



Diseño e implementación de evaluaciones de impacto

Curso a distancia sobre evaluación de impacto de políticas públicas en América Latina y el Caribe: intervenciones eficaces para el desarrollo social

14 de octubre de 2020

Agustina Hatrick (ahatrick@caf.com)

¿Qué es una evaluación de impacto?

- Es una metodología de evaluación que permite identificar los cambios sobre individuos, familias o instituciones causados por un programa, iniciativa, política pública o proyecto específico.
- Las evaluaciones de impacto nos ayudan a atribuir **causalidad** al establecer empíricamente en qué medida cierto programa, y solo ese programa, ha contribuido a cambiar un resultado.

La clave es que
permite



- Cuantificar los cambios
- Aislar otros factores que podrían tener también **influencia** en los beneficiarios del programa

¿Qué es una evaluación de impacto?

- El **impacto de un programa** es la diferencia en la variable de resultado para la misma unidad (persona, hogar, comunidad, etc.) cuando ha participado en el programa y cuando no lo ha hecho.

Situación con programa

vs.

Situación sin programa

PREGUNTA CENTRAL

¿Qué hubiese pasado con aquellos que recibieron el programa si no lo hubiesen recibido?



CONTRAFACTUAL

Escenario hipotético que representa **qué le hubiera ocurrido** a la unidad de análisis en **ausencia** de la intervención.

Gran reto de la evaluación de impacto

Construir un grupo de comparación para reconstruir qué les hubiese sucedido a los beneficiarios del programa en ausencia de este.



**Grupo de
tratamiento**



vs.

**Grupo de
control**



¿Por qué hacer una eval. de impacto?

- **Evidencia rigurosa:** contar con evidencia confiable y sólida sobre qué funciona, qué no, y por qué
- **Redirección de presupuesto:** sabiendo qué funciona se puede redirigir el presupuesto en forma eficiente (dejar de invertir en lo que no tiene impacto)
- **Escalar políticas:** proporcionar evidencia atrae inversiones (organismos internacionales, gobiernos nacionales, provinciales o locales) y genera consensos para ampliar el alcance de los programas
- **Sostenibilidad:** tener evidencia permite que los programas que son efectivos sean sostenibles (no estén sujetos a los vaivenes de la política)

¿Cuándo hacer una eval. de impacto?

- Hay una pregunta relevante de política pública que necesita ser respondida.
- No hay suficiente evidencia sobre la efectividad de la intervención o la evidencia proviene de otros contextos (falta de validez externa).
- No es ni muy pronto ni muy tarde: esperar a que el programa madure, pero antes de escalarlo. El programa debe estar maduro, no perfecto.
- Se dispone del tiempo, conocimiento, recursos financieros y capacidades para hacer una evaluación de calidad.
- Es metodológicamente factible.

Evaluación del modelo COAR



Evaluación de impacto de los Colegios de Alto Rendimiento (COAR)

Los COAR son un modelo de servicio educativo diseñado e implementado por el Ministerio de Educación del Perú cuyo objetivo es proporcionar un servicio educativo con altos estándares a estudiantes vulnerables.

El modelo



- Sistema de **escuelas internado dirigido a estudiantes de alto desempeño académico** de escuelas públicas.
- 3° a 5° año de secundaria.
- Sigue el currículo nacional + Bachillerato Internacional.
- Enfoque de desarrollo integral.
- Son gratuitos: incluyendo matrícula, vivienda, alimentación, vestimenta, laptop, libros y otros materiales.
- Administrado por el Ministerio Nacional de Educación.

¿Por qué evaluar los COAR?



Rápido crecimiento de los COAR a nivel nacional.



