comunidades pobres y vulnerables asentadas en la Zona Seca Central de Myanmar y constituyen un buen ejemplo de cómo aplicar el criterio de pertinencia de manera apropiada y con un enfoque *ad hoc*.

La evaluación de pertinencia tomó en consideración la capacidad de respuesta de la intervención a las necesidades y prioridades de las partes interesadas, y cómo se involucró a los grupos de partes interesadas a lo largo del diseño y la ejecución del proyecto.

La evaluación hizo hincapié en la pertinencia a lo largo del tiempo, entre otras cosas la adaptabilidad del programa a los cambios en las circunstancias que tuvieron lugar durante la ejecución. Además de la pertinencia durante todo el tiempo que duró el proyecto, la evaluación analizó igualmente cómo se manifestó la pertinencia en coyunturas críticas, comparando la valoración de pertinencia desde el examen de término medio con la pertinencia manifiesta en el momento de la evaluación.

La evaluación concluyó que la intervención es altamente pertinente para las necesidades y el contexto, y especialmente sólida a la hora de adaptarse a los cambios en el entorno y en las políticas a lo largo del tiempo. Aplica la pertinencia de manera apropiada, al usar y examinar las necesidades y el contexto en las etapas tanto de diseño como de ejecución, para arrojar una comprensión equilibrada de la intervención y su evolución.

Fuente: FAO (2020^[4]), *Evaluation of "Improving farmer livelihoods in the dry zone through improved livestock health, productivity and marketing"*, <http://www.fao.org/3/ca8463en/ca8463en.pdf>

Recuadro 4.3. Evaluación de la pertinencia de iniciativas de salud en Bolivia

Esta evaluación examinó el apoyo otorgado al sector salud de Bolivia entre 2009 y 2020. Para evaluar la pertinencia, la evaluación no solamente analizó las solicitudes oficiales aprobadas del gobierno boliviano, en especial del Ministerio de Salud; también cuestionó la capacidad de respuesta frente a necesidades más específicas de las partes interesadas, incluidos los beneficiarios. Concluye lo siguiente:

Pertinencia: Los proyectos responden a las iniciativas y solicitudes de las instituciones y autoridades aprobadas por el Ministerio de Salud, lo que significa que la cooperación italiana responde efectivamente a lo que el gobierno boliviano afirma que es necesario. Sin embargo, no existe un acceso apropiado para identificar las necesidades de los proyectos en general, lo que termina creando problemas de pertinencia y consistencia. Los proyectos con mayores logros entre los evaluados son aquellos en los que se puede demostrar al menos un conocimiento adecuado de las necesidades específicas que se deben atender.

Fuente: Eurecna Spa (2020^[5]), Bolivia - Evaluation of health initiatives (2009-2020), http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf

Coherencia

Definición de coherencia:

Coherencia: ¿En qué medida es compatible la intervención?

Compatibilidad de la intervención con otras intervenciones de un país, sector o institución .

Nota: Medida en que otras intervenciones (en particular políticas) promueven o socaban la intervención, y viceversa. Engloba la coherencia interna y la externa . La coherencia interna aborda las sinergias e interrelaciones entre la intervención y otras intervenciones llevadas a cabo por la misma institución o administración, así como la compatibilidad de la intervención con las normas y reglas internacionales pertinentes a las que se ha adherido dicha institución o administración. La coherencia externa tiene en cuenta la compatibilidad de la intervención con las intervenciones de otros actores en el mismo contexto. Engloba la complementariedad, la armonización y la coordinación con otras intervenciones, y la medida en que la intervención añade valor, sin resultar redundante.

¿Qué es la coherencia y por qué es importante?

En el mundo actual, es necesario prestar mayor atención a la coherencia, con especial hincapié en las sinergias (o las compensaciones) entre las áreas de políticas y en la coordinación intergubernamental, concretamente en los contextos de conflicto y de respuesta humanitaria, y al abordar la emergencia climática.

En consonancia con la Agenda 2030 y los ODS, este nuevo criterio fomenta un enfoque integrado y ofrece un importante lente para evaluar la coherencia, incluyendo las sinergias, la coordinación intergubernamental y la alineación con las normas y los estándares internacionales. Es un espacio en el cual se abordan las compensaciones y las tensiones, y se identifican situaciones en que la duplicación de esfuerzos o las inconsistencias de los enfoques de ejecución de políticas entre diferentes instancias

del gobierno o entre distintas instituciones son susceptibles de socavar el avance general.

Este criterio incentiva asimismo a los evaluadores a entender la función de una intervención dentro de un sistema en particular (una organización, un sector, un área temática o un país), en oposición a adoptar una perspectiva exclusivamente centrada en la intervención o la institución. Si bien la coherencia externa busca esclarecer si los objetivos de política de los actores están alineados con los objetivos de desarrollo internacionales y en qué medida, está incompleta si no se tienen en cuenta los intereses, la influencia y el poder de otros actores externos. En consecuencia, una perspectiva de economía política más amplia resulta de gran valor a la hora establecer la coherencia de las intervenciones.

Además, se están diversificando las fuentes de financiamiento (tanto internacionales como nacionales) del desarrollo sostenible. La referencia a "las normas y los estándares internacionales" fomenta el análisis de la consistencia de la intervención con los compromisos propios del actor en cuestión en el marco del derecho o los acuerdos internacionales, como las normas anticorrupción o los convenios sobre derechos humanos. Esto se aplica a aquellos acuerdos que la entidad ya ha suscrito y que, por lo tanto, quedan cubiertos bajo la coherencia interna. Anteriormente, este tipo de coherencia no se analizaba lo suficiente. Las normas y los estándares internacionales también se pueden evaluar bajo el criterio de pertinencia desde el punto de vista de la capacidad de respuesta de la intervención a las prioridades globales, lo que constituye una óptica complementaria.

Cómo entender la coherencia: elementos de análisis

La coherencia se compone de dos dimensiones: coherencia interna y coherencia externa.

Coherencia interna

La coherencia interna de las políticas considera dos factores: la alineación con los marcos de política más amplios de las instituciones; y la alineación con otras intervenciones realizadas por la misma institución, incluyendo las de otros departamentos responsables de llevar a cabo intervenciones de desarrollo o intervenciones que pueden afectar al mismo contexto operativo². Debe evaluar en qué medida se encuentran armonizadas entre sí estas actividades, si existe una duplicación de esfuerzos o actividades, y si las intervenciones se complementan mutuamente.

En los gobiernos nacionales (o, donde proceda, en otros niveles de gobierno), los desafíos para la coherencia surgen entre diferentes tipos de políticas públicas, diferentes niveles de gobierno y diferentes partes interesadas (tanto estatales como no estatales, comerciales y no comerciales). Esto debe tenerse muy en cuenta en la evaluación de la coherencia para entender dónde encaja la intervención dentro de este panorama general y el grado en que se encuentra alineada con las políticas que rigen el contexto más amplio.

Por ejemplo, las Directrices para la Evaluación de la AOD del Ministerio de Relaciones Exteriores del Japón apoyan a los comisionados y los entes de ejecución de las evaluaciones en la evaluación de la coherencia de las estrategias y actividades diplomáticas y de desarrollo del gobierno del Japón a todos los niveles (Ministerio de Asuntos Exteriores del Japón, 2019^[6]). Dichas directrices proporcionan un marco y una guía para ayudar a los evaluadores a valorar las interconexiones, la complementariedad y la coherencia de las estrategias diplomáticas y de asistencia oficial al desarrollo (AOD) de ese país. Esto fomenta un análisis integral de la actividad del Japón y su apoyo a diferentes sectores y países.

La coherencia de las políticas puede enfocarse desde un punto de vista horizontal. Por ejemplo, en el nexo acción humanitaria-desarrollo-paz puede haber una fuerte necesidad de coherencia, ya que un mismo actor puede llevar a cabo intervenciones que pertenecen al ámbito de la política de desarrollo, militar y de seguridad. En el terreno del medio ambiente, esto podría regir también para la necesidad de coherencia en el nexo agua-energía-alimentación. En otros contextos, se podría tener en cuenta la forma en que algunas áreas que no pertenecen al campo del desarrollo, tales como el comercio, afectan a la intervención.

Desde una perspectiva vertical, la coherencia de las políticas puede abordarse entre los diferentes niveles de una institución o entre las diferentes áreas de financiamiento para el desarrollo de un mismo gobierno (por ejemplo, su institución de cooperación bilateral, su institución financiera de desarrollo [IFD] y sus programas de apoyo multilateral). Podría tener en cuenta, además, la medida en que la intervención apoya o socava los objetivos de política a diferentes niveles geográficos. Por ejemplo, puede tener en cuenta en qué medida una intervención de desarrollo local está alineada con los objetivos de desarrollo nacionales y con las intervenciones a nivel nacional o viceversa.

Coherencia externa

La coherencia externa tiene principalmente en cuenta dos consideraciones: la alineación con los compromisos de política externa; y la coherencia con las intervenciones de otros actores en un determinado contexto.

Desde la perspectiva de la política, la coherencia externa considera la alineación de la intervención con los compromisos de política externa como los ODS y la medida en que estos han sido considerados en el diseño y la ejecución de la intervención. Es importante evaluar el compromiso de una institución con

los ODS a estas alturas, dado que la meta 14 del objetivo 17 de los ODS exhorta a “mejorar la coherencia de las políticas para el desarrollo sostenible”. Se trata de una consideración importante, ya que sintetiza la medida en que tanto la alineación de las políticas como la rendición de cuentas respecto de los ODS se transversalizan y se ejecutan en la práctica.

Al examinar la ejecución de una intervención en un contexto específico, los evaluadores deben considerar la coherencia con las intervenciones de otros actores. Por ejemplo, ¿cómo prestan servicios los diferentes actores, existen superposiciones u omisiones? La coherencia considera la medida en que la intervención añade valor en comparación con otros actores y cómo evita la duplicación de esfuerzos.

La evaluación de la coherencia externa debe mantenerse enfocada en la intervención o institución específica que está evaluando, y, a su vez, situarla en el contexto más amplio de los actores humanitarios y de desarrollo sostenible. Esto implica determinar si las intervenciones están diseñadas en el marco de los sistemas y las estructuras existentes, como los mecanismos de coordinación nacionales o sectoriales, y si se valen de ellos.

Conexiones con otros criterios

La coherencia está particularmente vinculada con la pertinencia, la eficacia y el impacto.

- Mientras que la pertinencia evalúa la intervención a nivel de las necesidades y prioridades de las partes interesadas y los beneficiarios que están directamente implicados, la coherencia escala al siguiente nivel y examina la compatibilidad de la intervención con el sistema más amplio. Tanto la pertinencia como la coherencia tienen en cuenta la medida en que la intervención está alineada con el contexto, pero lo hacen desde perspectivas diferentes.
- La coherencia a menudo es una óptica valiosa a través de la cual empezar a examinar los efectos no intencionales, que se pueden recoger bajo la eficacia y el impacto. Aunque la intervención puede lograr sus objetivos (eficacia), es posible que estos logros se vean revertidos por otras intervenciones (no coherentes) según el contexto.
- Igualmente, existen conexiones con la eficiencia: las intervenciones incoherentes pueden ser duplicativas, lo que supone un desperdicio de recursos.

Cómo integrar la inclusión

La coherencia interna proporciona un lente valioso a través del cual medir la inclusión, especialmente en vista de que está relacionada con los compromisos, las normas y los estándares de derechos humanos. Los evaluadores pueden analizar la compatibilidad de la intervención con las normas que rigen la inclusión y la igualdad a nivel nacional o institucional para las instituciones de ejecución, así como con las perspectivas de organizaciones locales como los grupos de base de poblaciones indígenas y las organizaciones de personas con discapacidad. La valoración de la coherencia puede ofrecer información útil sobre el valor y la coherencia de las actividades que aspiran a reducir la exclusión, beneficiar a los sectores marginados y vulnerables, y provocar una transformación en las desigualdades de género.

El análisis de la inclusión en relación con la coherencia se debe tomar en consideración cuando los evaluadores estudian si el impacto fue inclusivo y la intervención fue pertinente, ya que existen sinergias entre estas tres áreas de indagación evaluativa.

Desafíos de evaluar la coherencia y cómo enfrentarlos

El cuadro que aparece a continuación identifica varios de los principales desafíos que entraña la evaluación de la coherencia, entre ellos, los desafíos relacionados con la amplitud de ámbito, el mandato y la disponibilidad de datos. Se ofrecen sugerencias para enfrentar estos desafíos, tanto para los evaluadores como para los gestores de evaluaciones.

Cuadro 4.2. Desafíos de evaluar la coherencia

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
La coherencia de las políticas supone un desafío en muchos contextos y países, ya que, para una evaluación rigurosa, se requiere acceso a un gran volumen de documentos de política, datos y fuentes de datos.	<p>Los desafíos únicamente pueden resolverse realizando evaluaciones cualitativas basadas en las fuentes de datos e informaciones disponibles. Además, es importante valorar la evaluabilidad al inicio de una evaluación y ser abiertos y transparentes en cuanto a las restricciones. Cuando no se pueda acceder a las fuentes de información, se deberá dejar constancia de ello.</p> <p>Comprender el entorno legislativo y las políticas organizacionales, y utilizar el conocimiento para diseñar planes de recolección y análisis de los datos primarios apropiados.</p>	Los gestores de evaluaciones deben considerar la evaluabilidad y las restricciones de datos y acceso al ordenar una evaluación. Es necesario prever estas carencias en la mayor medida posible y analizarlas y mitigarlas con el equipo de evaluación desde un inicio.
Las organizaciones restringen el acceso a la información al amparo de la legislación o las políticas institucionales de protección de la información.		Considerar las restricciones y el acceso a la información al diseñar y ordenar una evaluación. Estos temas deberán tratarse regularmente con el equipo de evaluación para apoyar el acceso a la información y desarrollar estrategias apropiadas de mitigación de riesgos.
La evaluación de la coherencia en todas las áreas de política pertinentes puede expandir desmesuradamente el ámbito de la evaluación.	Colaborar con el gestor de la evaluación para cerciorarse de que las prioridades y el ámbito estén claramente definidos al inicio de la evaluación.	<p>Decidir cómo debe priorizarse la coherencia, definiendo el ámbito y priorizando las áreas de política que se considerarán. Esto dependerá del propósito y el uso de la evaluación.</p> <p>Al evaluar la coherencia de las políticas en un determinado sector, las instituciones nacionales deben coordinar tanto como sea posible la evaluación de la coherencia en ese sector.</p>
Los mandatos de algunos departamentos de evaluación pueden comprender únicamente las iniciativas de asistencia humanitaria y de desarrollo, restringiendo el ámbito de la evaluación e impidiéndoles evaluar los aspectos relacionados con la coherencia entre los diferentes niveles de gobierno.	Poner especial empeño en la etapa de diseño de la evaluación en definir y revisar el ámbito de la evaluación, y en operativizar la exploración de la coherencia en función de los mandatos de los departamentos de evaluación competentes, o dejar constancia de las restricciones para hacerlo.	<p>La coherencia requiere del reconocimiento de que otras áreas de política, más allá del desarrollo, pueden impactar en una evaluación. En teoría, esto confiere a los evaluadores mayor libertad para explorar áreas que trascienden la influencia inmediata de la intervención. Sin embargo, la operativización presentará desafíos cuando se examinen otras áreas de política que trascienden el mandato de la institución. Se trata de una consideración de evaluabilidad que idealmente debería tratarse antes de la evaluación.</p> <p>Las instituciones deben determinar si evalúan o no la coherencia en las áreas restringidas. Si se considera prioritario, los evaluadores o gestores podrían negociar con otros departamentos para llevar a cabo una evaluación conjunta.</p>

Ejemplos de evaluaciones de coherencia

Aunque la coherencia es un criterio nuevo para el CAD de la OCDE, ha sido incluida en muchas evaluaciones a lo largo de los años. El criterio de coherencia se utiliza rutinariamente asimismo en las evaluaciones humanitarias. Esta sección contiene una muestra transversal de ejemplos que revelan cómo se ha evaluado la coherencia en una evaluación estratégica de la coherencia de las políticas de desarrollo de Noruega, la asistencia a los esfuerzos de recuperación tras un desastre natural en Filipinas y una evaluación de país o de cartera en Montenegro.

