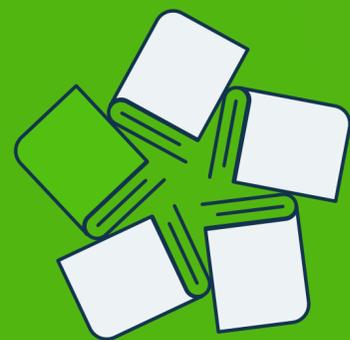




CAJA DE HERRAMIENTAS
PARA EVALUAR
INICIATIVAS EN EDUCACIÓN



MATERIAL GUIADO



EVALUACIÓN

Evaluación de resultados, impacto y costo-efectividad

➔ Sobre la evaluación	4
➔ Evaluación de resultados	11
➔ Evaluación de impacto	18
➔ Análisis de costo-efectividad	26

La última etapa del análisis consiste en la evaluación de efectos del programa una vez este haya sido implementado y/o se encuentre en el proceso de finalización. Esta evaluación suele comprender tres componentes: **i)** la evaluación de resultados, **ii)** la evaluación de impacto y **iii)** el análisis costo-efectividad.

Esta etapa tiene como objetivo responder cinco preguntas fundamentales sobre el programa:

01 ¿En qué medida se han cumplido los objetivos establecidos?

02 ¿Son los resultados observados atribuibles al programa?

03 ¿Existen atributos individuales que moderen los efectos del programa?

04 Dados los efectos del programa, ¿tiene este un costo razonable que permita su sostenimiento y/o ampliación?

05 ¿Qué componentes del programa se pueden mejorar?

Responder estas preguntas no solo sirve como evaluación de lo sucedido, sino que proporciona información valiosa para la continuidad del proyecto y su aplicación en diversos contextos.

Sobre la evaluación

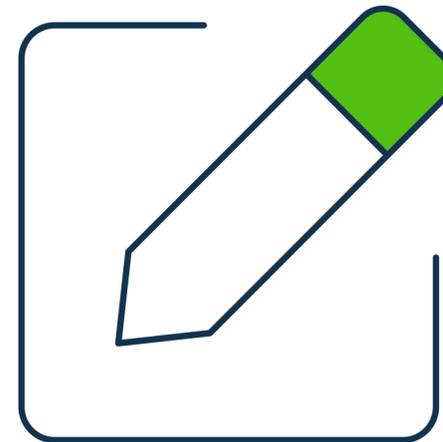
¿Qué es la evaluación?

Las evaluaciones son apreciaciones periódicas y objetivas de un proyecto, política o programa planeado o en curso. También constituyen un proceso integral de observación, recolección, medición, análisis sistemático e interpretación de información acerca del desempeño y efectividad de un programa, así como de comunicación de información, que provee insumos para la formulación de políticas y toma de decisiones¹.

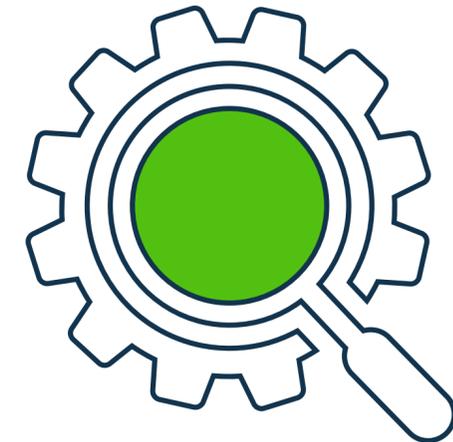
Ejemplos de preguntas para la toma de decisiones pueden ser:



¿Cómo asignar los recursos?



¿Se debe crear un nuevo programa o modificar uno existente?



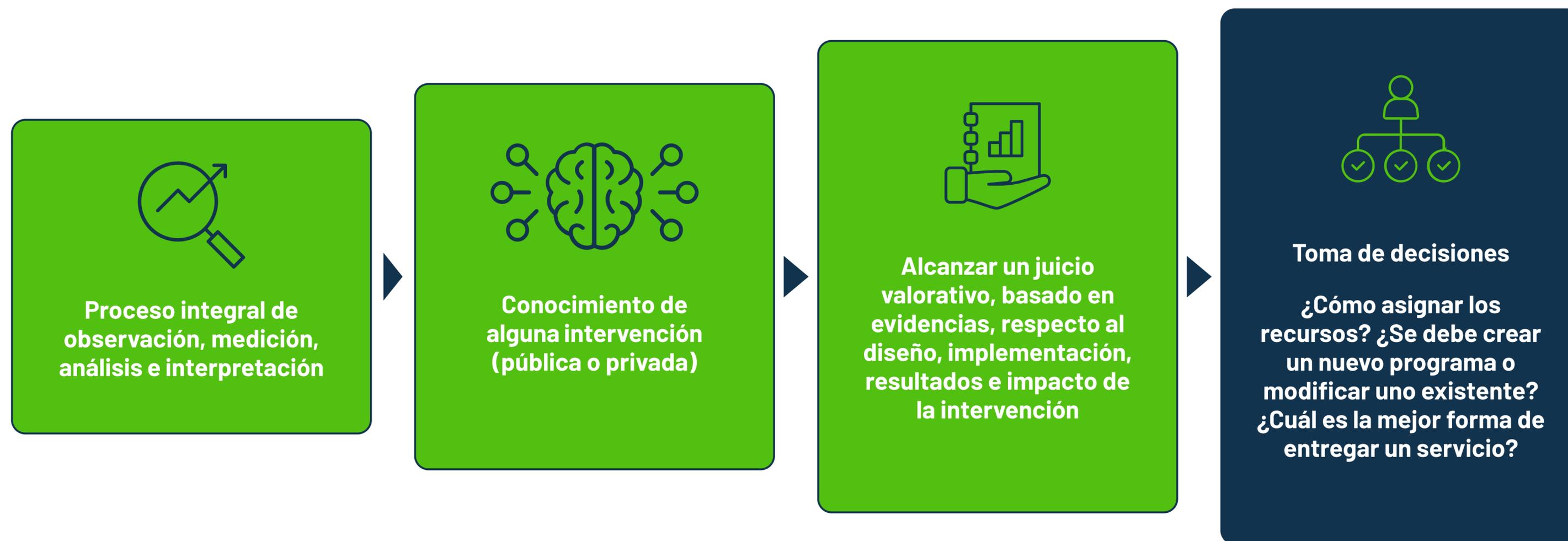
¿Cuál es la mejor forma de entregar un servicio?

