Guía para la evaluación comprensiva de programas educativos

Rafael López-Meseguer. Profesor del Área de Métodos de Investigación de la Universidad Internacional de la Rioja.

Cómo citar esta guía

López-Meseguer, R. (2026). Guía para la evaluación comprensiva de programas educativos. AULA UNO.01.

Qué es esta guía

Esta guía orienta la aplicación del modelo de evaluación comprensiva a un programa educativo. Explica cómo comprender su funcionamiento, analizar la adecuación entre medios y fines, recoger las valoraciones de sus participantes y examinar los cambios que produce. Organiza el trabajo en cuatro fases y señala las tareas y los productos de cada una.

A quién se dirige

Se dirige a profesionales de la educación que promueven, gestionan o evalúan programas: equipos directivos, profesorado, responsables de programas, personal investigador y equipos de evaluación.

¿Necesitas asesoramiento profesional?

Si lo que busca es asesoramiento profesional para llevar a cabo una evaluación comprensiva puede contactar con rafael.lopezmeseguer@unir.net

01

Definir el programa y el propósito

Qué significa evaluar de forma comprensiva

El Modelo de evaluación comprensiva de programas educativos es un modelo consolidado y avalado por la literatura científica especializada (López-Meseguer y Valdés Fernández, 2020). Una evaluación comprensiva es el resultado del análisis de la coherencia lógica y la correspondencia causal de un programa. Para ello, la evaluación se orienta a conocer (1) la adecuación del programa en términos de medios-fines, (2) la valoración del programa por parte de los participantes del mismo, (3) y la correspondencia entre los cambios que el programa prescribe y los que realmente produce. A continuación, se exponen sucintamente cada uno de los elementos que componen este modelo de evaluación.

Qué se entiende por programa

Un programa es habitualmente entendido como un plan destinado a desarrollar una serie de actividades en beneficio de un colectivo determinado. Nosotros, sin embargo, entendemos que un programa es un tipo de relación social que establece una serie de fines con los que orientar las acciones de los participantes y que han de ser reconocidos por los mismos. De acuerdo con esta definición, genérica, no importa la fase de desarrollo en que se encuentre, la magnitud y el tipo de programa, o el número de participantes con el que cuente. En el ámbito educativo, la mayoría de programas proceden de la iniciativa de los poderes públicos correspondientes (a nivel supranacional, nacional, regional o local), de la actividad académica investigadora, de la responsabilidad social de algunas instituciones de la sociedad civil, o bien como resultado de innovaciones que tienen lugar de manera continuada en centros educativos. A pesar de la multitud de programas procedentes de cada una de estas instancias, la mayoría de estos quedan sin evaluar, impidiendo de ese modo la retroalimentación del sistema a partir de buenas prácticas consolidadas que puedan ser objeto de conocimiento por el conjunto de la comunidad educativa.

Para qué evaluar

En términos generales, el modelo de evaluación comprensiva de programas educativos permite disponer de información suficiente para poder afirmar si el programa produce cambios y cómo, lo que ayudará en primer término a la mejora de los programas objeto de evaluación, y lo que permitirá, a su vez, la posibilidad de generar buenas prácticas que faciliten la escalabilidad de dichos programas. En segundo lugar, este modelo persigue también obtener medidas estadísticas sobre los efectos o cambios del programa, y trata de imputar causalmente, cuando sea posible, si esos efectos se deben a la acción del programa o bien tienen que ver con otras razones o causas a tener en cuenta.

Quién debe participar

La evaluación no es una mera actividad académica sino un oficio, y el oficio de evaluar requiere de equipos profesionales formados a tal efecto. Así pues, junto con el equipo de investigadores capaces de diseñar e implementar las técnicas de evaluación más adecuadas, se requiere contar con profesionales de la gestión del mundo educativo que sean capaces de comprender y conciliar los intereses de responsables políticos, de financiadores de programas, de equipos directivos y, sobre todo, del profesorado y del alumnado de los centros que participan en programas objeto de evaluación.

02

Comprender el programa antes de medirlo

La evaluación comprensiva se orienta a partir de una serie de principios de actuación, que beben o tienen su raíz en el concepto de comprensividad (*verstehen*) desarrollado por Max Weber.

Comprender los significados de los participantes

Partimos de que toda actuación educativa, para ser evaluada, necesita primero ser comprendida. Comprender un programa educativo lleva consigo el intento de captar su individualidad, es decir, atender y dar prioridad al significado subjetivo que le otorgan los participantes del programa. Los programas educativos funcionan o dejan de funcionar en la medida en que aquello que prescribe el programa es o no puesto en práctica (con las presumibles variaciones según cada caso) por sus participantes, y ello tiene lugar a través de un proceso de internalización. De ahí que, recabar información sobre las percepciones de los participantes sobre el programa, en sus diferentes fases de desarrollo, se convierta en un objetivo prioritario de la evaluación comprensiva. Complementariamente, en el ámbito educativo, suele ser habitual servirse de pruebas estandarizadas de tipo competencial, diagnóstica (TRI), u otras realizadas ad hoc, con las que medir los progresos de los participantes en función de las características del programa objeto de evaluación. La combinación de ambos elementos es lo que nos llevará a obtener una medida comprensiva del programa.