Recuadro 4.4. Evaluación de los esfuerzos noruegos para garantizar la coherencia de las políticas de desarrollo

El gobierno noruego se ha comprometido en varias ocasiones a garantizar la coherencia de las políticas para el desarrollo en sus intervenciones. Por “coherencia de las políticas” se entiende “trabajar en aspectos más amplios del desarrollo, además de la asistencia al desarrollo, por ejemplo, en el comercio, la migración, las inversiones, el cambio climático y la seguridad”. Esta evaluación plantea “dilemas” de política ahí donde existen claras contradicciones entre las políticas de los diferentes actores o áreas de política. El informe hace un llamado a “un diálogo más abierto y transparente en torno a los dilemas y las prioridades reales, ya que hay algunos dilemas concretos y decisiones que tomar”.

El propósito de la evaluación era arrojar luz sobre los dilemas de política y los desafíos de asegurar la coherencia de las políticas para el desarrollo. Se considera la coherencia tanto en el contexto nacional individual como a escala mundial. Se exploran los elementos tanto internos como externos de la coherencia:

- coherencia interna de las políticas que rigen las intervenciones lideradas por el Ministerio de Relaciones Exteriores y otros actores noruegos;
- coherencia externa para establecer la medida en que los objetivos de política de los actores estaban alineados con los objetivos de desarrollo internacionales.

Tomando Myanmar como estudio de caso, la evaluación considera que existen inconsistencias entre las políticas de desarrollo y otras áreas de política. Se seleccionó a Myanmar debido a la diversidad de actores de desarrollo noruegos, tradicionales y no tradicionales, presentes en el país y a los potenciales conflictos que surgen entre sus prioridades de política a nivel local. La evaluación adopta un interesante enfoque metodológico para el análisis de los dilemas de política. Examina aquellas áreas donde las organizaciones no gubernamentales (ONG), organizaciones políticas o instituciones internacionales han identificado un desafío al detectar políticas incoherentes en las prioridades e intervenciones noruegas. Por ejemplo, la evaluación puso de manifiesto dos dilemas: la implicación de empresas noruegas en Myanmar y su compromiso con la construcción de la paz; y la participación en la explotación de los recursos del país y su compromiso con la paz. Este análisis pone de relieve asimismo un gran desafío para la evaluación de la coherencia y la adopción de medidas concretas para garantizarla. De hecho, el Ministerio de Relaciones Exteriores no cuenta con el mandato de intervenir en las decisiones de política de otros ministerios, complicando su capacidad para evaluar y resolver las inconsistencias. En el análisis de la escala global, la evaluación tuvo en cuenta asimismo la incoherencia externa con los compromisos globales de Noruega. Encontró desafíos para la alineación de las políticas nacionales con las prioridades de desarrollo internacionales.

La evaluación demuestra una aplicación minuciosa del criterio de coherencia, al ilustrar los desafíos de poner en práctica intervenciones coherentes a nivel de políticas que a su vez sean sensibles tanto al contexto operativo como a las prioridades institucionales.

Fuente: Norad (2018^[7]), *Evaluation of Norwegian Efforts to Ensure Policy Coherence for Development*, <https://www.norad.no/contentassets/4ac3de36fdd4229811a423f4b00acf7/8.18-evaluation-of-norwegian-efforts-to-ensure-policy-coherence-for-development.pdf>

Recuadro 4.5. Evaluación de la coherencia de la asistencia para la reconstrucción tras un desastre natural en las Filipinas

Esta evaluación examina los programas de asistencia internacional de Canadá y su desempeño como donante en situaciones de desastres naturales. Ilustra cómo se puede evaluar la coherencia externa e internamente para una intervención. Analiza el programa de reconstrucción tras el paso del tifón Haiyan por un valor de 20,5 millones de dólares canadienses, así como la relación bilateral entre Canadá y Filipinas desde 2013-2014 hasta 2018-2019.

Evaluación de coherencia

Si bien la evaluación se llevó a cabo antes de que la coherencia fuese incorporada como criterio de evaluación del CAD de la OCDE, es indudable que la evaluación se enfocó en la coherencia, aunque no se mencione explícitamente. La evaluación considera la coherencia en relación con los elementos de análisis tanto externos como internos.

- **Coherencia externa:** la evaluación examina el desempeño de Canadá como donante, su comportamiento en los contextos operativo y político, y la medida en que esta intervención era compatible con dichos contextos. La evaluación destacó la alineación de la asistencia canadiense con las prioridades y el programa de recuperación nacional del gobierno filipino, y tomó en consideración la complementariedad de la intervención con otros programas que operan en la región.
- **Coherencia interna:** la evaluación examinó los programas de asistencia internacional de Canadá y los contextualizó frente a una muestra más amplia de programas de asistencia internacional. Puso un énfasis especial en la alineación del programa con la política canadiense en general y evaluó la medida en que la intervención encajaba con el enfoque canadiense típico para la respuesta de socorro en casos de desastre.

Además, revela la medida en que la coherencia y la pertinencia se superponen estrechamente, ya que la intervención se relaciona de manera constante con las necesidades y las prioridades de las partes interesadas y del contexto, ambos elementos de análisis de la pertinencia. Sin embargo, la evaluación va un paso más allá al analizar no solamente la capacidad de respuesta de la intervención, sino también su alineación y su consideración en el diseño y la ejecución de una intervención clara y coherente con la política canadiense, compatible con los intereses canadienses de carácter general en las Filipinas y consistente con sus otros programas de socorro en casos de desastre.

Fuente: Global Affairs Canada (2019^[8]), *Evaluation of Natural Disaster Reconstruction Assistance in the Philippines, 2013-14 to 2018-19*, <https://www.international.gc.ca/gac-amc/publications/evaluation/2019/endra-earcn-philippines.aspx?lang=eng>

Recuadro 4.6. Evaluación de la coherencia de la cooperación para el desarrollo de Eslovenia con Montenegro

Esta evaluación describe la ejecución de la cooperación para el desarrollo entre la República de Eslovenia y Montenegro entre 2013 y 2016. El objetivo de la evaluación era “generar evidencia del desempeño (esto es, de la medida en que se lograron los objetivos de la cooperación para el desarrollo), analizar las razones de los logros y los fracasos en el desempeño, y plantear recomendaciones para la formulación de políticas, programas y proyectos en el futuro”.

La evaluación incluyó la coherencia como uno de ocho criterios (también se incluyeron la gestión de programas, la entrega y el valor añadido para Eslovenia), formulando la siguiente pregunta:

- ¿Las contradicciones con otras políticas han impedido la ejecución y el logro de los objetivos de desarrollo, o las políticas se refuerzan mutuamente?

La evaluación examinó la coherencia externa con otras políticas vinculadas a la cooperación para el desarrollo, entre ellas el proceso de negociación con la UE y las inversiones en infraestructura ambiental. La evaluación concluyó que los programas estaban alineados con otras políticas.

La evaluación examinó asimismo la complementariedad con las actividades de otros donantes y halló que la falta de sistemas de coordinación suficientes entre los donantes en Montenegro constituía un obstáculo para lograr la complementariedad o prevenir la duplicación.

Fuente: Ministerio de Relaciones Exteriores de Eslovenia (2017^[9]), *Evaluation of Slovenia's Development Cooperation with Montenegro 2013-2016 period: Final Report*, <https://www.gov.si/assets/ministrstva/MZZ/Dokumenti/multilateral/razvojno-sodelovanje/Development-cooperation-with-Montenegro-evaluation-final-report.pdf>

Eficacia

Definición de eficacia:

Eficacia: ¿La intervención está logrando sus objetivos?

Medida en que la intervención ha logrado, o se espera que logre, sus objetivos y sus resultados, incluido cualquier resultado diferencial entre los distintos grupos.

Nota: El análisis de la eficacia supone tener en cuenta la importancia relativa de los objetivos o los resultados. El término eficacia se utiliza también como un consolidado de la medida en que una intervención ha logrado, o se espera que logre, impactos pertinentes y sostenibles de manera eficiente y coherente.

¿Qué es la eficacia y por qué es importante?

La eficacia ayuda a evaluar la medida en que una intervención está logrando, o ha logrado, sus objetivos. Puede proporcionar información sobre el logro o no de los resultados programados de una intervención, el proceso que se siguió para hacerlo, qué factores fueron determinantes en este proceso y si hubo efectos no intencionales. La eficacia se centra en los resultados más cercanamente atribuibles a la intervención, y es importante diferenciarla del impacto, que examina los efectos de carácter general y los cambios más amplios.

Para examinar el logro de los objetivos en la cadena de resultados o en la ruta causal se requiere una clara comprensión de los propósitos y objetivos de la intervención. Por ello, utilizar el lente de la eficacia puede ayudar a los evaluadores o los gestores u oficiales de programas y otras personas a formular (o evaluar) objetivos claros. Del mismo modo, la eficacia puede ser de utilidad para los evaluadores cuando tienen que identificar si el logro de los resultados (o la falta de ellos) se debe a defectos en la ejecución de la intervención o en su diseño.

Bajo el criterio de eficacia, los evaluadores deben identificar también los efectos no intencionales. Idealmente, los gestores de proyectos habrán identificado los riesgos durante la etapa de diseño, y los evaluadores pueden hacer uso de este análisis al iniciar su evaluación. La exploración de los efectos no intencionales es importante para identificar los resultados tanto negativos (por ejemplo, exacerbación de la dinámica del conflicto) como positivos (por ejemplo, innovaciones que redundan en favor de la eficacia). Es posible que las instituciones que ordenan las evaluaciones quieran proporcionar orientaciones a los evaluadores sobre los estándares mínimos que rigen la identificación de los efectos no intencionales, especialmente cuando estos tienen que ver con violaciones de los derechos humanos u otras consecuencias graves involuntarias.

La definición de eficacia incentiva a los evaluadores y gestores a formular importantes preguntas en torno a la distribución de los resultados entre los diferentes grupos, sea esta intencional o no. Esto es deliberado y la intención es reforzar las consideraciones de equidad, en concordancia con la prioridad de los ODS de “no dejar a nadie atrás”. Exhorta a los evaluadores a examinar los aspectos de equidad, esté incluida o no la equidad como un objetivo específico de la intervención. Un análisis de este tipo requiere de datos y puede implicar una inversión de recursos, lo cual a menudo está justificado por las valiosas conclusiones que puede proporcionar la evaluación³.

Al extraer conclusiones acerca de la eficacia, las evaluaciones deben concentrarse en los resultados que son más importantes a la luz del contexto y para los destinatarios de la evaluación. El término “importancia relativa” hace hincapié en el mensaje de que, al extraer las conclusiones sobre la eficacia, debemos ejercer el juicio evaluativo y ponderar la importancia de los objetivos y los resultados alcanzados y no alcanzados, entre ellos las consecuencias no intencionales.

Cómo entender la eficacia: elementos de análisis

La definición de la eficacia engloba los siguientes conceptos clave: logro de los objetivos, cambios en la importancia de los objetivos y los resultados, resultados diferenciados entre los grupos y comprensión de los factores que influyen en los efectos directos.

Logro de los objetivos

El énfasis principal de la evaluación de la eficacia sigue estando en establecer si una intervención ha logrado sus resultados previstos a distintos niveles de la cadena de resultados (generalmente a nivel de los productos y efectos directos, pero también en algunos casos de los impactos). La cadena de resultados debería estar especificada como parte del diseño de la intervención y constituye el referente más importante para la gestión, el seguimiento y la evaluación de la intervención.

Es sumamente difícil evaluar la eficacia de una actividad si los objetivos declarados o los resultados programados de la actividad son vagos o ambiguos, o si han variado en el transcurso de la intervención sin actualizarse ni reestructurarse. Los gestores de la intervención deberían al menos explicar por qué han variado los objetivos y cuáles son los objetivos nuevos. Si esto no ha sucedido, los evaluadores tendrán que consultar los documentos de la intervención o entrevistar a las partes interesadas, a efectos de recrear la lógica que subyace a los cambios en la intervención a lo largo del tiempo. Sobre la base de la lógica reconstruida, los evaluadores podrán entonces juzgar la pertinencia de los objetivos y la eficacia en el logro de estos.

Evaluar la eficacia también es importante en programas adaptables en que los cambios se efectúan de manera iterativa, en respuesta a la retroalimentación de las partes interesadas, los resultados que van emergiendo y las variaciones en el contexto. En los programas adaptables, el diseño y la ejecución de una intervención pueden sufrir múltiples cambios incrementales a lo largo del tiempo. En dichos casos, es importante que los evaluadores reflexionen y revisen las teorías de cambio en función de los sistemas más amplios en los que se inscribe la intervención y tengan en cuenta los registros de cómo se hicieron los cambios y por qué. La evaluación de eficacia debe reflejar los objetivos actuales de la intervención. El examen de la lógica y de la necesidad de cambios en la ejecución o los objetivos (a menudo recogidos en las teorías de cambio o en los marcos de resultados actualizados) debe fundamentar las evaluaciones de la eficacia.

Evaluar la eficacia puede implicar identificar los cambios observables en el grupo meta o en el entorno a lo largo de la ejecución de la intervención, como también establecer la causalidad de los cambios observados a distintos niveles; esto es, demostrar que los cambios se debieron a la intervención o que la intervención contribuyó a los cambios, en oposición a otros factores del entorno o, alternativamente, a otra intervención. Deben diseñarse metodologías para que el evaluador pueda llegar al fondo de cómo se produjeron los resultados y cuáles fueron las razones de los logros (los factores que los explican), de los logros insuficientes o de los fracasos.

Evaluar la eficacia incluye examinar los resultados de la intervención. El CAD de la OCDE define los resultados como los productos, efectos directos o impactos de la intervención (OCDE, 2002^[10]). Por consiguiente, la evaluación de la eficacia puede incluir también una evaluación de los efectos no intencionales, tanto positivos como negativos, que hubieran ocurrido como resultado de la intervención. En la ejecución de las intervenciones siempre existe el potencial para generar efectos sociales, económicos o ambientales no intencionales, o de generar efectos que no son voluntarios pero que hubieran podido preverse. Por eso, en las evaluaciones se debe tener cuidado de examinar los efectos que recaen fuera de los especificados en los objetivos de la intervención. Esto puede hacerse extensivo también al examen de los potenciales beneficios o riesgos que se derivan de estos efectos no intencionales (predecibles o impredecibles). La medida en que la intervención ha contribuido a la realización de las metas y los objetivos de desarrollo nacionales u otros según el contexto recae asimismo bajo el criterio de eficacia, mientras que el potencial para estas contribuciones se examinará bajo el criterio de pertinencia.

Ponderación de la importancia relativa de los logros alcanzados

Al evaluar la eficacia, los evaluadores exploran el logro (o la falta de este) de los diversos objetivos y resultados de la intervención. Si se cumplieron algunos de los objetivos, pero no todos, los evaluadores necesitarán examinar la importancia relativa de los objetivos alcanzados para extraer conclusiones acerca de la eficacia. Para ello pueden basarse en el análisis de la pertinencia, que debería haber incidido en las potenciales diferencias entre las prioridades y necesidades de las diversas partes interesadas. Por lo tanto, los evaluadores pueden concluir que la intervención fue eficaz en algunos aspectos, pero no en otros, o que fue eficaz desde la perspectiva de algunas de las partes interesadas, pero no de otras.

Resultados diferenciados

Los evaluadores deben considerar la inclusión y la equidad de los resultados entre los grupos beneficiarios, independientemente de que los beneficiarios sean personas o instituciones. El análisis de los resultados diferenciados puede incluir un examen de la medida en que la intervención garantizó un enfoque inclusivo en su diseño y ejecución. Por ejemplo, una evaluación podría examinar el proceso mediante el cual se formularon los objetivos de la intervención. Esto incluye determinar si los objetivos se formularon tras un proceso de diagnóstico de necesidades y consulta con las partes interesadas (incluyendo el grupo meta principal). Es posible que a través de este proceso se pueda determinar

asimismo si la intervención desperdició algunas oportunidades para generar resultados favorables para su población meta o beneficiaria, entre otras cosas contribuyendo a generar un cambio a largo plazo, por ejemplo, una reducción de las desigualdades. Los evaluadores pueden examinar los resultados no intencionales o imprevistos al igual que el resultado intencional, recordando que ciertos cambios y transformaciones son sutiles o a largo plazo, y pueden ser difíciles de cuantificar.