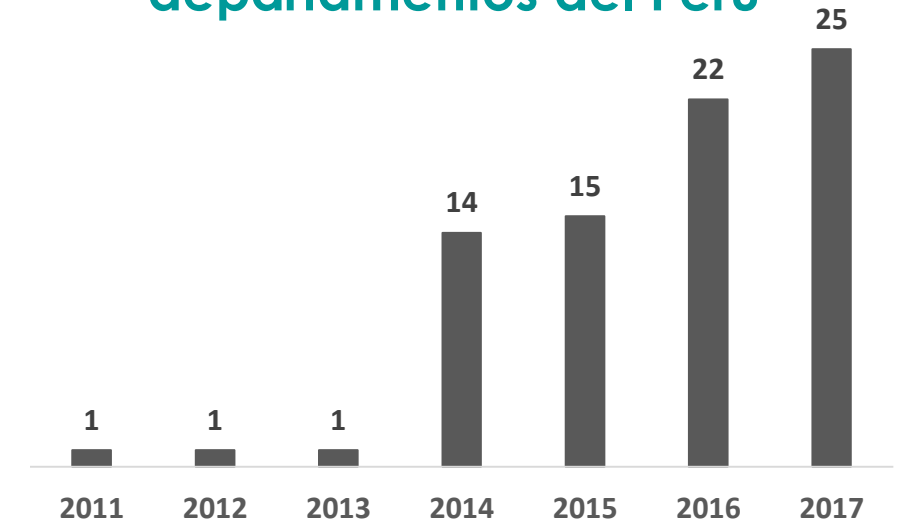
Inversión anual por estudiante es 3,5 veces mayor a la inversión en la Educación Básica Regular*.



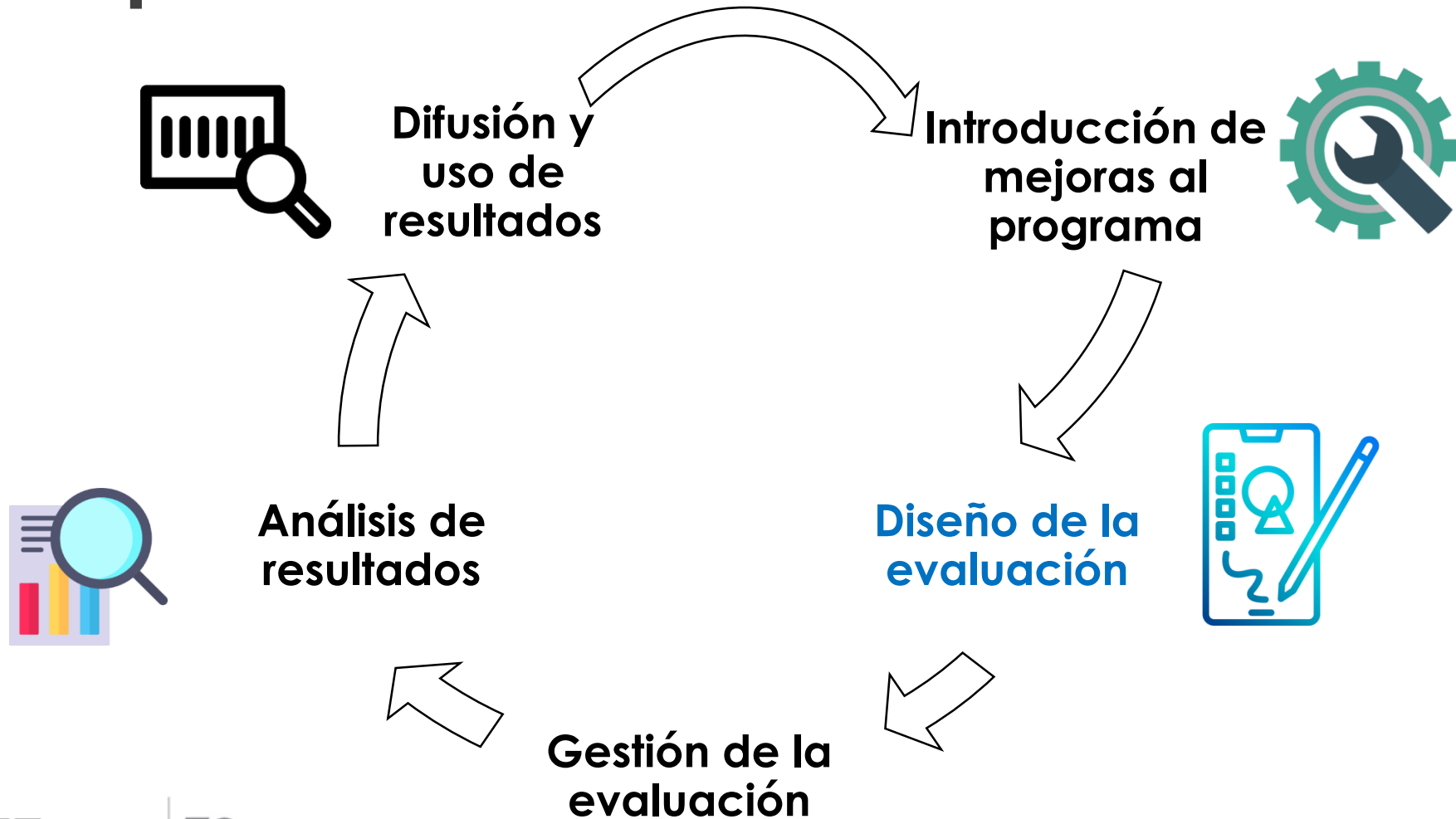
Evidencia mixta de la efectividad de los colegios de alto rendimiento.