Analizar la coherencia lógica y la correspondencia causal

La evaluación comprensiva se fundamenta en el análisis de la coherencia lógica y la correspondencia causal. Decimos que un programa es coherente lógicamente cuando (1) sea posible afirmar, en base al conocimiento científico existente y a las reglas de la experiencia común, que aquello que el programa prescribe tiene fundamento lógico y, (2) cuando los fines del programa hayan sido comúnmente reconocidos por los participantes en grado suficiente. Por otro lado, decimos que un programa se corresponde causalmente cuando (1) se atestigua que los participantes del programa se comportan conforme a las prescripciones del mismo, es decir, que el programa produce cambios en los comportamientos o en la realidad educativa, y (2) cuando se pueda estipular que esos cambios se deben al programa objeto de evaluación.

La coherencia lógica, por tanto, prestaría mayor atención al proceso, en cuanto que buscaría averiguar la correcta adecuación de objetivos, prácticas y resultados desde la perspectiva del evaluador y de los participantes. La comprobación de la correspondencia causal, por su parte, estaría más centrada en los resultados, ya que buscaría analizar el grado en que la situación típica, es decir, lo que fija el programa, ocurre en la realidad, ya sea a través de la perspectiva de los participantes o de otro tipo de pruebas estandarizadas con las que medir la eficacia del programa. La combinación de ambos aspectos es lo que nos permite obtener una visión comprensiva del programa evaluado.

Aplicación práctica Recabe las percepciones de los participantes y examine si reconocen los fines del programa, si ponen en práctica sus prescripciones y si los cambios observados pueden atribuirse al programa.

03

Articular la metodología y organizar el recorrido

Combinar las técnicas según las necesidades de información

La medida técnica de nuestras evaluaciones se fundamenta en un proceso de articulación metodológica, según el cual se desarrolla una estrategia particular que satisfaga las necesidades de información que se plantean de la combinación de los objetivos específicos de la evaluación que emanan de los requerimientos del programa y de las premisas en las que se fundamenta el modelo de evaluación comprensiva para poder hacer afirmaciones sobre su funcionamiento.

Para hacer eso posible, la evaluación comprensiva combina el análisis de los juicios de valor de los participantes acerca del programa (desde una perspectiva cuantitativa y cualitativa), junto con otro tipo de pruebas estandarizadas con las que tratar de obtener una medida de los progresos de los participantes. En ese sentido, hay que señalar que desentrañar la lógica de un programa y comprobar su correspondencia causal no puede reducirse a la aplicación de una técnica de investigación particular, por poderosa que ésta pueda ser desde el punto de vista estadístico, sino que ha de ser fruto de una articulación de diferentes metodologías de investigación-evaluación.

Cómo utilizar la guía

Recorra las cuatro fases con el equipo evaluador y los responsables del programa. Todas son necesarias para fundamentar las afirmaciones sobre su funcionamiento; revise la lógica del programa cuando resulte necesario.

Fase 1. Sistematizar la lógica del programa

Formule las hipótesis de cambio, revise la adecuación entre medios y fines y concrete los indicadores de proceso y resultado.

Fase 2. Diseñar la evaluación

Acuerde el diseño, los indicadores, las técnicas y el presupuesto en función del programa y los recursos disponibles.

Fase 3. Gestionar el trabajo de campo

Coordine los centros, programe las fechas, aplique las técnicas y documente las incidencias.

Fase 4. Analizar y difundir los resultados

Interprete los procesos y los cambios, elabore el informe y organice la comunicación de los resultados.

04

Fase 1. Sistematizar la lógica del programa

Propósito Valorar la consistencia teórica del programa y hacer posible su evaluación.

Qué debe tener en cuenta

La implementación del programa constituye el elemento central a la hora de que un programa despliegue sus efectos eficazmente. La relación que existe entre si el programa se ha conducido según lo planeado y los resultados producidos es lo que habitualmente se conoce como teoría del programa. De hecho, la mayor parte de las incongruencias entre los objetivos planteados y los resultados finalmente alcanzados suelen deberse a algún tipo de problema de implementación (*implementation failure*). Consecuentemente, la revisión de la lógica de un programa -en aquellos casos que resulte necesario- constituiría la primera fase del proceso de evaluación. A través de este proceso, que se lleva a cabo conjuntamente con los responsables del programa, se formula el programa en términos de las hipótesis de cambio que plantea y se analiza su adecuación y plausibilidad en términos de adecuación entre medios y fines. A este proceso lo denominamos sistematización de la lógica del programa. En definitiva, estamos hablando de evaluar cómo se ha programado el programa en consonancia con nuestro modelo de implementación comprensiva. Esta fase es crucial a la hora de valorar si el programa tiene consistencia teórica, o bien si resulta necesaria la reformulación de los objetivos del mismo con vistas a que el programa resulte más eficaz y hacer posible su evaluación.