Factores influyentes

El análisis de los factores que influyen en los resultados es importante porque ayuda a los evaluadores a establecer *por qué* una intervención pudo o no pudo alcanzar sus objetivos, lo cual ayuda a los asociados a identificar las áreas susceptibles de mejora. Dichos factores pueden ser internos a la intervención o externos. Pueden comprender factores a nivel de: gestión, recursos humanos, aspectos financieros, aspectos normativos, modificaciones a la ejecución o desviaciones de los planes. La calidad de la ejecución (y el cumplimiento de los protocolos de ejecución) a menudo es un factor que contribuye a la eficacia y debe describirse antes de evaluar la eficacia, la eficiencia, el impacto y la sostenibilidad.

Desde el punto de vista externo, los evaluadores deben considerar los efectos positivos y negativos asociados al contexto de la intervención, que, a su vez, influyen en que se logren o no los resultados. Esto puede incluir evaluar la capacidad de adaptación de la intervención en respuesta a los cambios en el contexto. Las evaluaciones pueden examinar asimismo la oportunidad de los resultados (por ejemplo, el escalonamiento del apoyo a los entornos o grupos meta, que favoreció la entrega de los resultados).

Conexiones con otros criterios

La eficacia está relacionada con otros criterios, especialmente la pertinencia y el impacto:

- Bajo la pertinencia, se identifican los objetivos de la intervención; bajo la eficacia se determinan los avances hacia el logro de dichos objetivos. Por supuesto, es posible que una intervención que carece de pertinencia se ejecute no obstante con eficacia. En caso de una desconexión de esta naturaleza, los evaluadores necesitarán ejercitar su buen criterio a la hora de extraer conclusiones en general, ya que no se puede simplemente ignorar los hallazgos de un criterio a favor de otro.
- La eficacia y el impacto son criterios complementarios que inciden en diferentes niveles de la cadena de resultados. La eficacia se ocupa del logro de resultados en función de los objetivos de la intervención, es decir, a nivel de productos y efectos directos; mientras que el impacto se enfoca en los resultados de carácter general, es decir, cuáles son los resultados declarados a nivel general y qué contribuye a ellos. En general, los gestores y los evaluadores de las intervenciones deben cerciorarse de que exista una distinción manifiesta entre los diferentes niveles de resultados (es decir insumo, producto, efecto directo e impacto) y de que no queden dudas respecto de qué aspectos serán evaluados bajo cada criterio.

Cómo integrar la inclusión

La definición de la eficacia incentiva un análisis en profundidad de la equidad entre los diferentes grupos. Los evaluadores deben valorar la medida en que la intervención ha sido inclusiva para los diferentes grupos de beneficiarios y cómo se han incorporado principios centrales como la equidad, la no discriminación y la rendición de cuentas en todas las etapas, desde el diseño hasta los resultados de la intervención. Al tenor del principio de “no dejar a nadie atrás”, hay que poner especial énfasis en la medida en que la intervención ha suplido las necesidades de los sectores en mayor desventaja. Es importante examinar el logro y la distribución de los resultados en relación con los diversos grupos de beneficiarios y explicar cualquier diferencia que se observe.

Por otra parte, los evaluadores deben analizar si los resultados contribuyen a la eliminación de la desigualdad, cómo y por qué. Bajo este criterio, los evaluadores deben examinar cómo impactan determinadas actividades en el bienestar de grupos específicos y si estas actividades constituyen oportunidades de empoderamiento para los y las participantes.

Desafíos de evaluar la eficacia y cómo enfrentarlos

El cuadro a continuación identifica los desafíos a nivel de disponibilidad de información, formulación de los objetivos y atribución de los resultados, y sugiere formas de hacerles frente tanto para los evaluadores como para los gestores de evaluaciones.

Cuadro 4.3. Desafíos de evaluar la eficacia

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
Ausencia de datos apropiados, informativos o correctamente desglosados.	La disponibilidad de datos (por ejemplo, la falta de una línea de base o de evaluaciones de término medio) puede impedir la medición del logro de los objetivos, ya que exacerba las dificultades para atribuir la causalidad y complica las posibilidades de determinar quién está contribuyendo a qué. La información de base puede ser reconstruida, pero con resultados menos precisos y fiables que si se dispusiera de información de base correctamente recopilada. Otros enfoques de los que pueden valerse los evaluadores son: 1) la triangulación y validación de los datos disponibles; 2) el uso de métodos alternativos; y 3) la construcción de una historia de contribución.	Idealmente, la disponibilidad de información debería asegurarse en la etapa de diseño de la intervención, para cerciorarse de que se recojan datos apropiados, informativos y correctamente desglosados a lo largo del ciclo de vida de la intervención. Los gestores de las evaluaciones deben estar al tanto de la disponibilidad de información y de la posibilidad de usar fuentes alternativas, y trabajar con los evaluadores en el desarrollo de enfoques alternativos cuando no existan datos disponibles. Recolectar los diagnósticos de otras evaluaciones e informes de investigación del área. Revisar y considerar cómo aplicar las actividades previas de elaboración de indicadores en áreas de trabajo relacionadas, entre ellas las metas de los ODS, y cómo aprender de dichas actividades.
Trabajar con objetivos o propósitos precariamente formulados, ambiguos o difíciles de medir.	Cerciorarse desde la fase inicial de la evaluación de que los objetivos de la intervención sean verificables y realistas, y estén alineados con los estándares internacionales en vigor para las intervenciones de desarrollo. Los indicadores para medir el logro de los objetivos también deben validarse en función de criterios generalmente aceptados como los indicadores SMART (específicos, mensurables, alcanzables, pertinentes y oportunos). Idealmente, este proceso debe acometerse de manera participativa entre los evaluadores y los gestores de evaluaciones.	Trabajar conjuntamente con los evaluadores en la identificación de criterios claros para medir el logro de los objetivos en virtud de los estándares generalmente aceptados (por ejemplo, indicadores SMART). En evaluaciones con una perspectiva de género, considerar pasar de indicadores SMART y RACER a indicadores CREAM (claros, pertinentes, económicos, adecuados y monitoreables) y SPICED (subjetivos, participativos, interpretables, contrastados y comparados, empoderadores, diversos y desglosados). Mejorar la definición y la concreción de los objetivos y los resultados, en el marco de la teoría del programa o la teoría del cambio definida en ese momento. En caso de que no exista una teoría del cambio, intentar reconstruir una que sea válida para contextualizar la lógica interna de la intervención. Mejorar la coherencia de la formulación <i>a posteriori</i> , reubicando el proyecto sobre la base de prácticas participativas y con la implicación de las partes interesadas.
Atribución de los resultados	El tema de la atribución de la causalidad o contribución representa un desafío para la evaluación de la eficacia, especialmente en el largo plazo. La reconstrucción de una historia de contribución y la creación de un modelo apropiado de triangulación de los flujos de datos puede ayudar al evaluador a identificar si es posible atribuir los resultados a la intervención sometida a evaluación.	Tener en cuenta las opiniones, evaluaciones y percepciones de las partes interesadas que conocen bien la intervención o la entidad, y que podrían ser potenciales asociadas en la ejecución de la intervención o en el logro de los resultados de desarrollo.

Ejemplos de evaluaciones de la eficacia

Esta sección presenta una muestra transversal de ejemplos de cómo se ha evaluado la eficacia en un caso de asistencia electoral y en la evaluación de un programa de país en Cabo Verde.

Recuadro 4.7. Evaluación de la eficacia y la inclusión de la asistencia electoral australiana

Desde los años 70 en adelante, Australia ha brindado asistencia electoral a una serie de países en desarrollo, con el fin de apoyarlos en sus procesos de transición democrática y consolidación de democracias formales, de profundo arraigo y funcionales. Esta evaluación examina la asistencia australiana a las elecciones nacionales en ocho países de la región de Asia y el Pacífico entre 2006 y 2016. Representa un buen ejemplo de la aplicación del criterio de eficacia, con referencia especial a los cambios en los objetivos a lo largo del tiempo y también a la eficacia en relación con la inclusión, el género y los grupos vulnerables.

La evaluación se enfocó en los objetivos generales de la asistencia electoral australiana y pasó luego a examinar la eficacia y la inclusión en tres áreas centrales: el fortalecimiento de los sistemas de gestión electoral, la promoción de la participación y la ayuda con la organización de elecciones.

La evaluación observó también que los objetivos de Australia en materia de asistencia electoral variaron durante el período de ejecución y dio cuenta de este giro en los objetivos en su análisis. Por otra parte, la evaluación examina atentamente los riesgos asociados a la provisión de una asistencia electoral eficaz e inclusiva, dando cuenta de los factores externos que pueden inhibir la eficacia de las intervenciones actualmente y en el futuro.

En el análisis de los objetivos, la evaluación examinó específicamente la inclusión de los resultados, lo que resulta útil para ilustrar la medida en que las evaluaciones pueden examinar los resultados diferenciados entre grupos de personas al evaluar la eficacia. El logro de los resultados se analizó utilizando un lente de género y la evaluación se valió de hallazgos cualitativos para concluir que las normas sociales y las barreras jurídicas e institucionales estaban dando lugar a unas cifras desproporcionadamente inferiores de participación electoral entre las mujeres.

La evaluación incidió asimismo en la inclusión de sectores vulnerables, en este caso de las personas con discapacidad, e identificó que la equidad de los resultados para los diferentes grupos también fue limitada debido a la existencia de barreras culturales, a pesar de las medidas adoptadas para mejorar el acceso físico. A partir de estos hallazgos, la evaluación pudo formular recomendaciones para intervenciones futuras centradas en incrementar la equidad de los resultados, lo que, a su vez, refuerza el logro general de los resultados de las intervenciones y los objetivos de desarrollo.

Fuente: Arghiros *et al.* (2017^[11]), *Making it Count: Lessons from Australian Electoral Assistance 2006-16*, <http://www.oecd.org/derec/australia/australia-electoral-assistance-2006-2016.pdf>

Recuadro 4.8. Evaluación de la eficacia de la estrategia y el programa de país del Banco Africano de Desarrollo en Cabo Verde

Esta evaluación del apoyo del Banco Africano de Desarrollo (BAfD) al sector electricidad en Cabo Verde constituye un ejemplo de la aplicación del criterio de eficacia en la práctica, incluyendo la evaluación de la pertinencia del diseño y de la ejecución de la intervención, con especial énfasis en los objetivos de género e inclusión social. La asistencia del BAfD tenía varios objetivos generales, entre ellos ayudar al país a promover un crecimiento inclusivo a través del desarrollo de infraestructura y reformas de buena gobernanza, y el objetivo estratégico de apoyar la transición del país hacia un desarrollo de “crecimiento verde”, con bajas emisiones de gases de efecto invernadero y resiliente al clima.

Para evaluar la eficacia de la intervención y si estos objetivos se habían cumplido, la evaluación valoró si la intervención había generado los siguientes efectos directos en el desarrollo de acuerdo a lo programado:

- aumento del acceso de la población a electricidad;
- aumento del porcentaje de energía renovable;
- aumento de la disponibilidad de superficie agrícola y abastecimiento de agua;
- mejoras a la buena gobernanza.

Para ello, la evaluación analizó factores internos adicionales como la solidez del diseño y los avances en la ejecución de la intervención, así como los factores externos que la afectaban. Además, la evaluación de la equidad de los resultados puso un énfasis especial en los aspectos de igualdad de género e inclusión social.

La evaluación halló que, si bien se observaba un firme avance a la hora de ayudar al país a promover un crecimiento inclusivo mediante el desarrollo de infraestructura y reformas de buena gobernanza, el objetivo estratégico de contribuir a la transición del país hacia un desarrollo de “crecimiento verde” o bajo en emisiones y resiliente al clima no contó con los recursos e instrumentos necesarios. La evaluación reveló que examinar la eficacia puede arrojar luz sobre si se logran o no los resultados programados de una intervención, así como sobre los factores que influyeron en la entrega de los resultados.

Fuente: IDEV (2018^[12]), *Cabo Verde: Evaluation of the Bank's Country Strategy and Program 2008–2017: Summary Report*, <http://www.oecd.org/derec/afdb/AfDB-2008-2017-cabo-verde-bank-strategy.pdf>

Eficiencia

Definición de eficiencia:

Eficiencia: ¿Se están utilizando adecuadamente los recursos?

Medida en que la intervención brinda, o podría brindar, resultados de forma rentable y oportuna.

Nota: “Rentable” se refiere a la conversión de insumos (fondos, conocimientos especializados, recursos naturales, tiempo, etc.) en productos, efectos directos e impactos de la manera más eficaz en función del costo, en comparación con alternativas viables en el contexto en cuestión. “Oportuna” se refiere a dentro del plazo previsto o de un plazo razonable en función de las exigencias impuestas por la evolución del contexto. Puede abarcar la determinación de la eficiencia operativa (es decir, de la correcta gestión de la intervención).

¿Qué es la eficiencia y por qué es importante?

Este criterio representa una oportunidad para verificar si los recursos de una intervención pueden justificarse a la luz de sus resultados, lo cual es de gran importancia práctica y política. La eficiencia es importante para muchos grupos de partes interesadas, entre ellos los gobiernos, la sociedad civil y la población beneficiaria. Un mejor uso de los limitados recursos implica que se pueden potenciar los logros con la cooperación para el desarrollo, por ejemplo a la hora de avanzar en los ODS, donde las necesidades son enormes. La eficiencia es de especial interés para los gobiernos que rinden cuentas a sus contribuyentes, quienes, con frecuencia, cuestionan la relación precio-calidad de diferentes políticas y programas, en especial de las decisiones relativas a la cooperación para el desarrollo internacional, que tienden a estar sujetas a un escrutinio más severo.

También en términos operativos, la eficiencia desempeña un papel importante. Muchas intervenciones se encuentran con problemas de viabilidad y ejecución, especialmente en relación con la forma en que se utilizan los recursos. La evaluación de la eficiencia contribuye a mejorar los incentivos de gestión para garantizar que los programas estén bien ejecutados, llamando a rendir cuentas a los gestores respecto de cómo se tomaron las decisiones y se administraron los riesgos.

Existen varios supuestos y aspectos importantes que se deben tener en cuenta:

- **Recursos:** deben entenderse en el sentido más amplio e incluir todos los costos económicos (humanos, ambientales, financieros y de tiempo). No es lo mismo que el presupuesto del programa o el dinero gastado.
- **Resultados:** también deben entenderse en un sentido amplio y general, cubriendo la totalidad de la cadena de resultados: productos, efectos directos e impactos. Dependiendo del tipo de evaluación, algunas organizaciones asocian la eficiencia solo con los productos; sin embargo, el criterio se define y conceptualiza aquí para fomentar la evaluación de la eficiencia también en relación con los efectos de carácter general, es decir, los impactos, aunque esto a menudo puede ser un desafío.
- **Evaluabilidad:** La capacidad de evaluar la eficacia, el impacto, la coherencia y la sostenibilidad influye en la información que se puede obtener sobre la eficiencia.
- **Eficiencia:** Se refiere a la elección entre varias alternativas viables susceptibles de generar resultados similares dentro de los recursos dados. Para poder hacer comparaciones de costo-efectividad, deben identificarse alternativas que sean genuinamente viables y comparables en términos de calidad y resultados.

Por todo esto, el análisis de la eficiencia debe estar firmemente enraizado en un análisis del contexto, ya que, en algunos casos, puede resultar más costoso llegar a los beneficiarios previstos, pero también más importante y justificable en razón de los impactos en el desarrollo.

Cómo entender la eficiencia: elementos de análisis

Evaluar la eficiencia supone examinar tres áreas clave: la eficiencia económica, la eficiencia operativa y la oportunidad.