¹ Gertler et al., 2011.

Con una evaluación adecuada, es posible responder a estas cuestiones.

Existen distintos tipos de evaluación, cuya diferencia radica en la evaluación de los diferentes componentes del programa. Algunos evalúan el diseño, otros los resultados y algunos ambos componentes.

Con la evaluación se conoce en profundidad una intervención, sea esta pública o privada, lo que permite alcanzar un juicio valorativo, basado en evidencias, respecto de su diseño, implementación, resultados e impacto. Por lo tanto, la evaluación es esencial para lograr el objetivo de formular políticas basadas en evidencia empírica.



¿Para qué sirve evaluar?

La evaluación orienta la toma de decisiones hacia la asignación eficiente y efectiva de recursos y ayuda a identificar obstáculos.



En cuanto a la **rendición de cuentas, transparencia y negociación**, sirve para:

- Conocer cómo se están utilizando los recursos y si se están logrando los objetivos planeados.
- Orientar decisiones de inversión.



En cuanto a la **mejora del diseño y operación de programas y proyectos de inversión**, la evaluación sirve para:

- Generar mayor conocimiento sobre el tipo de inversiones necesarias para obtener mejores resultados.
- Determinar si lo que se está haciendo funciona y los mecanismos que llevan a este resultado.
- Optimizar el proceso de operación de un programa o proyecto.

Factores a tener en cuenta

En la evaluación es importante tener en cuenta las cuatro etapas previas del trabajo: diagnóstico, diseño, seguimiento y monitoreo, e implementación. Esto permite verificar si en cada una se han cumplido los objetivos y se han abordado las preguntas que deben resolverse en cada etapa, que se recogen a continuación.

 **Figura 1.** Objetivo y preguntas de las primeras cuatro etapas



Diagnóstico

En la etapa de diagnóstico se identifica y plantea el problema principal, junto con sus causas y efectos, con respecto a la situación en cuestión. Lo anterior se realiza mediante un exhaustivo estudio sobre el estado del arte, y se ilustra por medio de un árbol de problemas.

Preguntas que se resuelven en esta etapa:

¿Cuál es el problema en cuestión?

¿Se han identificado problemáticas similares en otros contextos? ¿Cuáles fueron sus causas y efectos? ¿Cómo las solucionaron? ¿Qué podemos aprender y/o destacar?

¿Cuáles son las causas inmediatas del problema en cuestión? ¿Qué factores, sin ser causas directas, contribuyen a agravar la situación-problema?

¿Cuáles son las consecuencias directas del problema identificado?

Dé clic aquí para
más detalles de
esta etapa



Diseño

En la etapa de diseño se identifican los objetivos, la cobertura, las posibles alternativas de solución a la problemática, mediante un análisis de impacto, factibilidad y viabilidad, y se realiza el estudio técnico del proyecto. Para esto, es fundamental construir la teoría de cambio y la matriz de marco lógico.

Preguntas que se resuelven en esta etapa:

¿Cuáles son las metas y objetivos del proyecto?

¿Cuál es la población objetivo?

¿Cuál es el área de estudio y de influencia?

¿Cuáles son las posibles alternativas de solución?

¿Qué tan viables son las alternativas de solución?

¿Cuál es el valor agregado o la transformación que genera el proyecto?

¿Cuáles son los productos y actividades del proyecto?

¿Cuáles son los resultados a corto y largo plazo del proyecto?

Dé clic aquí para
más detalles de
esta etapa





Seguimiento y monitoreo

El seguimiento y monitoreo es el análisis continuo de la ejecución del programa para establecer el grado en que las actividades, procesos y metas se cumplen de acuerdo con lo planificado.

Preguntas que se resuelven en esta etapa:

¿Mi proyecto está cumpliendo con el progreso en la ejecución?

¿Cómo construir un plan de supervisión de los procesos y la operación de mi programa?

Dé clic aquí para
más detalles de
esta etapa



Implementación

La implementación responde al tipo de preguntas normativas, es decir, compara lo que efectivamente se está llevando a cabo con lo que debería estar pasando. Se evalúan actividades y si se están alcanzando los objetivos planeados.

Preguntas que se resuelven en esta etapa:

¿Mi proyecto está cumpliendo con lo necesario para el logro de las metas y objetivos?

¿Mi proyecto está cumpliendo la implementación de lo que se planificó?

Dé clic aquí para
más detalles de
esta etapa

Para identificar si cumplió con estos factores en el diseño de su programa, por favor diligencie la [encuesta de autoevaluación](#).

Tipos de evaluación

Se tratan aquí dos tipos de evaluación:

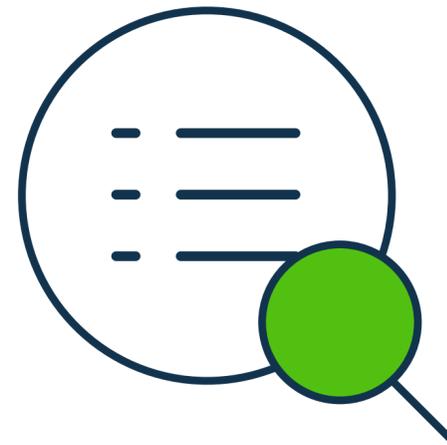


Evaluación de resultados



Evaluación de impacto

En esta etapa también se trata el análisis de costo-efectividad, como se indicó al inicio, aunque no sea este un tipo de evaluación estrictamente hablando.



Evaluación de resultados

Este tipo de evaluaciones responde a **preguntas descriptivas**, aquellas acerca de lo que se está llevando a cabo. Estas preguntas pueden involucrar procesos, condiciones, relaciones organizacionales y perspectivas de los socios o accionistas del proyecto.



Evaluación de impacto

Este tipo de evaluaciones responde a **preguntas de causalidad**, es decir, en qué medida el programa o proyecto afecta directamente el resultado planteado al que se quiere llegar.

Evaluación de resultados

¿Qué es una evaluación de resultados?

La evaluación de resultados se enfoca en los efectos brutos del programa, es decir, en la medida en que se cumplen los objetivos específicos del programa o proyecto sin descontar la influencia de factores externos a la intervención. Permite medir y hacer seguimiento de los resultados del programa, tanto de corto como de largo plazo.