*Fuente: información provista por MINEDU.

Hoy en día el sistema COAR tiene presencia en todos los departamentos del Perú

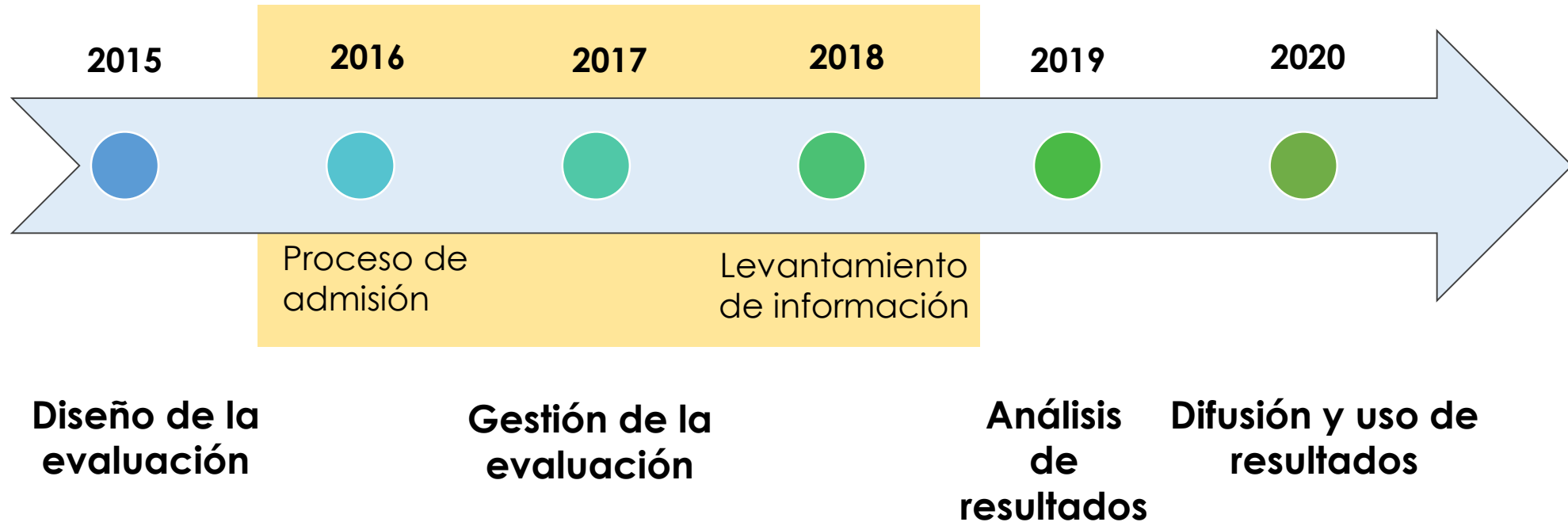


Etapas de una evaluación



Línea de tiempo

Implementación COAR-
Cohorte 2016-2018



Diseño de una evaluación

- 1) Definir actores clave, equipo de evaluación, roles y responsabilidades.