Eficiencia económica

Se trata del elemento principal para el análisis de la eficiencia. La eficiencia económica se refiere aquí a la ausencia de desperdicio y la transformación de los insumos en resultados de la manera más eficiente en función de los costos. Incluye la evaluación de la eficiencia de los resultados a todos los niveles de la cadena de resultados: productos, efectos directos e impactos. Supone también una evaluación de la medida en que se hicieron las elecciones apropiadas y se resolvieron las compensaciones tanto en la etapa de diseño como en la ejecución. Estas elecciones incluyen la forma en que se asignaron los recursos entre los diferentes grupos meta y los diferentes períodos de tiempo, así como las opciones que había disponibles para la adquisición de insumos en las condiciones del mercado.

Eficiencia operativa

La eficiencia operativa también es un importante elemento que considerar. Incide en la medida en que se invirtieron adecuadamente los recursos durante la ejecución de la intervención. Algunas de las preguntas que ayudan a explorar la eficiencia operativa son las siguientes: ¿Se invirtieron los recursos humanos y financieros de acuerdo a lo programado, de manera apropiada y en su totalidad (o, por el contrario, se asignaron recursos indebidamente o se gastó menos o más de lo presupuestado)? ¿Se redirigieron los recursos a medida que cambiaban las necesidades? ¿Se gestionaron los riesgos? ¿Se tomaron decisiones que ayudaron a mejorar la eficiencia en respuesta a la aparición de nueva información? ¿Fueron óptimas las decisiones de logística y adquisiciones?

Oportunidad

Estrechamente relacionada con la eficiencia tanto económica como operativa, la oportunidad parte de la pregunta de si se lograron los resultados dentro de los plazos establecidos y en qué medida. Asimismo, se trata de una ocasión para verificar si los plazos eran realistas o apropiados en primer lugar. Además, ¿se modificaron razonablemente los plazos en el transcurso de la intervención, dado que, en muchas intervenciones, es probable que surjan factores y cambios externos al programa? Los evaluadores deben analizar si se hicieron esfuerzos para superar los obstáculos y mitigar los retrasos en la gestión de la intervención a medida que evolucionaba la situación.

Conexiones con otros criterios

Como ya hemos mencionado, los diferentes criterios están conectados entre sí y deben verse como lentes alternativos para examinar la intervención, y no como límites rígidos. Algunas de las interconexiones con otros criterios son las siguientes:

- **Pertinencia y eficiencia:** Un aspecto central de la pertinencia operativa es si el diseño de la intervención respondía bien al contexto y permitía tener en cuenta consideraciones de viabilidad y capacidad. En términos prácticos, el hecho de que el diseño fuera viable y se pudiera poner en práctica también tiene un efecto directo en la eficiencia. Por lo tanto, en este aspecto específico, el evaluador puede terminar examinando ambos criterios juntos.

- **Eficiencia y resultados:** Dado que la eficiencia supone evaluar en qué medida los recursos utilizados se convirtieron en resultados, es preciso considerar todos los aspectos de los resultados (esto es, las preguntas que surgen durante la evaluación de la eficacia, el impacto y la sostenibilidad). La eficiencia operativa está estrechamente ligada a la eficacia y el impacto. Con frecuencia, examinar la medida en que las cosas están funcionando adecuadamente en una intervención supone analizar la eficacia y la eficiencia simultáneamente. Esto es especialmente aplicable, por ejemplo, a la hora de identificar los cuellos de botella y cómo abordarlos, o para garantizar que los recursos se asignen ahí donde se necesitan.

Cómo integrar la inclusión

Mediante el lente del criterio de eficiencia, los evaluadores pueden establecer cómo se entiende e integra la inclusión en la gestión de la intervención, y el grado en que el uso de los recursos refleja las experiencias y los resultados diferenciados para diferentes personas. El costo de lograr resultados a menudo varía de acuerdo a los beneficiarios, siendo los “más rezagados” a los que resulta más difícil (y caro) acceder. El análisis de la eficiencia debe estar impregnado por lo tanto de una clara comprensión de las desigualdades y las dinámicas de poder según el contexto, así como de una comprensión de la medida en que la intervención es compatible con la necesidad de un cambio transformador para combatir las desigualdades subyacentes. El análisis de la eficiencia es un proceso clave que considerar, independientemente de si el compromiso de “no dejar a nadie atrás” (y el compromiso con el objetivo de la Agenda 2030 de generar un cambio transformador para los grupos marginados) ha sido auténtica y efectivamente operativizado a todos los niveles de la gestión, la toma de decisiones y la asignación de recursos.

Aquí, el análisis puede incluir cómo y por qué se asignan los recursos entre los diferentes grupos seleccionados por la intervención, y la medida en que la asignación de recursos se basó en las necesidades y el trabajo con los grupos marginados. Los evaluadores pueden considerar si se generan resultados inclusivos y equitativos a un costo razonable, qué se entiende por “costo razonable” y cómo se establece y varía el costo para los diferentes grupos de beneficiarios. Por ejemplo, si la intervención se compromete a beneficiar a sectores específicos, ¿se han asignado y justificado los recursos suficientes para hacerlo correctamente?

Establecer qué voces se escuchan y se tienen en cuenta al tomar las decisiones en torno a cómo se diseñan las políticas, cómo se invierten los recursos y quién tiene el control y la supervisión de estos procesos es esencial. Cuando la lógica y los planes de la intervención incluyen cambiar estructuras y dinámicas de poder basadas en la desigualdad, los evaluadores deben analizar en qué medida han logrado su cometido o si, por el contrario, han reforzado de manera no intencional las estructuras y dinámicas de desigualdad existentes.

Es importante contemplar asimismo si las intervenciones recogen información de seguimiento pertinente y desglosada que permita a los entes de ejecución tomar decisiones pertinentes sobre el énfasis de las actividades o los objetivos y los recursos asignados para el desarrollo inclusivo.

Desafíos de evaluar la eficiencia y cómo enfrentarlos

La forma adecuada de aplicar el criterio de eficiencia dependerá completamente de la naturaleza de la intervención en sí y diferirá según el proyecto, programa, programa de país o política. El siguiente ejemplo del Ministerio de Relaciones Exteriores neerlandés muestra una aplicación del criterio en las políticas del sector de agua, donde la eficiencia se entiende en gran medida en términos de la coordinación y los aspectos prácticos de la planificación y la creación de asociaciones como parte de una compleja red de relaciones (Recuadro 4.11).

El Cuadro 4.4 identifica algunos de los desafíos centrales de la evaluación de la eficiencia y sugiere

formas de responder a ellos tanto para los evaluadores como para los gestores de evaluaciones.

Una decisión primordial es si deben usarse mediciones económicas tradicionales y las herramientas asociadas, por ejemplo, el análisis del costo-beneficio, de las tasas de retorno o de costo-efectividad, o las evaluaciones comparativas (*benchmarking*), entre otras, para evaluar la eficiencia⁴. Esto depende del propósito de la evaluación, la intervención y los resultados intencionales, la viabilidad, la disponibilidad de datos y recursos, y los destinatarios previstos.

La utilidad de las distintas herramientas de análisis dependerá asimismo del enfoque utilizado en las etapas de diseño y aprobación en el seno de la institución pertinente, ya que esto tendrá importantes implicaciones para la disponibilidad de la información necesaria para acometer diferentes tipos de análisis. En algunos bancos multilaterales de desarrollo y en algunos programas de inversiones de capital en el sector público, se establecen pautas muy claras para el diagnóstico económico, social y ambiental *ex ante*, y cada vez en mayor medida para la aplicación de un análisis de género a dichos proyectos y programas. Para evaluar la eficiencia en la etapa de diagnóstico o de aprobación del proyecto (normalmente *ex ante*, es decir, previamente a la ejecución de la intervención), es conveniente utilizar las mismas herramientas que durante la evaluación. Las normas y pautas adoptadas por la institución también determinarán en parte qué datos e indicadores se encuentran a disposición del evaluador (por ejemplo, si las tasas de retorno se estimaron durante el diagnóstico económico, de existir uno, y si se identificaron opciones alternativas).

Cuadro 4.4. Desafíos de evaluar la eficiencia

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
Encontrar comparaciones adecuadas	Los evaluadores suelen tener dificultades para encontrar las comparaciones perfectas, de modo que utilizan las opciones de las que disponen. Sin embargo, tienen que ser cautelosos a la hora de hacer comparaciones y deberían llevar a cabo un análisis de sensibilidad. Por ejemplo, la eficiencia en situaciones frágiles y afectadas por conflictos solo se debe comparar con entornos difíciles parecidos.	Cuando los gestores de evaluaciones tienen acceso a datos o análisis que podrían arrojar información general de utilidad para los evaluadores, planificar la incorporación de estos recursos en la documentación que se entrega a los evaluadores.
Lidiar con los diferentes conceptos de eficiencia y las metodologías y herramientas asociadas	La literatura sobre los conceptos económicos de eficiencia es extensa y los evaluadores pueden consultarla para decidir qué metodologías y herramientas son más apropiadas de cara a la evaluación en cuestión. El documento de trabajo del BMZ contiene un excelente resumen al respecto (Palenberg, 2011 ^[13]).	
Falta de conocimientos especializados adecuados sobre el análisis de la eficiencia	Cerciorarse de que exista un entendimiento respecto de lo que se requiere para llevar a cabo un análisis de eficiencia y cómo incorporarlo en el proceso de evaluación (tanto en términos de enfoque como de conocimientos especializados).	Cerciorarse de que los términos de referencia de las evaluaciones especifiquen los conocimientos especializados que se necesitan sobre el análisis de la eficiencia.
Datos y mediciones	Los datos disponibles sobre los beneficios, los resultados y los costos son difíciles de obtener. Por ejemplo, las evaluaciones del impacto pueden proveer estimaciones cuantificadas de los cambios en los efectos directos de la intervención, a mientras que omiten en general las estimaciones sobre los costos y, desde luego, sobre los costos totales.	Cerciorarse de que los evaluadores tengan acceso a los datos necesarios para el análisis de eficiencia (por ejemplo, datos sobre el costo-beneficio de una intervención o una acción dada en términos de presupuesto, recursos humanos, tiempo, etc.).

Consideraciones de tiempo	Tanto en el diagnóstico económico como en la evaluación de la eficiencia, las decisiones entre los períodos de tiempo en que se invierten los recursos y se entregan los resultados son cruciales. Los resultados y los costos totales (por ejemplo, los costos ambientales) pueden acumularse a lo largo de muchos años, de modo tal que es posible que la eficiencia real durante el transcurso del programa o en el momento de la evaluación no refleje el panorama completo.	La eficiencia se ve afectada por una serie de decisiones tomadas durante el transcurso del programa o de la intervención, entre otras cosas el modo en que los gestores respondieron a los giros en las prioridades, anticiparon los riesgos o reaccionaron frente a ellos, lidiaron con los sobrecostos, etc.
---------------------------	--	--

Ejemplos de evaluaciones de la eficiencia

Esta sección incluye una muestra transversal de ejemplos que demuestran cómo se ha evaluado la eficiencia en iniciativas concretas de electrificación rural, subsidios a los insumos agrícolas y una importante cartera de programas de gestión del agua.

Recuadro 4.9. Evaluación de la eficiencia del apoyo del Banco Mundial para electrificación rural

Este estudio incidió en el programa de préstamos del Banco Mundial para apoyar iniciativas de electrificación rural en varios países entre 1980 y 2006, e incluyó entre otras cosas una revisión de cartera que abarcó 120 proyectos, 10 estudios de caso nacionales y diferentes análisis de estadísticas de encuestas vigentes (encuestas de energía y de hogares de 12 países).

Recurriendo a los voluminosos conjuntos de datos de los que dispone un banco multilateral de desarrollo de gran envergadura con un largo historial de trabajo en el sector, el estudio consideró la eficiencia desde diversos ángulos, entre otras cosas comparando los costos y los beneficios. La revisión de cartera se basa en estimaciones de las tasas de retorno de los proyectos de la cartera de proyectos del banco, ya que estos se incluyen rutinariamente en los informes finales validados de los proyectos de infraestructura. Yendo un paso más allá, los impactos en el bienestar de los hogares se calculan a partir de estimaciones del excedente del consumidor, lo que permite a los evaluadores valorar los beneficios de las intervenciones y los costos relacionados. La evaluación demuestra un enfoque de evaluación del impacto, elementos de equidad (haber llegado a la población pobre) y análisis de las compensaciones entre sostenibilidad, eficiencia económica y asignativa, impacto general e impacto por grupos meta.

Fuente: GEI (2008^[14]), *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits*, <https://openknowledge.worldbank.org/handle/10986/6519>

Recuadro 4.10. Evaluación de la eficiencia de los subsidios para insumos agrícolas en el África subsahariana

Este estudio evaluó los subsidios otorgados para insumos agrícolas en Ghana, Malawi, Tanzania y Zambia, entre otras cosas los llamados “subsidios inteligentes”, que están especialmente diseñados para maximizar los efectos al menor costo posible.

Evaluación de la eficiencia

Dada la naturaleza del tema, la eficiencia económica (utilizada aquí en el sentido de si los subsidios valen su costo en diversos aspectos) es un aspecto central de este estudio.

El estudio se basa en estimaciones de cambios en la productividad agrícola, para luego analizar si son proporcionales a los costos, en el sentido de generar un retorno razonable, y qué factores, como la selección de los beneficiarios, fueron determinantes. Por ejemplo, en el caso de Zambia el estudio señala que:

El mayor uso de insumos agrícolas parece tener efectos considerables en la producción de maíz, pero la producción adicional conlleva un precio muy alto. Las mejores estimaciones disponibles sugieren que sería más económico importar maíz para el consumo en el medio urbano, en vez de incrementar la producción en Zambia a través del Programa de Apoyo a los Fertilizantes (ZFSP, por sus siglas en inglés) en su forma actual. Los subsidios para insumos aún podrían justificarse si las aparentes ineficiencias se compensasen con consideraciones de equidad o un desarrollo sostenible a largo plazo del sector de insumos. Sin embargo, dichos efectos son dudosos.

Como se puede ver, si bien examina la eficiencia, el estudio también considera naturalmente otros criterios, incluyendo algunas dimensiones de la sostenibilidad (por ejemplo, qué estrategia de salida se emplea) y el impacto o la eficacia de los subsidios, así como el impacto general.

Fuente: Baltzer y Hansen (2011^[15]), *Agricultural input subsidies in Sub-Saharan Africa: Evaluation Study*, <http://www.oecd.org/derec/49231998.pdf>

Recuadro 4.11. Evaluación de la eficiencia a nivel de políticas: la política de asistencia al desarrollo neerlandesa para mejorar la gestión del agua

Esta revisión evaluó la política de asistencia al desarrollo para mejorar la gestión integrada de los recursos hídricos impulsada por el Ministerio de Relaciones Exteriores de los Países Bajos durante el período comprendido entre 2006 y 2016. Se trata de una prioridad de política que había cobrado una importancia cada vez mayor en los Países Bajos e incluía diversas asociaciones con organizaciones multilaterales, gobiernos nacionales, ONG y el sector privado. Los dos principales criterios utilizados en la evaluación fueron la eficacia y la eficiencia de la política.

Evaluación de la eficiencia

La pregunta de evaluación específicamente relacionada con la eficiencia se formuló de la siguiente manera:

- ¿Se organizó y operativizó adecuadamente la ejecución de la política para apoyar el logro de los resultados centrales previstos en relación con los grupos de usuarios del agua, la calidad técnica y el mantenimiento de la infraestructura física?

Para responder a esta pregunta, los evaluadores consideraron la medida en que el logro de los resultados centrales se desvió del plan (en términos de costos y plazos) y por qué. Evaluaron los resultados a todos los niveles de la cadena de resultados, entre ellos a nivel de calidad, efectos directos y sostenibilidad. La evaluación identificó, por ejemplo, que la planificación participativa de la gestión del riego en ocasiones sufrió demoras y sobrecostos debido a la complejidad social y técnica de los problemas locales, que agravó los factores políticos e institucionales, generando lentitud en el desempeño y sobrecostos.