En este tipo de evaluación es necesario diferenciar entre el producto o servicio entregado y los resultados. Por ejemplo, para un programa de educación sexual mediante talleres de formación, los productos serían los talleres de educación sexual, mientras que el resultado serían los cambios en indicadores como la tasa de embarazo adolescente o la prevalencia de enfermedades de transmisión sexual.

La evaluación de resultados permite identificar problemas o limitaciones que se presentaron en la ejecución del programa o proyecto, que impidieron el logro de los objetivos en el tiempo previsto y con los costos estimados.

La evaluación de resultados examina:

A

Los cambios en las condiciones de los beneficiarios después de un determinado tiempo de estar expuestos a la intervención.

B

Los factores que afectan el resultado dado en los beneficiarios expuestos.

C

Las contribuciones generadas por la organización en pro de un mayor resultado.

D

Las contribuciones generadas por las alianzas con socios para modificar el resultado².

² PNUD, 2002.



¿Cómo se realiza una evaluación de resultados?

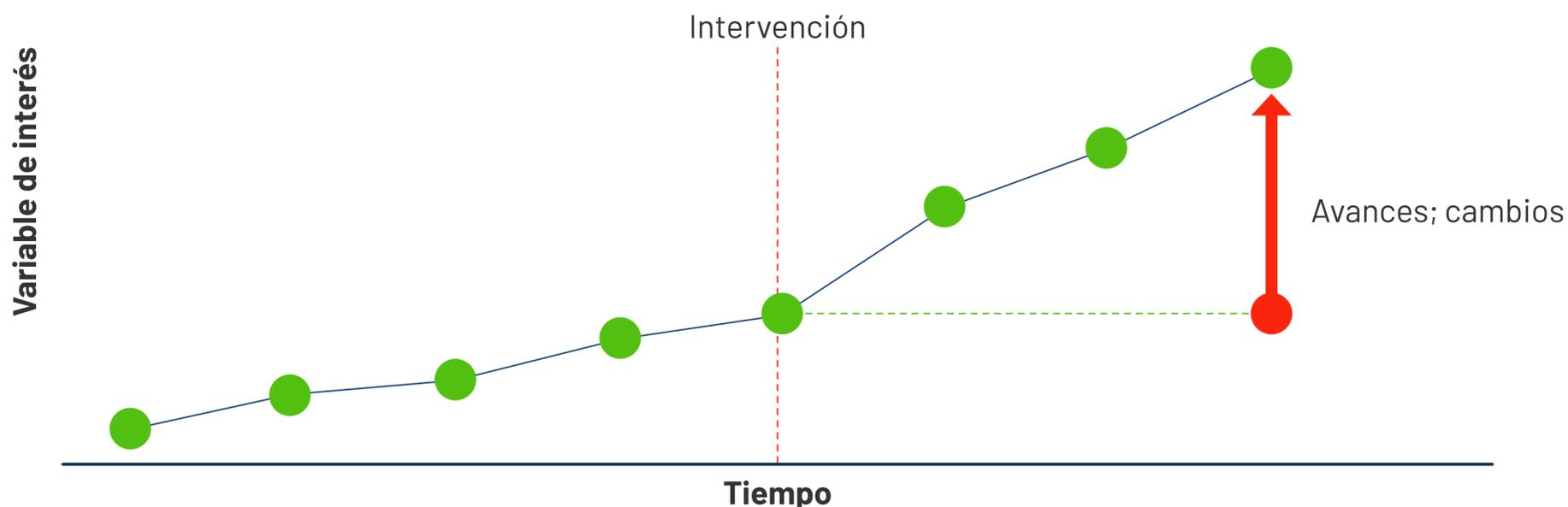
Para llevar a cabo una evaluación de resultados, es esencial medir los cambios percibidos por los beneficiarios y beneficiarias a lo largo del tiempo.

La medición de los resultados del programa se realiza comparando el indicador de interés antes (línea base) y después (seguimiento) de la intervención.

Para recolectar la línea base se puede recabar la información en el periodo inicial (t_0), o bien preguntar en retrospectiva en un momento dado (t_1) sobre las condiciones que se tenían al inicio. Para recolectar la línea de seguimiento se investiga cómo han cambiado las condiciones de los beneficiarios después de un determinado tiempo (t_2) de estar expuestos al programa.

En situaciones en las que: **i)** no se cuenta con información sobre la línea de base o información sobre los avances y el estado del resultado que normalmente hubiera sido recolectada mediante el seguimiento del resultado, o **ii)** los indicadores no cumplan con los criterios requeridos o no hayan sido sometidos a un seguimiento permanente, se recomienda consultar otras fuentes de información

 **Figura 2.** Gráfica modelo de medición de resultados



Fuente: elaboración propia.

como la documentación del proyecto junto con fuentes de información alternas para recolectar una línea de base y un conjunto de indicadores basados en dichas fuentes.

Adicionalmente, un resultado puede contar con varias dimensiones que no pueden ser reunidas en un único indicador o en un juicio de valor coherente. En estos