Dirección de Evaluación de Impacto y Aprendizaje de Políticas



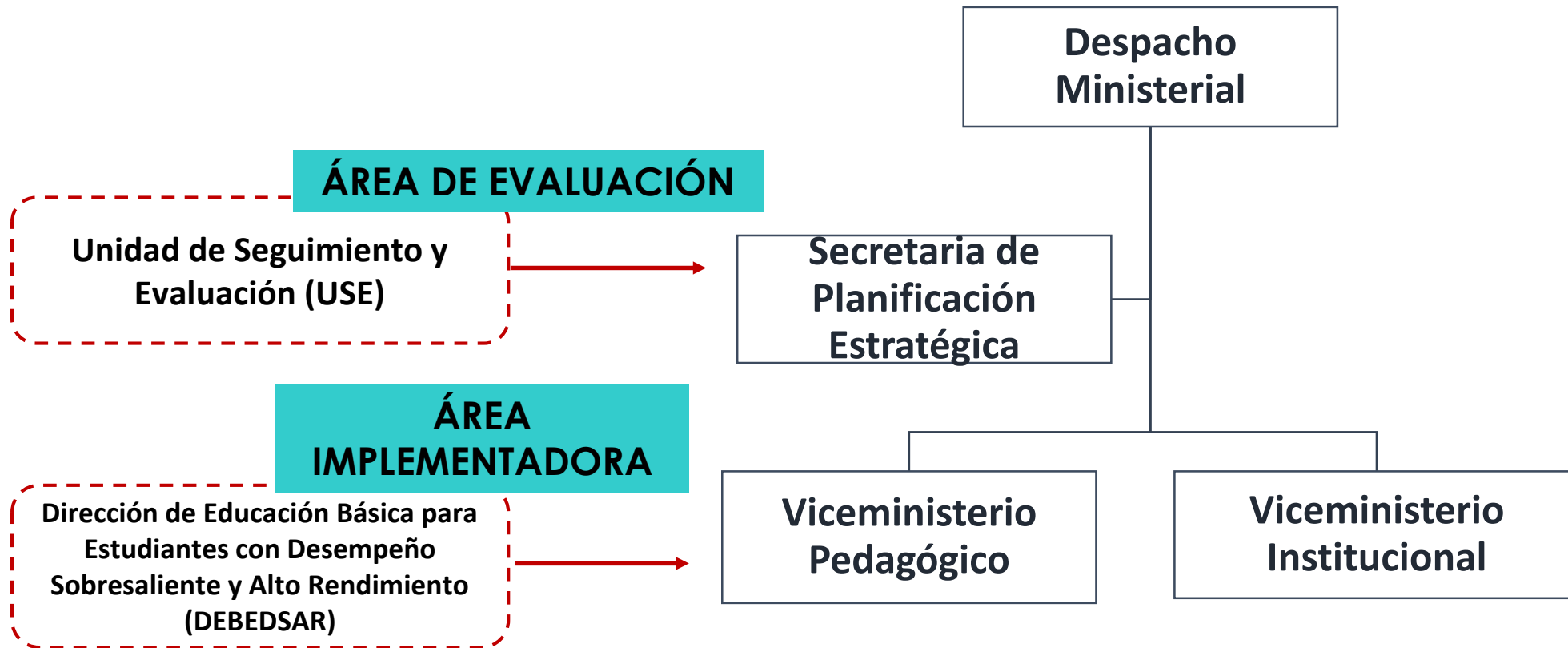
Unidad de Seguimiento y Evaluación (USE)

Dirección de Educación Básica para Estudiantes con Desempeño Sobresaliente y Alto Rendimiento (DEBEDSAR)



Consultoras externas

MINEDU



Desafíos

- **Plazos de la evaluación y clima de inestabilidad.** Durante el período de evaluación, hubo una serie de acontecimientos políticos que afectaron al sector educación:
 - huelga docente en 2017
 - alta rotación de ministros y de equipos dentro del Ministerio
- ✓ Compromiso institucional del MINEDU con la evaluación de políticas públicas.
- ✓ Importancia del equipo de CAF y del contacto continuo con la USE.
- ✓ Importancia de la elaboración de Nota Metodológica.