Fuente: IOB (2017^[16]), *Tackling major water challenges: Policy review of Dutch development aid policy for improved water management*, <https://english.iob-evaluatie.nl/publications/policy-review/2017/12/01/418-%E2%80%93-iob-%E2%80%93-policy-review-of-dutch-development-aid-policy-for-improved-water-management-2006-2016-%E2%80%93-tackling-major-water-challenges>

Impacto

Definición de impacto:

Impacto: ¿Qué diferencia genera la intervención?

Medida en que la intervención ha tenido o se prevé que tenga importantes efectos de carácter general, positivos o negativos, intencionales o no.

Nota: El impacto tiene en cuenta la trascendencia última y los efectos potencialmente transformadores de la intervención. Trata de identificar efectos sociales, ambientales y económicos de la intervención a largo plazo y con un alcance mayor que los ya identificados siguiendo el criterio de la eficacia. Además de los resultados inmediatos, el impacto trata de detectar consecuencias indirectas, secundarias y potenciales de la intervención. Para ello, se examinan los cambios holísticos y permanentes en los sistemas o las normas, así como los potenciales efectos en el bienestar de las personas, los derechos humanos, la igualdad de género y el entorno.

¿Qué es el impacto y por qué es importante?

El criterio de impacto invita a considerar la gran pregunta: “¿Por qué importa?”. Aquí es donde se consideran los efectos de desarrollo últimos de una intervención, donde los evaluadores analizan si la intervención creó o no un cambio que de verdad es importante para la gente. Representa una oportunidad para adoptar una perspectiva más amplia y una visión más integral. De hecho, es fácil dejarse absorber por los aspectos del día a día de una intervención en particular y simplemente ceñirse al marco de referencia de quienes vienen trabajando en ella. El criterio de impacto desafía a los evaluadores a ir más allá e indagar qué cambios se han logrado y a quién favorecen. La importancia de ello ha sido destacada por el Manual de Evaluación de la Agencia Sueca de Cooperación Internacional para el Desarrollo (Asdi) (Molund y Schill, 2004^[17]):

“El criterio de impacto ofrece un importante correctivo para lo que de otra manera podría devenir en un interés muy estrecho en las intenciones de quienes planifican y gestionan las intervenciones, y el correspondiente descuido de las perspectivas de los grupos meta y otras partes interesadas principales.”

Aunque el uso del término “impacto” es habitual, es importante advertir que a menudo existe confusión en torno a cómo se interpreta, lo que podría afectar la forma en que las partes interesadas entienden la evaluación. En primer lugar, en un contexto político, puede usarse de manera informal para referirse a los “resultados” en el sentido más amplio, abarcando tanto la eficacia como el impacto tal como están definidos aquí, así como otros aspectos del desempeño. En segundo lugar, en los últimos años, con frecuencia ha sido confundido con el término “evaluación del impacto”, que se refiere a las metodologías específicas para establecer las relaciones causales estadísticamente significativas entre la intervención y los efectos observados⁵. Cuando se utiliza de esta forma, el impacto puede referirse a los resultados en cualquier punto de la cadena de resultados, incluidos los productos, y en casi todos los casos hace referencia a los efectos *deseados*. Por estas razones, es importante aclarar con las partes interesadas desde un inicio cómo entienden ellos el término y explicar cómo se está usando en el contexto de la evaluación para examinar los efectos de carácter general.

Las preguntas que podría abordar el criterio de impacto son las siguientes:

- ¿Ha generado la intervención un cambio importante en las vidas de los beneficiarios previstos?

- ¿De qué manera la intervención generó efectos de carácter general (por ejemplo, cambios en las normas o los sistemas)?
- ¿Se beneficiaron de la intervención todos los grupos meta por igual, incluidos los más desaventajados y vulnerables?
- ¿Es transformadora la intervención, es decir, genera cambios perdurables en las normas, entre ellas las normas de género, y los sistemas, intencionalmente o no?
- ¿Está la intervención llevando a otros cambios, como resultados “escalables” o “replicables”?
- ¿De qué manera contribuirá la intervención a cambiar la sociedad para mejor?

Cómo entender el impacto: elementos de análisis

La definición de impacto engloba algunos conceptos centrales: los efectos de carácter general, la trascendencia, los impactos diferenciados, los efectos no intencionales y el cambio transformador.

Trascendencia

El criterio de impacto recoge la pregunta “¿Por qué importa?” de una evaluación. Examina la trascendencia de la intervención y sus resultados de carácter general, es decir, cuánto significó para los involucrados.

La definición está pensada para incentivar a los evaluadores a considerar diferentes perspectivas, de acuerdo a las circunstancias. El evaluador debe pensar detenidamente en el contexto, así como en las necesidades y las prioridades de los beneficiarios previstos de la intervención, los objetivos de política acordados de las instituciones pertinentes y la naturaleza de la intervención en sí. Este elemento de análisis puede aplicarse también al evaluar los resultados no intencionales o imprevistos de una intervención.

Al valorar la “trascendencia”⁶, los evaluadores deben ser conscientes de la importancia de considerar diferentes perspectivas y utilizar un enfoque sistemático basado en las necesidades de las partes interesadas. Además, deben tomar medidas para evitar que sus sesgos y juicios de valor (implícitos) afecten su evaluación de la trascendencia de la intervención.

Impacto diferenciado

Para acatar el mandato de los ODS de “no dejar a nadie atrás” y salvaguardar los derechos humanos, incluida la igualdad de género, es importante evaluar los impactos diferenciados entre los grupos. Los impactos positivos generales pueden esconder efectos negativos o distribucionales significativos. Es de suma importancia tener esto en cuenta en la etapa de diseño de la evaluación o, de hecho, en la etapa de diseño de la intervención, para garantizar que sea posible monitorear y después evaluar el impacto por grupo meta. Para ello, se requiere una planificación temprana del diseño y de la evaluación con el fin de garantizar la disponibilidad de datos desagregados donde sea factible, y puede suponer asimismo examinar una serie de parámetros relacionados con la exclusión y la inclusión. Supondrá un análisis granular de los datos desglosados cuando estén disponibles.

Efectos no intencionales

Los evaluadores deben considerar si una intervención tiene efectos no intencionales o imprevistos. Este análisis debe incluir el grado en que los impactos se buscaron o previeron al diseñar la intervención. Los efectos no intencionales pueden ser positivos o negativos. Cuando son positivos, los evaluadores deben considerar su trascendencia general y si existe un margen para la innovación o para aumentar o replicar el impacto positivo en otras intervenciones. Los evaluadores deben prestar especial atención a los

impactos negativos, en especial aquellos que pueden ser significativos, por ejemplo, entre otros, los impactos ambientales o los impactos no intencionales para grupos vulnerables.

Cambio transformador

La definición entiende por cambio transformador “los cambios integrales y perdurables en los sistemas o las normas”. Un cambio transformador puede interpretarse como un cambio que incide en las causas subyacentes o los impulsores sistémicos de la pobreza, las desigualdades, la exclusión y los daños ambientales, y la Agenda 2030 reconoce que se necesita un cambio transformador para lograr los objetivos de desarrollo sostenible. Cada vez es más común que las intervenciones deseen contribuir a un cambio transformador, y se insta cada vez en mayor medida a los evaluadores a responder preguntas sobre los efectos en las normas y los sistemas (sociales, económicos o políticos) al evaluar el criterio de impacto. Por ejemplo, una evaluación podría examinar la medida en que los roles de género tradicionales han sido transformados en algún sentido (ver el Recuadro 4.12).

Conexiones con otros criterios

Al igual que los demás criterios, el criterio de impacto interactúa conceptualmente con otros criterios:

- **Impacto y eficacia:** El impacto y la eficacia consideran ambos qué resultados generó la intervención, los efectos en la población beneficiaria y los resultados no intencionales. La diferencia entre ambos criterios dependerá en gran medida de cómo se diseñó la intervención y sus objetivos específicos. La eficacia generalmente incidirá en el logro de los objetivos declarados (a cualquier nivel de la cadena de resultados), mientras que el impacto se centrará invariablemente en los efectos de carácter general que, de lo contrario, no se hubieran tenido automáticamente en cuenta (al no haber sido incluidos como objetivos). Otra forma de entender la distinción entre ambos es que, con el tiempo, muchas intervenciones pueden ser calificadas de eficaces, pese a que no llegan al cambio transformador o de carácter general deseado. Si no se evalúa el impacto, estas tendencias pasarán inadvertidas. La formulación de los resultados previstos de una intervención variará debido a que las diferentes instituciones a menudo tienen distintos requisitos para definir la cadena de resultados. En algunas instituciones se exige que las intervenciones estén vinculadas a objetivos de carácter general, mientras que en otras se consideran solamente los efectos inmediatos. Para la aplicación de los dos criterios, es recomendable que las instituciones se cercioren de que los requisitos que rigen el diseño de la intervención están redactados de manera clara y coherente. Cuando en las intervenciones de menor envergadura, por ejemplo, los proyectos, no se formule de manera habitual la vinculación con los objetivos de carácter general, es importante que la política de evaluación disponga la obligatoriedad de realizar un análisis del impacto; en caso contrario, no se evaluarán los impactos de carácter general (ver el Recuadro 4.12 para conocer un ejemplo).
- **Impacto y coherencia:** El hecho de que la medición del impacto suponga adoptar una perspectiva integral significa que este criterio es naturalmente compatible con los factores de la coherencia, ya que los efectos generados por una intervención dependen casi invariablemente de otras intervenciones, objetivos de política, compensaciones y sistemas dentro de los cuales tiene lugar la intervención. Un ejemplo serían los programas de cooperación para el desarrollo que impulsan el fortalecimiento de los sistemas de salud pública en los países en desarrollo, donde el impacto se ve afectado no solamente por el programa, sino también por las políticas nacionales o internacionales de precios y las normas que regulan los productos farmacéuticos o la contratación del personal de salud.
- **Impacto y eficiencia:** Como se observó al tratar el criterio de eficiencia, para poder examinar la eficiencia en el sentido más amplio, los evaluadores tienen que adoptar una visión integral de los

resultados generados (por ejemplo, impacto y sostenibilidad) y cotejar los resultados con los recursos.

- Impacto y sostenibilidad: El impacto y la sostenibilidad consideran ambos en alguna medida si los resultados perdurarán en el tiempo. El impacto se ocupa especialmente de la dimensión temporal, en el sentido de examinar los cambios transformadores (que son perdurables por naturaleza). La sostenibilidad se enfoca en la continuación de los beneficios. Como criterio, la sostenibilidad es más amplia porque evalúa las condiciones para el logro de la sostenibilidad y las conexiones entre sostenibilidad económica, social y ambiental de una intervención.

Cómo integrar la inclusión

Cambio transformador, impacto diferenciado y trascendencia: todos estos conceptos están intrínsecamente ligados a la inclusión. Es importante establecer aquí qué impacto ha tenido lugar y para quién. ¿Hubo auténticas contribuciones a una transformación de los sistemas de opresión y podría esto significar un cambio perdurable para los grupos marginados y vulnerables? Los evaluadores deben tratar de establecer la medida en que estos impactos no intencionales obedecen a desigualdades estructurales en los sistemas más amplios y el impacto de las intervenciones sobre dichos sistemas.

La versión revisada del impacto enfatiza los resultados de una intervención a nivel general, entre ellos los impactos sociales y económicos a largo plazo. Supone un cambio transformador en los sistemas y las normas, incluidos, como consta en la definición, “los potenciales efectos en el bienestar de las personas, los derechos humanos, la igualdad de género y el entorno”. El impacto es el espacio donde los evaluadores pueden ver el panorama más completo, es decir, la medida en que una intervención contribuye y suma a un cambio transformador, la equidad, los derechos humanos y el empoderamiento.

Desafíos de evaluar el impacto y cómo enfrentarlos

De los seis criterios, el impacto es el que a menudo puede ser más difícil de evaluar y comprender. A continuación, presentamos un resumen de los principales desafíos que representa, así como algunas sugerencias sobre cómo proceder.

Como regla general, la evaluación de impacto tradicionalmente necesita de más recursos y considerablemente más datos que los demás criterios, y debe incorporarse al diseño de una evaluación únicamente si se tiene la certeza de que se dispondrá de dichos recursos o que su obtención para este fin será viable. Por otra parte, dado que evalúa si la intervención ha generado una diferencia y cómo, se trata del área que a menudo recibirá mayor atención por parte de los usuarios. En consecuencia, la inversión de tiempo y esfuerzo para incluirla en la evaluación es a menudo justificable.

Un desafío adicional para la evaluación de impacto tiene que ver con los plazos establecidos por las instituciones que ordenan las evaluaciones. Dichos plazos a menudo coinciden con la fecha de cierre de la intervención. Este requisito debería flexibilizarse para que el impacto pueda ser analizado durante un período de tiempo más largo, lo que permitirá comprender mejor los impactos que solo se hacen evidentes una vez finalizada la intervención.

El cuadro siguiente identifica varios de los desafíos centrales que supone la evaluación del impacto y sugiere alternativas para enfrentarlos, tanto para los evaluadores como para los gestores de evaluaciones.

Cuadro 4.5. Desafíos de evaluar el impacto

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
Claridad en torno a qué significa impacto como concepto para las partes interesadas y para los usuarios pertinentes de la evaluación	Acordar desde un inicio qué definición se está usando, ya sea la terminología del CAD de la OCDE o un concepto diferente, y ceñirse sostenidamente a ello. Verificar regularmente en las conversaciones con las partes interesadas si su interpretación del término ha variado.	
Claridad en torno a cuál fue el impacto previsto para esta intervención y cómo debía generarse	Depende del diseño original de la intervención y de la disponibilidad de un marco de resultados claramente definido, acompañado de una teoría del cambio. Si estos no existieran, una evaluación del impacto basada en la teoría (algo muy conveniente) sería difícil de obtener, de modo tal que la teoría deberá reconstruirse, lo cual incluye consultar con grupos meta específicos. Algunas intervenciones tienen tradicionalmente objetivos que son muy amplios y se han inflado más allá del ámbito que la intervención puede conseguir.	
Identificar el grado en que la intervención genera el impacto	<p>Mapear las rutas de contribución o atribución (según el caso) de la intervención a los resultados cuando estas sean menos difíciles de rastrear o explicitar a los niveles más altos. El uso de metodologías como el análisis de la contribución puede ser de utilidad aquí⁷.</p> <p>Consultar en guías o metodologías específicas cómo medir el impacto, explorar opciones y tomar decisiones deliberadas y realistas, sustentadas en el propósito de la evaluación, el contexto, las preguntas por responder, la disponibilidad de datos y los recursos disponibles para la evaluación y su viabilidad. Una pregunta clave es si las metodologías se limitarán a evaluar la magnitud del impacto de la intervención o si también identificarán por qué ocurrió y cómo.</p>	Verificar que los recursos disponibles para la evaluación y las competencias de los involucrados (en el uso de las metodologías) sean suficientes para evaluar el impacto.
Disponibilidad de datos, incluyendo valores de base e indicadores	Para evaluar los efectos de carácter general, los datos pueden representar un desafío mayor, entre otras cosas debido a la falta de valores de base e indicadores. Se trata en parte de un tema de evaluabilidad que se debe tomar en consideración al inicio de la evaluación, pero que también fundamenta la elección de los métodos adecuados para basarse en los diferentes tipos de datos cualitativos y cuantitativos de la mejor forma posible.	Establecer la disponibilidad de datos durante el proceso de elaboración de los términos de referencia de la evaluación.
La intervención tiene efectos no intencionales o imprevistos significativos.	Con frecuencia es difícil recabar información sobre los resultados no intencionales o imprevistos, ya que a menudo no se mencionan en los marcos y la recolección de datos de seguimiento. Para recoger lo que resulta más importante para los beneficiarios y otras personas (positiva o negativamente) afectadas por la intervención, los evaluadores deben permanecer abiertos a ir más allá de los documentos formales y los datos previstos del proyecto. La trascendencia relativa de los efectos probablemente variará para las diferentes partes interesadas.	Facilitar a los evaluadores el espacio necesario para escuchar a las personas afectadas (positiva o negativamente) por la intervención y apoyarlas a la hora de llamar la atención sobre los efectos no intencionales que son importantes para ellas.