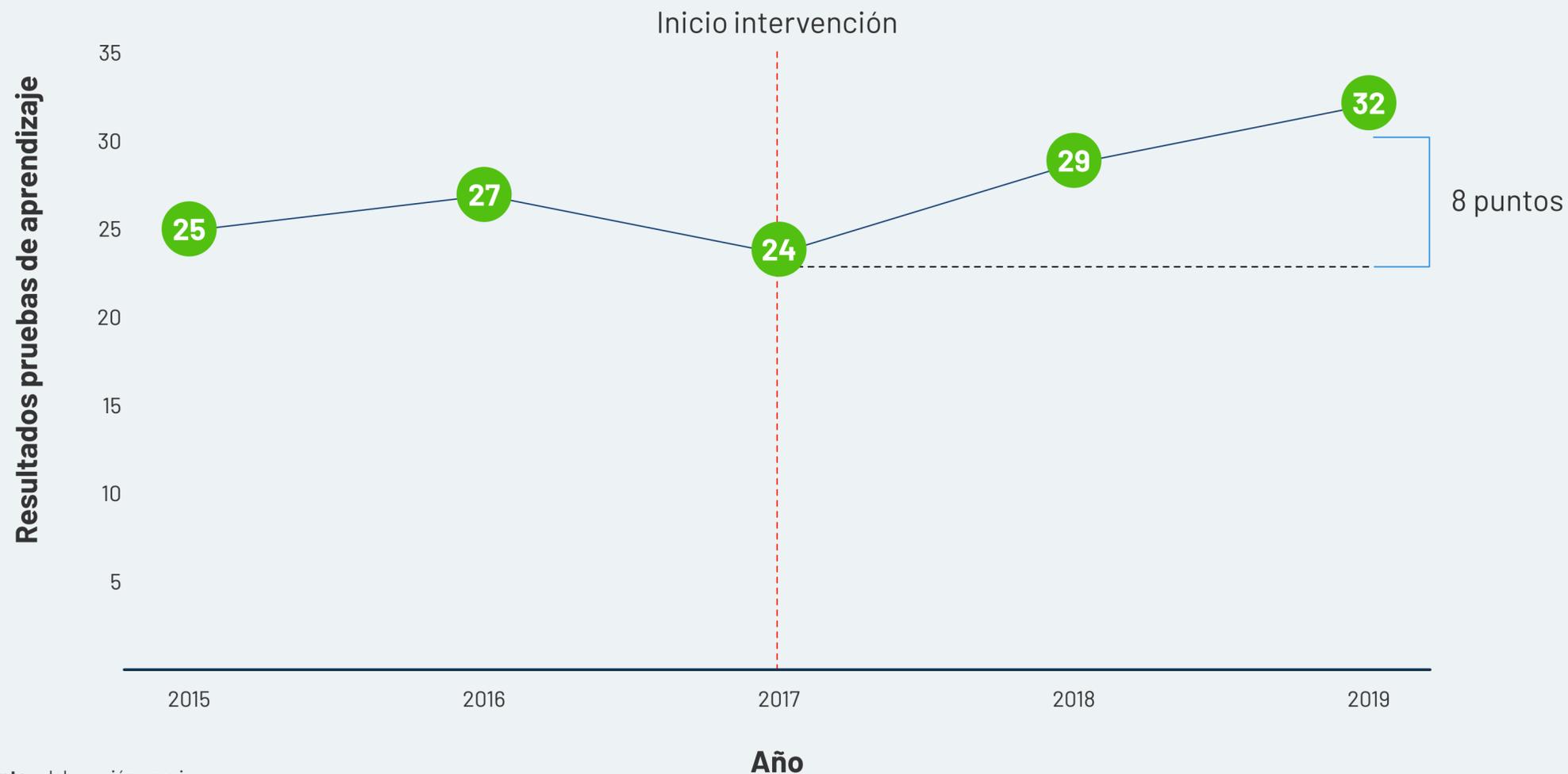
casos, puede ser útil desglosar el resultado, separando sus distintas partes, o enfocarse en los cambios a nivel local o institucional implícitos en el resultado, en lugar de en los cambios a nivel general, agregado y nacional³.

³ Ibidem.

Ejemplo aplicado

Un ejemplo aplicado puede ser el siguiente, donde se relacionan años sin intervención y con ella, con variación de resultados en pruebas de aprendizaje.

Figura 3. Ejemplo de gráfica de medición de resultados



Fuente: elaboración propia.

Suponga que una intervención que brinda formación docente con un único colegio beneficiario comenzó su implementación en el año 2017. Además, asumamos que la principal variable de resultado a medir son los puntajes obtenidos en pruebas estandarizadas. La evaluación de resultados en este caso consistiría en ver la diferencia entre los puntajes en el periodo posterior a la intervención (32 puntos) y el periodo previo a esta (24 puntos). De esta manera, es posible afirmar que la diferencia en los resultados en términos de puntaje en la prueba estandarizada es de 8 puntos.

Es importante resaltar que esto no implica causalidad, y que, por tanto, esta diferencia no es directamente atribuible a la intervención, razón por la cual resulta necesario realizar una posterior evaluación de impacto.

Para realizar la medición se recomienda:



Establecer objetivos específicos

Antes de poder evaluar los resultados, es fundamental definir con claridad los objetivos de desempeño que se espera alcanzar.



Determinar los criterios a evaluar

Es crucial identificar de manera precisa aquello que debe ser evaluado, evitar ambigüedades y elegir los “verdaderos resultados”.



Elegir los instrumentos adecuados

Para obtener información precisa es necesario que el instrumento sea el apropiado. Por ejemplo, no es útil evaluar la satisfacción del cliente mediante la evaluación del presupuesto de gastos. Existen instrumentos para este fin, como pequeñas encuestas aplicadas al momento de la venta. Aunque múltiples instrumentos pueden estar relacionados, su aplicación debe ser específica y estar diseñada correctamente para evaluar el criterio correspondiente.



Realizar evaluaciones periódicas

No significa evaluar en todo momento, sino evaluar cuando se debe hacerlo. Si el objetivo planteado requiere revisar cada semana, diariamente, o en cualquier

otro intervalo de tiempo, es necesario hacerlo para controlar que se está logrando el cometido. De aquí se desprende la importancia de realizar una adecuada planeación previa.



Utilizar los resultados de la evaluación

El objeto principal de las evaluaciones es servir como instrumento para la toma de decisiones y la implementación de acciones. Estas acciones podrán ser preventivas o correctivas, pero deben llevarse a cabo. Es decir, de nada sirve realizar el mejor proceso de evaluación si la información obtenida no nos es útil. Las evaluaciones deben servirnos para tomar acciones que permitan:

- Corregir cuando sea necesario.
- Mejorar lo que es posible mejorar.
- Comprobar que se hace lo correcto.

Tampoco se puede perder de vista que los factores que pueden influir en el resultado son múltiples y complejos, se relacionan entre sí y varían constantemente. Los pasos para identificar esos factores son⁴:

01

Realizar la recolección de información y analizarla

Para esto se hará uso de fuentes externas como variables de contexto, entrevistas y grupos focales.

02

Identificar los factores contribuyentes más importantes que impulsan el cambio

Los factores contribuyentes pueden ser oportunidades proactivas que deben ser capitalizadas o restricciones negativas que deben ser eliminadas. Se debe limitar el análisis a cinco o seis factores contribuyentes, de modo tal que puedan concentrarse en los factores clave.

03

Examinar fuentes locales de conocimiento en relación con los factores que influyen en el resultado

Se debe sintetizar la documentación existente sobre el desarrollo del área de interés.