Cómo llevarla a cabo

- Trabaje con los responsables del programa para formular las hipótesis de cambio que plantea.
- Analice la adecuación y la plausibilidad de la relación entre los medios y los fines, apoyándose en el conocimiento científico y en evaluaciones del mismo ámbito.
- Determine si es necesario reformular los objetivos y concrete los indicadores de proceso y de resultado de la implementación.

Qué debe obtener

- Un informe en el que se formule y pueda visualizarse la lógica funcional del programa. En él se indican específicamente los indicadores de proceso y de resultado de la implementación del programa.
- Una serie de recomendaciones para la mejora del programa, con carácter previo a la evaluación. Estas recomendaciones serán fruto del conocimiento científico existente sobre la materia que aborda el programa, así como de otras evaluaciones llevadas a cabo en el mismo ámbito que permitan establecer paralelismos sobre el funcionamiento.

05

Fase 2. Diseñar la evaluación

Propósito Seleccionar un diseño y unas técnicas adecuados a las hipótesis de cambio, al estado del programa y a los recursos disponibles.

Cómo llevarla a cabo

El segundo paso lo constituye la selección del diseño y las técnicas de evaluación. Para ello es necesario, en primer lugar, plantear los principales indicadores de la evaluación, una vez el programa está formulado en hipótesis de cambio. En segundo lugar, acuerda un diseño de evaluación junto con las técnicas de evaluación que acompañan ese diseño. El modelo de evaluación comprensiva combina técnicas de evaluación cuantitativas y cualitativas con vistas a obtener información sobre la implementación y los resultados del programa. A mayor cantidad de técnicas de evaluación desplegadas, mayor y mejor información sobre el funcionamiento del programa. En ese sentido, las posibilidades son variadas y pueden ir desde un diseño experimental de evaluación, cuando la consistencia del programa y la financiación lo permita, hasta un diseño basado en un estudio de caso o experiencia piloto, cuando no se disponga de una muestra suficiente, pero se persiga con la evaluación hacer emerger una determinada práctica innovadora, sistematizarla, y escalarla a otros interesados.

Seleccione entre los cuatro tipos de diseños evaluativos que se presentan a continuación y en la página siguiente.

A) Evaluación comprensiva de procesos

Mediante la articulación metodológica de técnicas cuantitativas y cualitativas, esta evaluación se dirige a corroborar la coherencia lógica, es decir, la correcta implementación del programa desde la perspectiva de los participantes. Es un tipo de evaluación formativa, que sirve fundamentalmente para mejorar el programa objeto de aplicación, y que es especialmente propicia para proyectos piloto.

B) Evaluación comprensiva de procesos y resultados

Mediante la articulación metodológica de técnicas cuantitativas y cualitativas, esta evaluación se dirige a corroborar la coherencia lógica y la correspondencia causal de un programa a partir de las percepciones de los participantes. Es un tipo de evaluación formativa y sumativa, que se inspira en la teoría del cambio, y sirve al propósito de tener una primera medida de la adecuación de los procesos, así como de los resultados, y que se recomienda para programas que todavía no cuentan con un alto grado de consolidación en el ámbito educativo.

C) Evaluación comprensiva estandarizada

Este tipo de evaluación suma a los elementos anteriores la realización de pruebas competenciales (en aquellos casos donde sea posible) o diagnósticas (TRI), con el objetivo de disponer de una mayor fiabilidad con respecto a los efectos del programa. Para la aplicación de dichas pruebas se lleva a cabo un diseño pretest-postest, con el que medir la evolución de los participantes tras la implementación del programa. Este tipo de evaluación es particularmente idónea para programas que se encuentran más consolidados.

D) Evaluación comprensiva de impacto

Este tipo de evaluación compara los efectos del programa en un grupo tratado (donde el programa es implementado), con los de un grupo de control (donde no se implementa). En caso de llevarse a cabo mediante un proceso previo de aleatorización de los grupos (evaluaciones experimentales), es posible aislar los posibles efectos que tengan que ver con factores externos al programa, y de esa manera imputar causalmente -de modo más preciso- sus efectos; mientras que, en las evaluaciones cuasiexperimentales, por su parte, el requisito de la aleatorización es rebajado, lo que provoca una menor precisión en dicha imputación. Este tipo de evaluación es especialmente propicia para programas que buscan ser implementados como políticas públicas y, por tanto, requieren un alto grado de generalización.