Ejemplos de evaluaciones de impacto

Esta sección contiene una muestra transversal de ejemplos que revelan cómo se ha evaluado el impacto de las intervenciones de: empoderamiento familiar, construcción de la paz, apoyo a investigaciones, planificación del uso del suelo y violencia contra la mujer y la niña.

Recuadro 4.12. Evaluación de la eficacia y el impacto del programa Más Familias en Acción de Colombia

Una evaluación de los impactos en las operaciones y el bienestar comunitario del programa Más Familias en Acción de Colombia ofrece un interesante ejemplo de cómo entender e identificar un cambio transformador a nivel general, entre otras cosas cómo se evalúan ciertos aspectos bajo los criterios de eficacia e impacto desde una óptica ligeramente distinta. Esta evaluación examinó la función de las *madres líderes*, mujeres capacitadas para ayudar a sacar adelante el programa, que participan empoderando y orientando a sus comunidades.

- En las operaciones: La evaluación examinó la función de las madres líderes en apoyar la eficacia del programa, destacando el rol primordial de las mujeres en la ejecución y entrega de las actividades.
- A nivel general: Adicionalmente, la evaluación identificó que la participación de las madres líderes en el programa generó efectos positivos en el empoderamiento de las mujeres a nivel individual, así como efectos potencialmente transformadores en las comunidades más amplias, al modificar las percepciones del rol de las mujeres en la sociedad. Estos cambios de carácter general en las normas pueden recogerse mediante el criterio de impacto.

Fuente: Economía Urbana e Ipsos (2019^[18]), *Evaluación de operaciones del programa "más familias en acción" y de resultados del componente de bienestar comunitario*, https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion_MFA_Informe_Resultados.pdf

Recuadro 4.13. Evaluación del impacto del apoyo de Asdi a la construcción de la paz en países afectados por conflictos y en situaciones posconflicto

Esta evaluación es un interesante ejemplo del uso del criterio de impacto para evaluar una intervención compleja durante un largo período de tiempo en contextos de fragilidad. El impacto se evaluó a nivel del efecto general de los esfuerzos de construcción de la paz de Asdi en términos de resultados directos, indirectos, intencionales y no intencionales, positivos y negativos.

Esta evaluación recogió los impactos de más largo aliento adoptando una perspectiva de muy largo plazo e incidiendo en el trabajo de construcción de la paz en Somalia, Ruanda, Bosnia y Guatemala desde principios de la década de 1990 hasta fines de la década de 2010. Además del impacto general, la evaluación se enfocó especialmente en las mujeres y el impacto en los sectores marginados. La evaluación utilizó la metodología del "cambio más significativo" para evaluar las contribuciones de la intervención.

Fuente: Bryld (2019^[19]), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf

Recuadro 4.14. Valoración del impacto en la evaluación de una política estratégica

En una evaluación del apoyo y los estudios que llevaron a cabo 21 autoridades danesas involucradas en 40 proyectos en 18 países asociados, un hallazgo clave relacionado con el criterio de impacto fue la trascendencia de los efectos de carácter general en lo que respecta al empoderamiento de las instituciones asociadas. La evaluación concluyó que uno de los logros más importantes de la intervención fue empoderar a las autoridades asociadas más allá de los objetivos formales del proyecto. Esto hace referencia a la capacidad de Dinamarca para interactuar y conectar con actores clave del sector público y privado, la sociedad civil y el sector del desarrollo, e influir en la generación de efectos directos más amplios como consecuencia de las políticas y los programas. Dichos logros son mutuamente beneficiosos para ambos asociados, pero en su mayor parte pasan desapercibidos durante el seguimiento de la intervención. No obstante, se trata de un efecto directo fundamental del programa para apreciar su valor añadido en toda su integridad.

Fuente: PEM Consult (2020^[20]), *Evaluation of the Danish Strategic Sector Cooperation*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationID=CBE77158-1D4D-46E3-A81F-9B918218FAFF

Recuadro 4.15. Evaluación de impacto, difusión y aumento del impacto de un enfoque integral de planificación del uso de la tierra en las Filipinas

Al examinar el impacto de una intervención sobre el uso de la tierra, esta evaluación consideró los efectos a mediano y largo plazo de la intervención en función de indicadores ambientales y socioeconómicos.

Se valió de un enfoque basado en la teoría de la intervención y un diseño de metodología mixta, sustentado en los datos de panel de una encuesta multiniveles, entrevistas cualitativas, debates con grupos focales, una revisión de la literatura, un análisis de los documentos de planificación del uso de la tierra e informaciones y datos geográficos. Los evaluadores realizaron una reconstrucción pormenorizada de la teoría del cambio de la intervención.

La pieza medular de esta evaluación es una valoración rigurosa del impacto, que midió y cuantificó los efectos en cinco áreas de impacto: desde mejoras en las estructuras y condiciones administrativas de las autoridades de planificación; pasando por mejoras en la gestión de los recursos naturales, medidas y actividades de gestión del riesgo de desastres y funcionamiento de la gobernanza local; hasta mejoras en el bienestar de la población afectada.

Los datos de panel se recabaron en dos momentos (2012 y 2016) en 3.000 hogares, distribuidos en 300 *barangays* o distritos de 100 municipios pertenecientes a 11 provincias de la región de las Bisayas. La evaluación incluyó hogares que habían recibido asistencia de la agencia alemana de cooperación al desarrollo GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit) y hogares que no habían recibido asistencia. La metodología de la evaluación de impacto también se basó en un diseño cuasi experimental sustentado en el enfoque de emparejamiento por puntuación de propensión para identificar “gemelos estadísticos” a partir de varias docenas de características de los municipios, *barangays* y hogares.

Fuente: Leppert et al. (2018^[21]), *Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines: From Development Cooperation to National Policies*, https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2018_Philippinen/DEval_Executive_Summary_Philippinen_2018_EN.pdf

Recuadro 4.16. Evaluación del impacto del programa Empoderamiento de la Niña (GE, por sus siglas en inglés) en Liberia

Esta evaluación realizada por International Rescue Committee ofrece un ejemplo concreto de la evaluación del impacto de una intervención multifacética para combatir la violencia contra la mujer y la niña, denominada “Empoderamiento de la niña: mentoría y transferencias en efectivo para promover el bienestar de las adolescentes en Liberia”. Esta intervención se llevó a cabo en el condado de Limba en 2016. Representa un valioso ejemplo de cómo los resultados específicos evaluados bajo los criterios tanto de eficacia como de impacto varían dependiendo de cómo se definieron los objetivos de la intervención.

Evaluación del impacto

El propósito específico del estudio es evaluar el impacto de un tipo de intervención en particular, el combate de la violencia contra las mujeres y las niñas en Liberia, que en este caso añadió un componente de transferencias en efectivo a una intervención que combina fortalecimiento económico con perspectivas de género transformadoras para las adolescentes desde muy temprana edad.

La evaluación recabó información detallada (de informaciones de base y finales) sobre la magnitud de la violencia contra las adolescentes y las niñas, y los efectos específicos de la intervención, tanto durante la intervención como al cabo de un año tras su finalización. Los efectos directos se midieron utilizando la metodología de ensayo controlado aleatorizado por conglomerados a nivel de aldeas. El diseño de la investigación trataba de establecer: 1) el impacto general del programa, en comparación con un grupo contrafáctico (el grupo de control); y 2) la eficacia de añadir un pago en efectivo como incentivo para la participación de las niñas (“programa GE+”), específicamente para evaluar si otorgar incentivos en efectivo a las niñas tiene beneficios de protección y empoderamiento que reducen el riesgo de violencia sexual, posiblemente mediante el incremento de la asistencia de las niñas a las sesiones del programa.

Los efectos en el empoderamiento y los cambios en las normas y las actitudes sociales frente a la posición de las niñas en la sociedad y a la violencia serían ejemplos de impactos transformadores.

Fuente: Hallman *et al.* (2018^[22]), *Girl Empower Impact Evaluation: Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing in Liberia*, <https://www.rescue.org/sites/default/files/document/4346/girlempowerimpacetevaluation-finalreport.pdf>

Sostenibilidad

Definición de sostenibilidad:

Sostenibilidad: ¿Se mantendrán los beneficios?

Medida en la que se mantienen o pueden mantenerse los beneficios netos de la intervención.

Nota: Incluye un examen de las capacidades financiera, económica, social, ambiental e institucional de los sistemas necesarios para que los beneficios netos se mantengan a lo largo del tiempo. Conlleva la realización de análisis de resiliencia, riesgos y posibles compensaciones. Dependiendo de la cronología de la evaluación, podría comportar el análisis del flujo real de beneficios netos o la estimación de la probabilidad de que los beneficios netos se mantengan a medio y largo plazo.

¿Qué es la sostenibilidad y por qué es importante?

La evaluación de la sostenibilidad permite a la comunidad evaluadora determinar si los beneficios de una intervención se mantendrán a nivel financiero, económico, social y ambiental. Si bien el concepto subyacente de continuación de los beneficios se mantiene, el criterio actualizado no solo es más conciso, sino que también es de ámbito más amplio con respecto a la definición anterior⁸. La sostenibilidad supone varios elementos de análisis (financieros, económicos, sociales y ambientales) y debe prestarse atención a la interacción entre ellos.

Es posible que surja una confusión entre sostenibilidad en el sentido de continuación de los resultados y sostenibilidad ambiental o el uso de los recursos teniendo en cuenta a las generaciones futuras. Aunque la sostenibilidad ambiental es una preocupación (y puede examinarse en función de varios criterios, entre ellos la pertinencia, la coherencia, el impacto y la sostenibilidad), el significado principal de los criterios no gira en torno a la sostenibilidad ambiental como tal; al describir la sostenibilidad, los evaluadores deben tener claridad respecto de cómo están interpretando el criterio.

La sostenibilidad debe considerarse en cada punto de la cadena de resultados y del ciclo de proyectos de una intervención. Los evaluadores deben reflexionar asimismo acerca de la sostenibilidad en términos de resiliencia y adaptación en entornos dinámicos y complejos. Esto incluye la sostenibilidad de los insumos (financieros o de otro tipo) una vez finalizada la intervención, así como la sostenibilidad de los impactos en el contexto más amplio de la intervención. Por ejemplo, una evaluación podría valorar si una intervención consideró las capacidades de los asociados y generó apropiación al inicio del período de ejecución, y si había la voluntad y la capacidad necesarias para mantener el financiamiento una vez finalizada la intervención. En general, los evaluadores pueden examinar las condiciones de sostenibilidad que fueron creadas o no en el diseño de la intervención y por las actividades de la intervención, y si se efectuaron adaptaciones cuando fue necesario.

Los evaluadores no solo deben examinar la sostenibilidad desde el punto de vista de los donantes y los flujos financieros externos. Las personas que ordenan una evaluación deben considerar también llevar a cabo una evaluación de la sostenibilidad antes del inicio de una intervención o mientras el financiamiento o las actividades se encuentran en curso. Al evaluar la sostenibilidad, los evaluadores deben: 1) tener en cuenta los beneficios netos, es decir, el valor general de los beneficios sostenidos de la intervención, incluidos los posibles costos permanentes; y 2) analizar las potenciales compensaciones y la resiliencia de las capacidades o los sistemas en los que se sostiene la continuación de los beneficios. Puede haber compensaciones, por ejemplo, entre la sostenibilidad fiscal de los beneficios y la sostenibilidad política (la continuación del apoyo político).

La evaluación de sostenibilidad ofrece información de gran valor sobre la continuación o la posible continuación de los beneficios netos de una intervención en el mediano a largo plazo, lo que ha demostrado ser un gran desafío en la práctica de acuerdo a lo establecido por diversos metaanálisis. Por ejemplo, en sectores como el agua y el saneamiento, o en tipos de intervenciones como las iniciativas de desarrollo impulsadas por la comunidad, los beneficios a menudo se desdibujan al cabo de un tiempo. La función de la evaluación aquí puede ser escudriñar los supuestos de la teoría del cambio con relación al logro de la sostenibilidad (Mansuri y Rao, 2013^[23]; White, Menon y Waddington, 2018^[24]).

Si todos estos aspectos se consideran cuidadosamente en una evaluación, pueden llevar a importantes revelaciones acerca de la forma en que las intervenciones pueden planear y ejecutar actividades para generar un cambio que asegure el desarrollo sostenible en el futuro. Las lecciones pueden poner de relieve la potencial escalabilidad de las medidas para garantizar la sostenibilidad de la intervención en el contexto actual o su potencial capacidad de réplica en otros contextos.

Un aspecto central de la sostenibilidad es la planificación del cierre de la intervención. Las evaluaciones deben evaluar si se ha elaborado y puesto en práctica una estrategia de cierre apropiada, lo cual aseguraría la continuación de los efectos positivos, por ejemplo, entre otras, las consideraciones financieras y de capacidad. Si la evaluación se está llevando a cabo *ex post*, el evaluador puede examinar además si la estrategia de cierre programada se ejecutó debidamente de cara a asegurar la continuación de los efectos positivos tal como se había previsto, dejando al mismo tiempo un margen para los cambios que podrían producirse en las condiciones del contexto, tal como se muestra en los ejemplos más adelante.

Un recurso de utilidad para entender de forma más adecuada la sostenibilidad y responder mejor a los desafíos comunes que esta plantea es una metaevaluación (Noltze, Euler y Verspohl, 2018^[25]) y una síntesis evaluativa (Noltze, Euler y Verspohl, 2018^[26]) realizadas hace poco por el Instituto Alemán de Evaluación de la Cooperación para el Desarrollo (DEval). Los dos estudios hacen referencia a los diferentes elementos de análisis de la sostenibilidad que se pueden examinar en una evaluación. La metaevaluación aboga fuertemente a favor de una evaluación de sostenibilidad que incorpore los principios de los ODS y destaque las áreas en que puede añadir valor a una evaluación. Esto incluye un análisis de la medida en que la identificación y la evaluación de los efectos no intencionales de un proyecto y de las interacciones o compensaciones entre las diferentes dimensiones de sostenibilidad pueden respaldar el aprendizaje y la rendición de cuentas al aplicar el criterio de sostenibilidad.

Cómo entender la sostenibilidad: elementos de análisis

Entender la definición de sostenibilidad significa entender los siguientes componentes: entorno propicio, continuación de los efectos positivos, y riesgos y compensaciones.

Cómo generar un entorno propicio para el desarrollo sostenible

Las evaluaciones pueden analizar las múltiples formas en que una intervención ha contribuido a promover un entorno propicio para el desarrollo, como la medida en que la intervención impulsó el fortalecimiento de los sistemas, las instituciones o las capacidades necesarios para apoyar la futura actividad humanitaria o de desarrollo. Esto incentiva a las evaluaciones a evaluar la capacidad del asociado para el desarrollo que se ha creado o fortalecido como resultado de la intervención, así como la resiliencia construida para lidiar con los cambios y las perturbaciones externas. Esto asegurará que los beneficios netos, tal como se describieron anteriormente, continúen de cara al futuro.

Las contribuciones de una intervención para promover un entorno propicio para el desarrollo podrían incluir, por ejemplo: capacidades fortalecidas (a nivel individual, comunitario o institucional); apropiación o voluntad política mejoradas; mayores compromisos financieros o presupuestarios a nivel nacional (y

subnacional, donde proceda); cambios en las políticas o estrategias; reformas legislativas; reformas institucionales; reformas a la gobernanza; mayor rendición de cuentas con relación al gasto público; o mejores procesos de consulta pública en la planificación del desarrollo.

Continuación de los efectos positivos: sostenibilidad real y proyectada

La sostenibilidad puede evaluarse a lo largo de distintos períodos. Los evaluadores pueden evaluar tanto la sostenibilidad real (es decir, la continuación de los beneficios netos creados por la intervención que ya son evidentes) como la sostenibilidad futura (es decir, los beneficios netos para las partes interesadas clave que probablemente continuarán en el futuro). Los evaluadores deben poner especial atención en seleccionar enfoques de evaluación apropiados para evaluar la sostenibilidad real o futura, dependiendo de la cronología de la evaluación y del horizonte temporal de los beneficios previstos. Muchos cambios de carácter general tardarán años o decenios en manifestarse por completo.