04

Resolver el problema de los efectos de segundo orden, no intencionales, que está teniendo el resultado

Hay que identificar información sobre la imposibilidad de lograr el resultado con los hallazgos obtenidos. Con esto se puede identificar si los productos planeados eran los más relevantes en relación con los factores contextuales que influyeron en el resultado.

⁴ Ibidem.



Al evaluar los resultados es crucial considerar tanto las contribuciones generadas por la organización como aquellas surgidas de las alianzas con socios para modificar el resultado.

Las **contribuciones de la organización** abarcan toda la gama de actividades e iniciativas, tanto en el marco de proyectos como fuera de estos⁵.

Esas contribuciones pueden establecerse mediante:

⁵ Ibidem.

A

Análisis de la coherencia entre la estrategia y la gestión de las operaciones generales por parte de la organización en relación con los cambios a nivel de resultados. Los evaluadores pueden formular preguntas de investigación como las siguientes: ¿la gestión se centra en los resultados o se circunscribe a los productos?, ¿las discusiones con el gobierno se centran en los resultados?, ¿puede el personal articular la estrategia orientada a los resultados?

B

La verificación de si la planificación y la gestión de diferentes intervenciones por parte de la organización han sido alineadas para aprovechar sinergias que contribuyan a los resultados. Esto supone examinar los proyectos, programas y actividades que se ejecutan fuera del ámbito de los proyectos o programas y evaluar si son coordinados o no de una manera que vincule sus productos al resultado en cuestión. Las personas encargadas de la evaluación deben considerar una serie de preguntas: ¿se consulta y coordina con los socios clave?, ¿se están llevando a cabo actividades comunes?, ¿se está intercambiando información entre proyectos/programas?, ¿los resultados están siendo tomados como base de la planificación de actividades conjuntas?

C

Determinación de si los productos individuales son eficaces o no en contribuir a los resultados. La pregunta clave es: ¿los productos individuales son eficaces en contribuir a los resultados? Quienes realicen la evaluación deben centrar su análisis en aquellos productos que hayan sido —o demuestren el potencial para ser— genuinamente estratégicos a la hora de lograr avances hacia el resultado.

Los resultados son influenciados por una compleja gama de factores. Lograr un cambio invariablemente requiere la acción concertada de varias instituciones y actores⁶.

Las **contribuciones generadas por las alianzas** pueden establecerse mediante:



El examen del **grado de apoyo mutuo** entre los socios basado en la afinidad y el interés común, definidos por misiones de seguimiento conjuntas, evaluaciones, la ejecución de actividades relacionadas con el proyecto, etc. Se busca determinar si las actividades de los socios en la búsqueda del mismo resultado son coordinadas o no.

6 Ibidem.



El examen de **cómo se formaron las alianzas** y cómo ha sido **su desempeño**.



La realización de **evaluaciones conjuntas**. Las evaluaciones de resultados se prestan para efectuarse como evaluaciones conjuntas porque los resultados solo pueden alcanzarse en alianza con otros.

Evaluación de impacto

¿Qué es una evaluación de impacto?

La evaluación de impacto es un tipo de evaluación que se centra en responder preguntas relacionadas con la relación causal entre un resultado y el programa o proyecto específico. Su objetivo es identificar cambios observables en el resultado de interés que puedan atribuirse de manera directa y exclusiva al programa o proyecto evaluado.

Preguntas que busca responder una evaluación de impacto

Todas las evaluaciones de impacto responden una pregunta de causa y efecto.

Estas preguntas pueden aplicarse a cualquier ámbito, siempre que busquen evaluar el impacto de una modalidad de un programa o una innovación en su diseño, no solo el programa en sí mismo.

Ejemplo aplicado

Algunos ejemplos de preguntas que una evaluación de impacto podría abordar incluyen:

¿Cuál es el efecto causal (impacto) de proporcionar becas a estudiantes en su asistencia escolar y su desempeño académico?

¿Mejorar las carreteras y vías de acceso incrementa el acceso al mercado laboral y aumenta los ingresos de los hogares?

Si se reemplazan pisos de tierra por pisos de cemento, ¿cuál es el impacto de este cambio en la salud de los niños?

¿El tamaño de las clases afecta el aprendizaje de los estudiantes?

Para responder preguntas de **causa y efecto**, es necesario entender el concepto de resultados potenciales o contrafactuales. Con el fin de realizar una definición formal de este problema supongamos que y_i es el resultado de interés (y) para el individuo i . Adicionalmente, consideremos los superíndices T y C para denotar la pertenencia a los grupos de tratamiento (beneficiarios del programa) y control (no beneficiarios), respectivamente. Así, el impacto de un programa o intervención está dado por:

$$\text{Impacto}_i = y_i^T - y_i^C$$

Esto es, el impacto del programa para el individuo i es la diferencia entre el resultado que este obtiene cuando es parte del grupo de tratamiento y el resultado que obtiene cuando es parte del grupo de control.

El problema para obtener este impacto radica en que, por definición, ningún individuo puede ser a la vez beneficiario y no beneficiario del programa a evaluar; no es posible observar y_i^T y y_i^C simultáneamente. La medición directa del impacto solo sería posible si para cada individuo tuviéramos un individuo igual en todas las medidas posibles (como un gemelo o clon), a excepción de tratamiento, al cual denominamos *contrafactual*.

El contrafactual es el **escenario hipotético** que habría tenido lugar en ausencia del programa. Particularmente, para ver el efecto del tratamiento, el contrafactual de interés es el resultado observado en el escenario en el cual los participantes del programa no hubieran participado en él.

7 Roy, 1951; Rubin, 1974; Bernal y Peña, 2011.

Ejemplo aplicado

Si se quisiera ver el efecto causal (impacto) de proporcionar becas a estudiantes en su asistencia escolar y desempeño académico, es necesario evaluar cómo hubiese sido la asistencia y el desempeño académico de los estudiantes beneficiarios de las becas si no se les hubiera dado la beca.

Si bien es cierto que no es posible observar un contrafactual para cada individuo, sí es posible construir contrafactuales si nuestro objetivo es analizar promedios. Más aún, por esto es por lo que cuando se habla del efecto de un programa o intervención se suele hablar del efecto promedio del tratamiento.