Acordar las decisiones de diseño

Este proceso de diseño de evaluación requiere de la colaboración de los equipos evaluadores y los responsables del programa, para establecer coherentemente los objetivos de la evaluación y las posibilidades técnicas que los recursos permitan desplegar.

- Establezca los objetivos de la evaluación con los responsables del programa.
- Seleccione los indicadores cuantitativos y cualitativos y las técnicas del trabajo de campo.
- Ajuste el diseño a las posibilidades técnicas y económicas del proyecto.

Qué debe obtener

- El diseño metodológico de la evaluación, ajustando los parámetros de la evaluación comprensiva al programa objeto de evaluación. Este diseño se acompañará de la estimación presupuestaria de todo el proyecto.
- La selección de indicadores de evaluación cuantitativos y cualitativos, así como las diferentes técnicas que se llevarán a cabo en el trabajo de campo.

07

Fase 3. Gestionar y realizar el trabajo de campo

Propósito Aplicar los procedimientos de evaluación en los centros seleccionados y dejar constancia de su realización.

Cómo llevarla a cabo

En la tercera fase, se lleva a cabo todo el procedimiento de gestión y trabajo de campo. En primer lugar, se realiza el contacto con el personal responsable del programa en los centros y se programan las fechas de aplicación de las técnicas de evaluación y, en segundo lugar, se aplican dichos procedimientos en los centros seleccionados. La duración de esta fase varía en función de las técnicas de evaluación seleccionadas y de la muestra de centros que se haya seleccionado.

- Contacte con el personal responsable del programa en cada centro seleccionado.
- Programe las fechas de aplicación de las técnicas de evaluación acordadas.
- Aplique los procedimientos en los centros seleccionados y registre las incidencias que se produzcan.

Qué debe obtener

Un informe del trabajo de campo llevado a cabo, en el que se incluirán las fechas de realización de cada una de las técnicas de evaluación realizadas junto con las incidencias que hubiera habido en la aplicación de esta fase.

Organizar el informe de campo

Documente cada aplicación con las fechas, las técnicas realizadas y las incidencias. Esta información permite describir cómo se desarrolló el trabajo de campo.

- Fechas de realización de las técnicas de evaluación.
- Técnicas aplicadas en los centros seleccionados.
- Incidencias durante la aplicación.

Duración Ajuste el calendario a las técnicas seleccionadas y a la muestra de centros. La duración de esta fase depende de ambas decisiones.

08

Fase 4. Analizar y difundir los resultados

Propósito Fundamentar las conclusiones sobre el funcionamiento del programa y comunicar los resultados de la evaluación.

Cómo llevarla a cabo

Finalmente, se llevan a cabo los análisis de resultados, que estarán guiados por los parámetros o prerrogativas de nuestra propuesta de evaluación. Una vez llevados a cabo los análisis se podrá afirmar si (I) lo que prescribe el programa realmente ocurre y cómo (a través de la perspectiva de los participantes); (II) si ocurre en un grado de probabilidad aceptable conforme a los fines que se propone el programa, (III) y si todo lo anterior tiene lugar como consecuencia de la aplicación del programa o si se debe a la concurrencia de otros factores que han de tenerse en cuenta.

Preguntas que debe responder el análisis

- ¿Ocurre realmente lo que prescribe el programa y cómo lo explican sus participantes?
- ¿Ocurre en un grado de probabilidad aceptable conforme a los fines del programa?
- ¿Los cambios se deben a la aplicación del programa o a otros factores que deben tenerse en cuenta?

Qué debe obtener y difundir

- Informe de resultados de la evaluación llevada a cabo. En este punto se entiende positivo acordar el derecho de la explotación de los datos, citando su procedencia a partir del proyecto, con el objetivo de facilitar a la comunidad investigadora los hallazgos y generar conocimiento científico.
- Se pueden organizar jornadas en la que se explicarán los resultados de la evaluación llevada a cabo y se organizará la difusión de resultados para diferentes medios de divulgación.

Utilizar los resultados

Oriente las conclusiones hacia la mejora del programa y la identificación de buenas prácticas que faciliten su escalabilidad. Cuando sea posible, examine en qué medida sus efectos pueden atribuirse a la acción del programa.

Referencia

López-Meseguer, R., & Valdés, M. T. (2020). La Evaluación Comprensiva de Programas Educativos: ¿Un Nuevo Paradigma Teórico? *Revista Iberoamericana De Evaluación Educativa*, *13*(2), 85-105. https://doi.org/10.15366/riee2020.13.2.005