Para evaluar la sostenibilidad real, el evaluador puede examinar la medida en que los efectos positivos generados por la intervención hayan continuado probadamente para las partes interesadas clave, entre ellas los beneficiarios previstos, una vez finalizada la intervención. Los evaluadores pueden examinar además si se han identificado y previsto oportunidades para apoyar la continuación de los efectos positivos de la intervención y si se han hecho planes para aprovecharlos y cómo, así como las posibles barreras que pudieran haber interferido en la continuación de los efectos positivos. Esto puede respaldar los hallazgos en el sentido de que la intervención habría demostrado tener capacidad de adaptación cuando fue necesario.

El examen de la sostenibilidad futura supone un enfoque ligeramente distinto. Una evaluación que examina las posibilidades de sostenibilidad futura evaluaría las probabilidades de que los efectos positivos de la intervención, ya sean previstos o actuales, continúen, generalmente asumiendo que las condiciones actuales se mantienen. La evaluación necesitará valorar la estabilidad y la permanencia relativa de cualquier efecto positivo ya generado, así como las condiciones para su continuación, como la sostenibilidad institucional, la sostenibilidad económica y financiera, la sostenibilidad ambiental, la sostenibilidad política y la sostenibilidad cultural.

Riesgos y posibles compensaciones

La evaluación de la sostenibilidad requiere de un examen no solamente de las probabilidades de continuación de los efectos positivos de la intervención, sino también de los potenciales riesgos y los costos permanentes asociados a una intervención. En consecuencia, los gestores de evaluaciones deben evaluar los factores que podrían favorecer la sostenibilidad de los beneficios netos en el tiempo, así como los factores que podrían inhibirla. El examen de los riesgos asociados a la sostenibilidad de una intervención puede incluir una valoración de los factores contextuales (positivos o negativos, identificables o previsibles) susceptibles de influir en la durabilidad de los resultados de la intervención.

Esto pone sobre el tapete adicionalmente el tema de las compensaciones, un importante elemento de los criterios revisados. La evaluación de las compensaciones asociadas a una intervención supone también un examen de la compensación entre el impacto instantáneo y los potenciales efectos o costos a más largo plazo, así como las compensaciones entre los aspectos financieros, económicos, sociales y ambientales. Por ejemplo, una evaluación puede establecer que una intervención propició el crecimiento económico, pero que dicho crecimiento es insostenible debido a los graves costos ambientales que pueden impactar negativamente en el crecimiento económico a más largo plazo. Esto se ajusta a la definición de desarrollo sostenible de los ODS y amplía el ámbito para que las evaluaciones examinen la sostenibilidad más allá de las probabilidades de continuación de los efectos positivos de una intervención.

Conexiones con otros criterios

La sostenibilidad se encuentra estrechamente relacionada con los demás criterios.

- La sostenibilidad se relaciona con la pertinencia en el sentido de que el nivel de pertinencia para las partes interesadas clave constituye un factor primordial para la apropiación y la aceptación de los eventuales beneficios, lo que a su vez impulsa la sostenibilidad.
- De igual manera, la coherencia puede ofrecer datos valiosos sobre la sostenibilidad, ya que examina otras intervenciones en un contexto dado, las mismas que podrían apoyar o socavar los beneficios de la intervención con el correr del tiempo.
- Eficacia e impacto: La evaluación de la continuación de los resultados de una intervención depende, en primer lugar, de que se hayan demostrado logros (eficacia) y, en segundo lugar, de que se hayan demostrado efectos de carácter general (impacto). Por lo tanto, la eficacia y el impacto pueden verse como criterios absolutos para evaluar la sostenibilidad dado que, si su análisis no demuestra que la intervención generó productos, efectos directos o impactos, no habrá beneficios claros que sostener. El Recuadro 4.19 ofrece un ejemplo de evaluación del impacto y la sostenibilidad en conjunto. El análisis de las sinergias entre impacto, eficacia y sostenibilidad a través de una evaluación de las condiciones que son suficientes y necesarias para la continuación de los resultados una vez finalizada la intervención permite a los evaluadores explorar la eficacia y el impacto a más largo plazo.
- Las preocupaciones relacionadas con la eficiencia también pueden socavar la sostenibilidad de los beneficios. Por ejemplo, cuando los costos a corto plazo son los que determinan las decisiones, en detrimento de los efectos a largo plazo, el efecto puede ser reducir la sostenibilidad.

Cómo integrar la inclusión

La definición revisada de sostenibilidad y su nota correspondiente hacen hincapié en las dimensiones “financiera, económica, social, ambiental e institucional” de la sostenibilidad, y cómo estas contribuyen a los beneficios actuales y a largo plazo de los resultados de la intervención. Los evaluadores deben considerar la medida en que se han tomado previsiones para la continuación de los beneficios para los diferentes grupos de beneficiarios y, si la evaluación se está realizando de manera *ex post*, la medida en que esto es manifiesto para los diferentes grupos. Aquí el énfasis debe recaer en el principio de “no dejar a nadie atrás” y en cómo los grupos marginados experimentan los beneficios positivos actuales, así como las compensaciones que pueden surgir entre los diferentes grupos.

Además, es pertinente que los evaluadores tengan en cuenta la medida en que la intervención ha construido un entorno habilitador para el desarrollo inclusivo y equitativo, atacando los problemas sistémicos subyacentes (“tratando la enfermedad, y no solo los síntomas”) de acuerdo a los criterios tanto de impacto como de sostenibilidad. Los aspectos de apropiación y empoderamiento de género son importantes aquí. La sostenibilidad de los sistemas requiere de mayores capacidades, de modo tal que los evaluadores deben establecer de quién son las capacidades que se crearon y cómo se relaciona esto con los sistemas y estructuras de desigualdad existentes. ¿Existen tanto capacidad como compromiso de parte de los diferentes grupos de partes interesadas para crear y mantener un entorno propicio para la igualdad de género y el empoderamiento de la mujer en el mediano a largo plazo? De no ser así, ¿cuáles son las barreras y se encuentran estas dentro del ámbito de la intervención?

Desafíos de la evaluación de sostenibilidad y cómo enfrentarlos

El cuadro que presentamos a continuación identifica los desafíos de la evaluación de sostenibilidad en términos de cronología, inexistencia de efectos positivos y otros factores que afectan la sostenibilidad, y

sugiere opciones para hacerles frente, tanto para los evaluadores como para los gestores de evaluaciones.

Cuadro 4.6. Desafíos de la evaluación de sostenibilidad

Desafío	Cómo enfrentarlo: evaluadores	Cómo enfrentarlo: gestores de evaluaciones
La intervención no ha generado los resultados previstos (evaluados bajo el criterio de eficacia) ni beneficios positivos no intencionales, ni ha realizado contribuciones al impacto.	Si la intervención no ha generado beneficios, el análisis de la sostenibilidad será redundante. Los evaluadores deben tener claridad y mantener una actitud abierta respecto de las limitaciones de la evaluación de la sostenibilidad y considerar enfocarse en otros criterios.	Las personas que gestionan la evaluación deben considerar hacer ajustes al ámbito de esta y redirigir los recursos hacia otros criterios, incluyendo indagar por qué no se generaron beneficios.
Cronología de la evaluación	Las evaluaciones a menudo tendrán lugar durante el ciclo de vida de una intervención, o es posible que se efectúen en un momento en que la sostenibilidad no es todavía evidente. Por eso, los evaluadores deben considerar enfocarse en si existen o no condiciones para asegurar la sostenibilidad si la intervención aún está en proceso de ejecución o ha concluido recientemente. Deben entregar una evaluación sólida de la sostenibilidad futura, identificando los factores en el entorno operativo que podrían fomentar la sostenibilidad. En ese sentido, el análisis de género es clave y generará información pertinente y de utilidad acerca de la consistencia de los cambios estructurales generados y su efecto en la reducción de las desigualdades.	Si la evaluación se está llevando a cabo mientras la intervención todavía está en marcha, los gestores deben tener claridad acerca de la medida en que priorizan las consideraciones de sostenibilidad y el propósito de analizar este aspecto. Por ejemplo, ¿se considerarán los hallazgos en la planificación del cierre de la intervención o de la transición?
Evaluación de los factores susceptibles de influir en la sostenibilidad	Los factores contextuales que apoyan (o socavan) la sostenibilidad futura pueden evaluarse desde un punto de vista cualitativo o cuantitativo. Incluyen, entre otros, la apropiación y el compromiso de las partes interesadas, la capacidad de absorción, la voluntad política y la disponibilidad de recursos a nivel nacional.	Los gestores pueden ayudar a identificar a las partes interesadas a ser entrevistadas o encuestadas.

Ejemplos de evaluaciones de sostenibilidad

Esta sección contiene ejemplos de un proyecto de desarrollo rural en Afganistán, uno de apoyo al presupuesto general y otro de salud materna, para demostrar cómo se ha evaluado la sostenibilidad.

Recuadro 4.17. Examen de la sostenibilidad en una evaluación *ex post* del Proyecto de Desarrollo Rural Intercomunitario del Japón en Afganistán

El Proyecto de Desarrollo Rural Intercomunitario se llevó a cabo como parte de la Iniciativa Ogata de la Agencia de Cooperación Internacional del Japón (JICA) con miras a establecer un modelo de desarrollo rural liderado por la comunidad que contribuyera a generar efectos directos en la reconstrucción y el desarrollo rural en las provincias de Balkh, Bamyan y Kandahar, en Afganistán. Una evaluación *ex post* examinó la sostenibilidad en relación con los esfuerzos de reconstrucción tras la guerra civil que azotó Afganistán.

Habiendo establecido que la intervención se había ejecutado con éxito y había cumplido con todos sus objetivos, la evaluación examinó después la sostenibilidad de los beneficios. Al valorar la sostenibilidad, la evaluación analizó si existían las condiciones políticas, institucionales, técnicas y financieras necesarias para asegurar la continuación de los efectos positivos. Se incluyó entre estas consideraciones la capacidad y la voluntad del gobierno afgano de continuar con la ejecución de la iniciativa al término del proyecto.

La evaluación halló que la intervención había establecido sistemas de seguimiento en colaboración con otros donantes, así como una fundación fiscal financiada por los donantes, complementando y desarrollando capacidades en el sistema nacional del país afectado por el conflicto. Además, se confirmó que el gobierno afgano estaba decidido a continuar con la iniciativa de desarrollo rural liderado por la comunidad, creando conglomerados de Consejos de Desarrollo Comunitario con apoyo de los donantes. Asimismo, el proyecto fue incorporado al Programa Nacional de Solidaridad. En consecuencia, la evaluación concluyó que sí existían las condiciones políticas, institucionales, técnicas y financieras requeridas para garantizar la continuación de los efectos positivos del proyecto, y que no había riesgos aparentes.

Fuente: Watanabe (2016^[27]), *Ex-Post Evaluation of Technical Co-operation Project "Inter-Communal Rural Development Project"*, https://www2.jica.go.jp/en/evaluation/pdf/2015_0603847_4.pdf

Recuadro 4.18. Lecciones de sostenibilidad de la evaluación del cese del apoyo al presupuesto general en Malaui, Ruanda, Uganda y Zambia

Este estudio, un interesante ejemplo del uso minucioso del criterio de sostenibilidad en una evaluación, examinó la medida en que la intervención se preparó adecuadamente para la continuación de los efectos positivos tras el término del apoyo al presupuesto y si impulsó la creación de un entorno propicio para el desarrollo. Este estudio se llevó a cabo después del retiro generalizado de las iniciativas de apoyo al presupuesto general, que se inició cerca de 2010, debido a una serie de escándalos en los países receptores y a presiones en los países donantes para llamar a rendir cuentas a los países receptores. Este desarrollo fue posterior a un giro a favor del apoyo al presupuesto que se dio a principios de la década de 2000 en reacción a la Declaración de París de 2005 sobre la Eficacia de la Ayuda. El apoyo al presupuesto era ampliamente considerado como una opción eficaz y sostenible para impulsar resultados de desarrollo. Una serie de evaluaciones y estudios de síntesis generó evidencia de que el apoyo al presupuesto había contribuido positivamente a generar importantes efectos directos en el desarrollo, por ejemplo, el aumento del gasto en los pobres, el incremento de la matriculación escolar y las mejoras en la administración de las finanzas públicas.

La evaluación se basó en un diseño comparativo de estudios de caso de Malaui, Ruanda, Uganda y Zambia para determinar si se habían aplicado estrategias adecuadas para asegurar la continuación de los efectos positivos asociados al apoyo al presupuesto general después del cierre de dicha iniciativa. Destacó que, si bien el nivel total de la asistencia se mantuvo constante en la mayoría de los países, las estructuras creadas desaparecieron en su mayor parte a raíz de la decisión de retirar el apoyo. La mayoría de los logros positivos se vieron afectados negativamente o incluso revertidos. Se estableció que el financiamiento autónomo de proyectos individuales era la nueva modalidad de apoyo prevaleciente en las carteras de asistencia bilateral, y que el financiamiento conjunto se circunscribía a unos pocos sectores y programas. Esto había generado externalidades negativas y una alta fragmentación de la ayuda, lo que dificultó la creación de un sentido de apropiación nacional o la subsanación de problemas de gobernanza sistemáticos más amplios.

Fuente: Orth, Birsan y Gotz (2018^[28]), *The Future of Integrated Policy-Based Development Cooperation: Lessons from the Exit from General Budget Support in Malawi, Rwanda, Uganda and Zambia*, http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_EN_The%20Future%20of%20Integrated%20Policy-Based%20Development%20Cooperation..pdf

Recuadro 4.19. Evaluación del impacto y la sostenibilidad del programa de salud materna del DFID

Este examen efectuado por la Comisión Independiente para el Impacto de la Ayuda Oficial al Desarrollo (ICAI, por sus siglas en inglés) de los resultados del Departamento para el Desarrollo Internacional del Reino Unido (DFID) en mejorar la salud materna en diferentes países representa un ejemplo de cómo evaluar de manera conjunta el impacto y la sostenibilidad. Los exámenes de la ICAI se rigen por un enfoque específico, que en este caso prioriza tres grandes áreas: impacto, equidad y sostenibilidad. Las preguntas del examen incluyen las siguientes:

- Impacto: ¿En qué medida ha maximizado el DFID el impacto de sus programas sobre la salud materna? ¿En qué medida son sólidas sus afirmaciones respecto de sus resultados de salud materna a nivel mundial, por países y a nivel del programa? ¿En qué medida sus inversiones se hacen eco de la evidencia global de las cosas que funcionan para mejorar la salud materna? ¿En qué medida sus programas han tenido los efectos directos requeridos para mejorar la salud materna?
- Equidad: ¿En qué medida los programas seleccionaron como grupo meta a las mujeres marginadas y a las que resulta difícil llegar? ¿En qué medida el DFID está contribuyendo a fortalecer los sistemas de salud y a mejorar el acceso de las mujeres a ellos?
- Sostenibilidad: ¿En qué medida es probable que los resultados de salud materna del DFID demuestren ser sostenibles? ¿En qué medida ha apoyado el DFID el desarrollo y la ejecución de políticas e instituciones nacionales que garanticen la sostenibilidad de los resultados?

Equidad incluye aquí la evaluación directa de efectos específicos para los pobres y las mujeres jóvenes, en consonancia con la naturaleza de la intervención. Dichas preguntas relativas a los resultados diferenciados y a poblaciones específicas pueden enfocarse desde distintos ángulos al amparo de las nuevas definiciones de pertinencia, eficacia e impacto. La sostenibilidad se evalúa en términos de “en qué medida es probable” que los resultados se sostengan, dado que la sostenibilidad real (*ex post*) solo se podría observar durante un período de tiempo más largo (en el futuro).

La metodología en este caso consiste en realizar una evaluación independiente (ajustada a las funciones y el mandato de la ICAI) de las aseveraciones del DFID en relación con los resultados de su trabajo de mejora de la salud materna, incluyendo su trabajo de creación de modelos. Se sostiene en documentos, entrevistas y visitas de campo.