Para entender la clave de por qué este ejercicio es viable al utilizar promedios sigamos con nuestro ejemplo del programa de becas. Supongamos que nuestros grupos de tratamiento y control están compuestos cada uno por tres individuos, a los cuales denominaremos con las letras A, B y C. Imaginemos entonces que nos interesa evaluar la correspondencia de los dos grupos en términos de las edades de sus miembros, las cuales se presentan en la siguiente tabla:

	TRATAMIENTO	CONTROL
Edad A	17	15
Edad B	15	16
Edad C	13	14

De esta manera, ambos grupos en nuestro ejemplo tienen una edad promedio de 15 años, por lo que son en promedio comparables en este ámbito, aun cuando sus individuos cuentan con edades distintas. Si de alguna manera logramos garantizar que esto mismo ocurra con todas las características de interés, no solo con la edad, entonces el grupo de control es un buen contrafactual para el grupo de tratamiento.

Usualmente, para conseguir la correspondencia entre los dos grupos se emplean modelos estadísticos que dependen del diseño del programa y del método de focalización que se llevó a cabo en este. Así, según el método de elección de los beneficiarios, los diseños de evaluación de impacto se clasifican como *experimentales* o *no experimentales*⁸.

En los **diseños experimentales** los individuos son asignados de manera aleatoria al grupo de tratamiento o el grupo de control. Al asignar los individuos beneficiarios del programa con este método se garantiza que estos son comparables en todas sus características observables y no observables.

Los diseños experimentales pueden asignar el tratamiento de diferentes maneras según la situación que se presente en el proyecto a evaluar, manteniendo siempre el criterio de aleatorización. Así, la aleatorización puede ser:

⁸ Morra-Imas y Rist, 2009.

01

Aleatorización simple: consiste en ordenar a los potenciales beneficiarios del programa de acuerdo con un número pseudoaleatorio generado de manera automática y con base en este orden seleccionar al grupo de tratamiento y el grupo de control. Usualmente, este método se conoce como el algoritmo de coordinado negativo.

02

Aleatorización por etapas: se realiza el mismo proceso que en la aleatorización simple, con la distinción de que el grupo de tratamiento es dividido en múltiples subgrupos con el fin de emplear tratamientos diferenciados o secuenciales. En este caso, el grupo de control no debe ser alterado, con el fin de contar con un número adecuado de individuos de comparación.

03

Aleatorización por estratos: los individuos son asignados al tratamiento de manera aleatoria, considerando números pseudoaleatorios generados de manera separada para cada uno de los niveles de una variable categórica (estrato) que denota una característica preexistente a la intervención que puede afectar la probabilidad de ser beneficiario del programa.

04

Aleatorización por bloques: los individuos son asignados al tratamiento de manera aleatoria, considerando números pseudoaleatorios generados de manera separada para cada uno de los niveles de una variable categórica (bloque) que denota una característica seleccionada por los investigadores que puede afectar la probabilidad de ser beneficiario del programa.

Los efectos de un programa con un diseño experimental pueden ser obtenidos mediante la diferencia en los promedios de los resultados del grupo de tratamiento y control, una vez haya transcurrido un tiempo considerable en la implementación del programa.

Los **diseños no experimentales** son aquellos en los cuales la asignación al tratamiento no se realiza de manera aleatoria y, por tanto, existen múltiples características observables y no observables que se relacionan con la inclusión de los individuos en el programa. Dadas estas disimilitudes, la diferencia de los resultados promedio de los dos grupos no brinda suficiente información del impacto de la intervención. Es necesario entonces emplear métodos estadísticos más robustos tales como:



Diferencias-en-diferencias (DiD)



Emparejamiento



Variables instrumentales



Regresión discontinua

Vale la pena resaltar que los dos principales factores para la elección del método no experimental son: el método de asignación al programa y la disponibilidad de la información. Por tanto, no es viable aplicar todos los métodos en cualquier situación.



Volviendo a nuestro ejemplo de becas, si estas no fuesen asignadas de manera aleatoria, es bastante probable que los participantes en el programa sean estudiantes que, en promedio, son más motivados o cuentan con mejor información sobre su contexto escolar (características no observables). Estas diferencias en motivación pueden redundar en diferencias en la deserción escolar, las cuales serían erróneamente atribuidas al programa al no tener en cuenta el mecanismo de selección de los beneficiarios.

Tipos de evaluaciones de impacto según el tiempo

Las evaluaciones de impacto pueden clasificarse como prospectivas o retrospectivas según la manera en que se midan las variables de resultado. Usualmente, las evaluaciones prospectivas son más robustas, en cuanto estas contemplan la medición de los resultados del programa desde su inicio. Las evaluaciones retrospectivas suelen realizarse cuando se tiene muy poca información sobre la implementación y diseño del programa.



Evaluaciones prospectivas

Estas evaluaciones se planifican y desarrollan paralelamente al diseño del programa, formando parte integral de su proceso de implementación.



Las evaluaciones prospectivas se caracterizan por:

A

Contar con una línea base

La línea base consiste en los datos sobre participantes y no participantes del programa que se recogen antes de ser implementado el programa. Tener estos datos permite:

- Medir las variables de interés o resultados antes de la implementación del programa.
- Comparar las características observables de los grupos de tratamiento y control con el fin de evaluar si ambos grupos, sin la intervención del programa, son en promedio iguales en sus características.
- Evaluar la efectividad de la focalización del programa, es decir, determinar si el programa efectivamente alcanzó a los beneficiarios planeados.

B

Contar con medidas de éxito del programa

Durante la etapa de diseño se deben hacer explícitas las medidas que permitan definir si el programa tuvo éxito en alcanzar sus objetivos. Estas medidas:

- Se fundamentan en la teoría de cambio.
- En la etapa de diseño del programa y de su evaluación se deben establecer con claridad los objetivos del programa y, por tanto, las medidas apropiadas para evaluar el éxito del mismo.
- Además de claridad en los objetivos que el programa debe alcanzar, debe haber claridad en las preguntas de investigación que ha de responder la evaluación.

C

Identificación *ex ante*

Los grupos de tratamiento y control deben identificarse antes de que se implemente el programa a evaluar. Cuando la evaluación se planifica desde el principio, antes de la implementación del programa, existen numerosas opciones para realizar evaluaciones de impacto válidas.

Casi siempre es posible obtener estimaciones válidas y precisas del contrafactual para cualquier programa, siempre que este tenga reglas claras y transparentes para la asignación del tratamiento, en el marco de una evaluación prospectiva.



Evaluaciones retrospectivas

Estas evaluaciones se caracterizan por:

A

Desarrollarse sin haber considerado la recopilación de información necesaria desde el inicio del programa. Esto impide que sea fácil determinar si el programa fue implementado con éxito y si sus participantes realmente se beneficiaron de él.

B

Evaluar el impacto del programa después de que el programa fue implementado (identificación *ex post*).

C

En muchos casos, en la etapa de diseño del programa no se recolecta una línea base. Cuando esto sucede, y el programa se desarrolla, suele ser demasiado tarde para recolectar el tipo de información necesaria. En estos casos hay muchas menos opciones para obtener estimaciones válidas del contrafactual.

D

La evaluación normalmente resulta dependiente de la existencia de reglas claras y precisas de asignación del tratamiento y operación del programa.

E

También depende en gran medida de la disponibilidad de datos e información con suficiente cobertura de los grupos de tratamiento y de control antes y después de la implementación del programa.

F

La factibilidad de una buena evaluación retrospectiva depende considerablemente del contexto del programa y su diseño, y nunca está garantizada.

G

En los casos en que se garantice la factibilidad de la evaluación, esta normalmente se basa en métodos cuasiexperimentales, por lo cual tiende a depender de supuestos muy fuertes sobre los datos y el programa en sí. Por ello, la evidencia encontrada suele ser objeto de debate.



Descargue aquí la plantilla para los
productos necesarios para la evaluación



¿Cómo decidir si hacer o no evaluación de impacto?

No todos los programas o proyectos requieren de una evaluación de impacto. Si se lleva a cabo es porque se quiere responder a una pregunta de tipo causa-efecto que ofrezca una información específica sobre el programa o proyecto.

Hay que tener en cuenta que, si no se cuenta con información disponible, se requiere recolectar datos, y esto puede hacer que la evaluación sea muy costosa.

Si se está empezando un nuevo programa, o se piensa en expandir o innovar uno ya existente, basta con responder una serie de preguntas que ayudarán a tomar la decisión de realizar una evaluación de impacto. En la siguiente figura se presenta el camino de preguntas que ayudará a determinar si se debe hacer o no una evaluación de impacto.

Figura 4. Preguntas para decidir si hacer o no una evaluación de impacto



Fuente: elaboración propia.



Descargue aquí la plantilla del formato de decisión



Análisis de costo-efectividad

El análisis costo-efectividad es un método que se utiliza para medir la razón entre el costo monetario de un programa o intervención y los resultados que este brinda, los cuales son obtenidos durante la etapa de evaluación de impacto. De esta manera, un análisis costo-efectividad permite cuantificar la efectividad obtenida por cada unidad de costo adicional, a través de un cálculo sencillo que consiste en la agregación de los costos y beneficios totales.

Adicionalmente, es posible comparar diferentes intervenciones a través de sus análisis de costo-efectividad y con base en estos determinar el programa que se presenta como una mejor alternativa en términos financieros.

Las evaluaciones costo-efectividad son bastante comunes en el análisis de programas de educación y de salud, en los cuales los impactos son medidos con variables tales como el desempeño académico de los estudiantes o el estado de salud de los individuos tratados. Sin embargo, en casos en los cuales los efectos del programa pueden ser traducidos a términos monetarios (por ejemplo, un programa que busque incrementar el salario de sus beneficiarios), es recomendable realizar también un análisis de costo-beneficio.

Requisitos para hacer un análisis de costo-efectividad

Realizar un análisis de costo-efectividad requiere hacer consideraciones detalladas sobre los impactos y costos del programa a estudiar.



En cuanto a los **impactos del programa**:

- Se requieren estimaciones precisas del impacto en todos los resultados de interés junto con la periodicidad de dichos efectos.
- Se debe realizar una revisión sistemática de la literatura relevante. Esto con el fin de identificar cuáles son los efectos obtenidos en intervenciones con estructuras y contextos comparables.



En cuanto a los **costos del programa**:

- En la medida en que contar con información retrospectiva implica mayores costos y márgenes de error, se recomienda hacer un seguimiento constante a los costos de implementación del programa durante su desarrollo.
- Es fundamental recolectar y contabilizar de manera desagregada, entanto esto permite evaluar de manera exacta cuáles fueron las compras y contrataciones, así como sus costos asociados.
- Es importante tener en cuenta todos los factores que regularmente afectan las medidas monetarias, tales como las tasas de cambio, la deflación en el año base, la inflación en el año de análisis, así como la transformación de todos los costos a valor presente.

Un análisis detallado del costo del programa debería identificar:

01

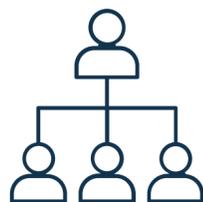
Lista de los componentes y materiales necesarios para la implementación del programa

02

Cantidades y costos unitarios de estos materiales



Los componentes del costo del programa se dividen en las siguientes categorías:



Administración del programa

Incluye personal contratado para trabajar en la implementación del programa, el costo de la infraestructura necesaria y cualquier costo del *overhead* que haya sido incluido.



Costos de focalización

Incluye todo costo incurrido en focalizar, identificar y crear conciencia sobre el programa en los potenciales beneficiarios.



Capacitación del personal

Incluye costos de capacitar al personal encargado de implementar el programa.



Capacitación de beneficiarios

Incluye todos los costos incurridos en capacitar a los beneficiarios o potenciales participantes del programa.



Costos de implementación

Incluye el costo de material distribuido entre los participantes, costos del personal que únicamente trabajó en las actividades implementadas, o el costo de crear y mantener los recursos desarrollados para la intervención.



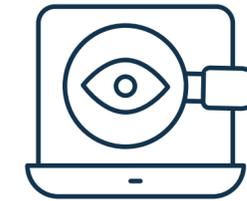
Costos de los beneficiarios

Incluye todos los costos incurridos por parte de los participantes del programa, como el costo de los bienes que los participantes tuvieron que comprar o adquirir, o el costo de oportunidad asociado al tiempo invertido por los participantes.



Costos evitados

Incluye los costos que fueron reemplazados o descontinuados como resultado de la intervención.



Costos de monitoreo

Incluye los costos incurridos en hacer seguimiento y monitoreo a los participantes y el personal del programa.