Fuente: ICAI (2018^[29]), *Assessing DFID's results in improving Maternal Health: An impact review*, <https://icai.independent.gov.uk/wp-content/uploads/ICAI-review-Assessing-DFIDs-results-in-improving-Maternal-Health-.pdf>

Referencias

- Arghiros, D. *et al.* (2017), *Making it Count: Lessons from Australian Electoral Assistance 2006-16*, Departamento de Relaciones Exteriores y Comercio del Gobierno de Australia, <http://www.oecd.org/derec/australia/australia-electoral-assistance-2006-2016.pdf> [fecha de consulta de la publicación original: 11 de enero de 2021]. [11]

- Baltzer, K. y H. Hansen (2011), *Agricultural Input Subsidies in Sub-Saharan Africa: Evaluation Study*, DANIDA, International Development Cooperation, Ministerio de Relaciones Exteriores de Dinamarca, <http://www.oecd.org/derec/49231998.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021). [15]
- Befani, B. y J. Mayne (2014), "Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation", *IDS Bulletin*, vol. 45/6, págs. 17 a 36, <http://dx.doi.org/10.1111/1759-5436.12110>. [31]
- Belcher, B. y M. Palenberg (2018), "Outcomes and Impacts of Development Interventions", *American Journal of Evaluation*, vol. 39/4, págs. 478 a 495, <http://dx.doi.org/10.1177/1098214018765698>. [34]
- Bryld, E. (2019), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, Asdi, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf (fecha de consulta de la publicación original: 11 de enero de 2021). [19]
- Chambers, R. et al. (2009), "Designing impact evaluations: different perspectives", núm. 4, 3ie, <https://www.3ieimpact.org/evidence-hub/publications/working-papers/designing-impact-evaluations-different-perspectives> (fecha de consulta de la publicación original: 11 de enero de 2021). [35]
- Economía Urbana e IPSOS (2019), *Evaluación de operaciones del programa "más familias en acción" y de resultados del componente de bienestar comunitario*, Departamento Nacional de Planeación, Bogotá D.C., https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion_MFA_Informe_Resultados.pdf (fecha de consulta de la publicación original: 15 de enero de 2021). [18]
- Eurecna Spa (2020), *Bolivia - Evaluation of Health Initiatives (2009-2020)*, Ministerio de Relaciones Exteriores y Cooperación Internacional de Italia, http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf (fecha de consulta de la publicación original: 11 de enero de 2021). [5]
- FAO (2020), *Evaluation of "Improving farmer livelihoods in the dry zone through improved livestock health, productivity and marketing"*, Organización de las Naciones Unidas para la Alimentación y la Agricultura, Roma, <http://www.fao.org/3/ca8463en/ca8463en.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021). [4]
- GEI (2008), *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits*, Banco Mundial, <http://dx.doi.org/10.1596/978-0-8213-7367-5>. [14]
- Gertler, P. et al. (2016), *La evaluación de impacto en la práctica: Segunda edición*, Grupo Banco Mundial, Washington D.C., <https://publications.iadb.org/en/impact-evaluation-practice-second-edition> (fecha de consulta de la publicación original: 12 de enero de 2021). [30]
- Global Affairs Canada (2019), *Evaluation of Natural Disaster Reconstruction Assistance in the Philippines, 2013-14 to 2018-19*, Global Affairs Canada, <https://www.international.gc.ca/gac-amc/publications/evaluation/2019/endra-earcn-philippines.aspx?lang=eng> (fecha de consulta de la publicación original: 12 de enero de 2021). [8]

- Hallman, K. et al. (2018), *Girl Empower Impact Evaluation: Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing in Liberia*, International Rescue Committee, <https://www.rescue.org/sites/default/files/document/4346/girlempowerimpacetevaluation-finalreport.pdf> (fecha de consulta de la publicación original: 12 de enero de 2021). [22]
- ICAI (2018), *Assessing DFID's Results in Improving Maternal Health: An Impact Review*, Comisión Independiente para el Impacto de la Ayuda del Reino Unido, <https://icai.independent.gov.uk/wp-content/uploads/ICAI-review-Assessing-DFIDs-results-in-improving-Maternal-Health-.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021). [29]
- IDD and Associates (2006), *A Joint Evaluation of General Budget Support Evaluation of General Budget Support: Synthesis Report*, <https://www.oecd.org/development/evaluation/dcdndep/37426676.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021). [2]
- IDEV (2018), *Cabo Verde: Evaluation of the Bank's Country Strategy and Program 2008–2017: Summary Report*, Departamento de Evaluación Independiente del Desarrollo, Banco Africano de Desarrollo, <http://www.oecd.org/derec/afdb/AfDB-2008-2017-cabo-verde-bank-strategy.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021). [12]
- IOB (2017), *Tackling major water challenges: Policy review of Dutch development aid policy for improved water management*, Departamento de Evaluación de Políticas y Operaciones, Ministerio de Relaciones Exteriores, Países Bajos, <https://english.iob-evaluatie.nl/documents/2017/12/01/dutch-development-aid-policy-for-improved-water-management> (fecha de consulta de la publicación original: 12 de enero de 2021). [16]
- Leeuw, F. y J. Vaessen (2009), *Impact Evaluations and Development: NoNIE Guidance on Impact Evaluation*, Red de Redes sobre Evaluación del Impacto, Washington D.C., <https://documents1.worldbank.org/curated/en/411821468313779505/pdf/574900WP0Box351BLIC10nonie1guidance.pdf> (fecha de consulta de la publicación original: 12 de enero de 2021). [33]
- Leppert, G. et al. (2018), *Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines: From Development Cooperation to National Policies*, Instituto Alemán de Evaluación del Desarrollo (DEval), Bonn, https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Berichte/2018_Philippinen/DEval_Executive_Summary_Philippinen_2018_EN.pdf (fecha de consulta de la publicación original: 12 de enero de 2021). [21]
- Mansuri, G. y V. Rao (2013), *Localizing Development*, Banco Mundial, <http://dx.doi.org/10.1596/978-0-8213-8256-1>. [23]
- Ministerio de Relaciones Exteriores de Eslovenia (2017), *Evaluation of Slovenia's Development Cooperation with Montenegro 2013-2016: Final Report*, Departamento de Cooperación para el Desarrollo de Eslovenia, <https://www.gov.si/assets/ministrstva/MZZ/Dokumenti/multilateral/razvojno-sodelovanje/Development-cooperation-with-Montenegro-evaluation-final-report.pdf> (fecha de consulta de la publicación original: 12 de enero de 2021). [9]

- Ministerio de Relaciones Exteriores del Japón (2019), *Japan ODA Evaluation Guidelines*, [6]
https://www.mofa.go.jp/policy/oda/evaluation/basic_documents/pdfs/guidelines11th.pdf (fecha de consulta de la publicación original: 18 de febrero de 2021).
- Molund, S. y G. Schill (2004), *Looking Back, Moving Forward Sida Evaluation Manual*, Sida, [17]
<https://www.oecd.org/derec/sweden/35141712.pdf> (fecha de consulta de la publicación original: 11 de enero de 2021).
- Noltze, M., M. Euler e I. Verspohl (2018), *Evaluation Synthesis of Sustainability in German Development Cooperation*, Instituto Alemán de Evaluación del Desarrollo (DEval), Bonn, [26]
http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_Evaluierungssynthese_EN_web.pdf (fecha de consulta de la publicación original: 12 de enero de 2021).
- Noltze, M., M. Euler e I. Verspohl (2018), *Meta-Evaluation of Sustainability in German Development Cooperation*, Instituto Alemán de Evaluación del Desarrollo (DEval), Bonn, [25]
http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_NH_Meta-Evaluierung_EN_web.pdf (fecha de consulta de la publicación original: 12 de enero de 2021).
- Norad (2018), *Evaluation of Norwegian Efforts to Ensure Policy Coherence for Development*, [7]
 Norad, <https://www.norad.no/en/publications/2018/evaluation-of-norwegian-efforts-to-ensure-policy-coherence-for-development/> (fecha de consulta de la publicación original: 11 de enero de 2021).
- OCDE (2002), *Evaluation and Aid Effectiveness No. 6, Glosario de los principales términos sobre evaluación y gestión basada en resultados*, OECD Publishing, París, [10]
<https://doi.org/10.1787/9789264034921-en-fr>.
- OCDE (2010), *Estándares de Calidad para la Evaluación del Desarrollo, Serie: Directrices y Referencias del CAD*, OECD Publishing, París, [1]
https://www.oecd.org/es/publications/estandares-de-calidad-para-la-evaluacion-del-desarrollo_9789264094949-es.html.
- OCDE (2011), *Evaluating Budget Support: Methodological Approach*, Red de Evaluación del Desarrollo del CAD, OECD Publishing, París, [3]
<https://www.oecd.org/dac/evaluation/dcdndep/Methodological%20approach%20BS%20evaluations%20Sept%202012%20with%20cover%20Thi.pdf> (fecha de consulta de la publicación original: 12 de enero de 2021).
- Orth, M., M. Birsan y G. Gotz (2018), *The Future of Integrated Policy-Based Development Cooperation: Lessons from the Exit from General Budget Support in Malawi, Rwanda, Uganda and Zambia*, Instituto Alemán de Evaluación del Desarrollo (DEval), Bonn, [28]
http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_EN_The%20Future%20of%20Integrated%20Policy-Based%20Development%20Cooperation..pdf (fecha de consulta de la publicación original: 12 de enero de 2021).
- Palenberg, M. (2011), *BMZ: Tools and Methods for Evaluating the Efficiency of Development Interventions | Managing for Sustainable Development Impact*, División de Evaluación, Ministerio Federal de Cooperación Económica y Desarrollo (BMZ), [13]
<http://www.managingforimpact.org/resource/bmz-tools-and-methods-evaluating-efficiency-development-interventions> (fecha de consulta de la publicación original: 12 de enero de 2021).

- PEM Consult (2020), *Evaluation of the Danish Strategic Sector Cooperation*, Departamento de Evaluación, Aprendizaje y Calidad, Ministerio de Relaciones Exteriores/DANIDA, Dinamarca, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationID=CBE77158-1D4D-46E3-A81F-9B918218FAFF (fecha de consulta de la publicación original: 12 de enero de 2021). [20]
- UNEG (2013), *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*, <http://www.uneval.org/document/detail/1433> (fecha de consulta de la publicación original: 12 de enero de 2021). [32]
- Watanabe, K. (2016), *Ex-Post Evaluation of Technical Cooperation Project "Inter-Communal Rural Development Project"*, JICA, https://www2.jica.go.jp/en/evaluation/pdf/2015_0603847_4.pdf (fecha de consulta de la publicación original: 12 de enero de 2021). [27]
- White, H., R. Menon y H. Waddington (2018), "Community-driven development: does it build social cohesion or infrastructure? A mixed-method evidence synthesis", núm. 30, 3ie, <https://www.3ieimpact.org/evidence-hub/publications/working-papers/community-driven-development-does-it-build-social-cohesion> (fecha de consulta de la publicación original: 12 de enero de 2021). [24]

Notas

¹ Los beneficiarios serían las "personas, grupos u organismos que se benefician, directa o indirectamente, de una intervención, sean o no objetivo de esta". Pueden utilizarse otros términos, como "titulares de derechos" o "personas afectadas".

² Puede haber casos en que dos instituciones participen en una intervención, una como asociado en la ejecución y la otra como entidad de financiamiento. Bajo esta configuración, es posible evaluar la coherencia interna desde la perspectiva de ambos asociados: el asociado de financiamiento y el asociado en la ejecución.

³ Una importante salvedad es que es mucho más probable que la evaluación de impacto sea de utilidad cuando está acompañada de un análisis de cómo se logró el impacto y qué se puede hacer para aumentarlo, a diferencia de cuando se considera como un simple ejercicio de contabilidad.

⁴ Ver un análisis detallado más completo de las opciones metodológicas para definir y evaluar la eficiencia en un documento de trabajo del BMZ elaborado por Palenberg (2011[24]), que identifica tres niveles: Nivel 0: describir y dar una opinión sobre algunos aspectos de una intervención de asistencia relacionados con la eficiencia. Nivel 1: identificar el potencial de mejora de la eficiencia de una intervención de ayuda. Esto permite formarse una imagen parcial de los procesos de ejecución, los costos de los insumos, la transformación de los insumos en productos o la transformación de los productos en efectos directos. Nivel 2: evaluar la eficiencia de una intervención de asistencia de forma tal que pueda contrastarse con intervenciones alternativas o de referencia. Se trata de un enfoque integral que comporta una estimación confiable de todos los beneficios y costos principales.

La evaluación de nivel 2 rara vez se aplica en la práctica. Incluso una organización con una enorme capacidad como el Banco Mundial señaló ya en el 2010 que se había producido un notable descenso en la aplicación del análisis de costo-beneficio en la etapa de diagnóstico, incluso en aquellos sectores

donde era más aplicable. El informe destacó los aspectos positivos de utilizar dicha profundidad de análisis (en términos de rigor y nexos con el desempeño subsiguiente del proyecto), pero también los desafíos involucrados (GEI, 2010[25]). La iniciativa del sector de electrificación rural (GEI, 2008[26]) es un ejemplo relativamente infrecuente de un análisis completo de nivel 2. Otras opciones disponibles para los evaluadores son los modelos multicriterio de la toma de decisiones. Es importante anotar asimismo que dichos tipos de análisis de la eficiencia son más propensos a ser aplicados en ciertos sectores, por ejemplo, la infraestructura, la salud y la agricultura. En el Recuadro 4.10 se proporciona un interesante ejemplo del sector agrícola que involucra subsidios a los insumos, sobre la base de un estudio de caso nacional que se sustenta en estimaciones económicas y datos de encuestas.

⁵ La literatura sobre este tema es abundante, pero ver, por ejemplo: UNEG (2013^[32]); Chambers *et al.* (2009^[35]); Leeuw y Vaessen (2009^[33]); Belcher y Palenberg (2018^[34]), y Gertler *et al.* (2016^[30]).

⁶ No confundir con trascendencia estadística, que a menudo surge en ciertos tipos de evaluaciones de impacto; ver Gertler *et al.* (2016^[30]), que analizan el concepto de cálculos de poder y otros conceptos técnicos relacionados en las evaluaciones de impacto cuantitativas.

⁷ Ver Befani y Mayne (2014^[31]).

⁸ La definición de sostenibilidad establecida en el Glosario de 2002 (OCDE, 2002^[10]) es: "Continuación de los beneficios de una intervención para el desarrollo después de concluida. Probabilidad de que continúen los beneficios en el largo plazo. Situación en la que las ventajas netas son susceptibles de resistir los riesgos con el correr del tiempo.

Aplicación minuciosa de los criterios de evaluación

Pertinencia, coherencia, eficacia, eficiencia, impacto y sostenibilidad son criterios de evaluación ampliamente utilizados, especialmente en la cooperación internacional para el desarrollo. Ayudan a determinar el mérito o el valor de diversas intervenciones, como estrategias, políticas, programas o proyectos. Esta guía tiene como objetivo ayudar a los evaluadores y a otros actores a comprender mejor estos criterios y mejorar su aplicación. Comienza describiendo qué son y cómo se pretende que se utilicen. A continuación, se explican las definiciones y los conceptos que sustentan cada criterio. Finalmente, se incluyen ejemplos que ofrecen al lector ideas concretas para su uso.

Los criterios fueron establecidos originalmente a principios de los años 2000 por la Red de Evaluación del Desarrollo (EvalNet) del Comité de Ayuda al Desarrollo (CAD) de la OCDE. Aunque se han utilizado ampliamente en evaluaciones y más allá, este documento es el primero en explorar de manera exhaustiva los conceptos en detalle, explicar su uso previsto y ofrecer orientaciones prácticas. Recoge el pensamiento actual y las mejores prácticas en evaluación, basándose en las aportaciones de expertos internacionales en evaluación de EvalNet y de otros ámbitos.



IMPRESA ISBN 978-92-64-37258-0
PDF ISBN 978-92-64-43340-3