Método de ingredientes para el análisis de costo-efectividad

Nuestra perspectiva va en línea con J-PAL, en cuanto consideramos que es recomendable seguir la guía propuesta por Dhaliwal et al. (2012) con el fin de realizar el análisis costo-efectividad. Esta guía facilita la recolección de los costos desagregados del programa usando un método de ingredientes, el cual consiste en listar los ingredientes del programa, estimar sus costos unitarios y determinar su utilización a través de la implementación de la intervención. De acuerdo con Dhaliwal et al. (2012), los costos asociados al diseño, evaluación e investigación sobre el programa no deben ser considerados en dicho análisis.

Una vez obtenidos todos los costos y efectos del programa, estos se pueden agregar con el fin de construir la razón costo-efectividad de este de la siguiente manera:

$$CE = \frac{\text{Impactos con el programa} - \text{Impactos sin el programa}}{\text{Costos con el programa} - \text{Costos sin el programa}}$$



Para poner esto en contexto, veamos de nuevo nuestro ejemplo del programa de becas. Supongamos, por un lado, que sabemos, a través de una evaluación de impacto, que la intervención disminuye la deserción escolar en 5 puntos porcentuales (p.p.). Por otro lado, asumamos que al adicionar nuestros costos fijos (10.000 USD) y variables (5.000 USD) obtenemos unos costos totales de 15.000 USD. Dada esta información, y considerando un costo de cero en ausencia del programa, la razón de costo-efectividad es de 5 p.p./15.000 USD. Esto implica que, por cada 1.000 USD invertidos, la intervención estaría logrando reducir la deserción promedio de los beneficiarios en 0,3%.



**Descargue aquí la plantilla para
costos de J-PAL**



**Descargue aquí el caso Aprendamos
todos a leer (ATAL) etapa 5**



Referencias

- Álvarez, H., Berlinski, S y Busso, M. (2019). *Remedial Education: Evidence from a Sequence of Experiments in Colombia*. IDB Working Paper Series. Obtenido de: <https://publications.iadb.org/en/remedial-education-evidence-sequence-experiments-colombia>
- Banco Mundial. (s.f.). *Árbol de problemas y Experiencias*. Documento no publicado.
- Bernal, R. y Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Bogotá: Universidad de los Andes.
- Banco Interamericano de Desarrollo [BID]. (s.f.). *Aprendamos todos a leer*. <https://www.iadb.org/es/sectores/educacion/aprendamos-todos-leer>
- Cohen, E. y Martínez, R. (s.f.). *Manual: Formulación, evaluación y monitoreo de proyectos sociales*. CEPAL. Obtenido de: <https://dds.cepal.org/re-desoc/publicacion?id=242>
- Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2013). *Modelo de términos de referencia para la evaluación de procesos de programas de desarrollo 2013*. Ciudad de México: CONEVAL.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2017). *Términos de referencia evaluación de procesos*. Ciudad de México: CONEVAL.
- Dhaliwal, I., Duflo, E., Glennerster, R. y Tulloch, C. (2012). *Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education*. Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT. Obtenido de: https://www.povertyactionlab.org/sites/default/files/research-resources/CEA%20in%20Education%202013.01.29_0.pdf
- Duflo, E., Glennerster, R. y Kremer, M. (2006). *Using Randomization in Development Economics Research: A Toolkit*. NBER Technical Working Paper Series. Obtenido de: https://www.nber.org/system/files/working_papers/t0333/t0333.pdf
- Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B. y Vermeersch, C. (2011). *La evaluación de impacto en la práctica*. Washington D.C.: World Bank. Obtenido de: <https://openknowledge.worldbank.org/entities/publication/f090e5a0-16f6-5795-a3b2-7f711c8b7eb7>
- Gertler, P., Martinez, S., Premand, P., Rawlings, L. y Vermeershc C. (2016). *Impact Evaluation in Practice. Second Edition*. World Bank. Obtenido de: <https://openknowledge.worldbank.org/handle/10986/25030>
- Glewwe, P. y Todd P. (2022). *Impact Evaluation in International Development: Theory, Methods and Practice*. World Bank. Obtenido de: <https://openknowledge.worldbank.org/handle/10986/37152>

Instituto Latinoamericano y del Caribe de Planificación Económica y Social [ILPES]. (1995). *Dirección de proyectos y programación de inversiones*.

Instituto Latinoamericano y del Caribe de Planificación Económica y Social [ILPES]. (1995). *Guía para la identificación y formulación de proyectos de educación*. Santiago de Chile: CEPAL.

López, J. (s.f.). *Guía para el diseño y evaluación de los proyectos*. Centro de Estudios Rurales y de Agricultura Internacional. Obtenido de: <https://cerai.org/wordpress/wp-content/uploads/2021/03/Gu%C3%ADa-para-el-dise%C3%B1o-y-evaluaci%C3%B3n-de-los-proyectos-CERAI.pdf>

Morra-Imas, L. G. y Rist, R. C. (2009). *El camino hacia los resultados: diseño y realización de evaluaciones eficaces para el desarrollo*. Washington D. C.: World Bank.

National Endowment for Science Technology and the Arts. (2016). *Using Research Evidence: A Practice Guide*. Obtenido de: <https://www.bl.uk/collection-items/using-research-evidence-a-practice-guide>

Ortegón, E., Pacheco, J. y Roura, H. (2005). *Metodología general de identificación, preparación y evaluación de proyectos de inversión pública*. Serie Manuales, núm. 39. Santiago de Chile: ILPES-CEPAL. Obtenido de: <http://repositorio.cepal.org/handle/11362/5608>

PNUD. (2002). *Informe sobre desarrollo humano 2002: Profundizar la democracia en un mundo fragmentado*. Madrid/Barcelona: Mundi-Prensa Libros. Obtenido de: <https://hdr.undp.org/system/files/documents/hdr2002espdf.pdf>

Roy, A. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2), pp. 135-146.

Rubin, D. (1974). Estimating Causal Effects of Treatment in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5), pp. 688-701.

Secretaría de la Función Pública. (2016). *Guía para la optimización, estandarización y mejora continua de procesos*. México: Gobierno de los Estados Unidos Mexicanos.

Shotland, M. (s.f.). *What is evaluation?* [Diapositivas de PowerPoint]. Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT.

USAID. (2015). *Manual de diagnóstico de problemas y diseño de programas*. Documento no publicado.