Desafíos

- **Alcance la evaluación.** Coordinación de diversos actores y alcance a nivel nacional.
 - ✓ Importancia de la contraparte técnica dentro del MINEDU → USE
 - ✓ Importancia del vínculo con el área implementadora (DEBEDSAR) y su participación durante toda la evaluación.

Diseño de una evaluación

- 2) Definir el tratamiento y población objetivo.
 - 3) Construir la [teoría del cambio](#) del programa.
 - 4) Definir preguntas de evaluación y variables de interés que sean relevantes para la política pública.
 - 5) Determinar la disponibilidad de información:
 - ¿Datos administrativos, encuestas, censos?
 - ¿Datos antes de la intervención, datos después de la intervención?
 - ¿A qué nivel de agregación?
 - 6) Identificar necesidad y factibilidad de recoger datos primarios:
 - Definir el tamaño de la muestra en función del **cálculo de poder** y presupuesto disponible.
-

Diseño de una evaluación

Variables de interés

- Capacidades básicas en matemática y comprensión lectora
- Habilidades socioemocionales
- Estrés académico y hábitos de estudio
- Expectativas académicas y laborales
- Hábitos saludables
- Estudios superiores

Disponibilidad de información

- Uso de datos administrativos (evaluación censal de estudiantes, censo educativo, matrícula, matriculación y admisión a universidades).
- Oportunidades de recolección de datos durante el proceso de admisión: ficha socioeconómica
- Levantamiento de datos primarios luego de la implementación del programa (fines de 2018)



Diseño de una evaluación

6) Seleccionar la metodología de evaluación más adecuada.

Reglas de
asignación del
programa

Disponibilidad de
datos

Factibilidad política y
logística

Son los elementos clave para determinar la **metodología de evaluación**.

- Evaluación experimental (RCT)
- Variables instrumentales (IV)
- Regresión discontinua (RD)
- Diferencias en diferencias (Diff-in-diff)
- Matching o emparejamiento
- Diferencias en diferencias emparejadas

Resumen de métodos de evaluación

Método	Reglas de asignación del programa	Grupo de comparación	Condiciones
Evaluación experimental	Asignación aleatoria	Unidades asignadas aleatoriamente al grupo de control	Balance de características observables y cumplimiento de la asignación
Variables instrumentales	Asignación aleatoria con cumplimiento imperfecto	Unidades asignadas aleatoriamente al grupo de control	Instrumento relevante y exógeno
	Asignación no aleatoria	Unidades que no recibieron el programa	
Regresión discontinua (sharp o fuzzy)	Asignación en base a un índice/puntaje y un punto de corte	Unidades con un puntaje apenas inferior (o superior) al punto de corte	Discontinuidad en la probabilidad de recibir el tratamiento; no manipulación del puntaje; balance de características observables alrededor del punto de corte
Diferencias en diferencias	Experimento natural o asignación no aleatoria	Unidades que no recibieron el programa	Tendencias paralelas
Matching	Asignación no aleatoria	Unidades que no recibieron el programa	Selección por observables

Resumen según disponibilidad de info

		Experimento	RD	Dif-en-Dif	Matching
Participación (o asignación al programa)		Indispensable			
Variable de corte			Indispensable		
Resultados	Después	Indispensable			
	Antes	Útil	Útil	Indispensable	Útil
Características de participantes y no participantes		Útil	Útil	Útil	Indispensable

Resumen de métodos de evaluación

- Una evaluación tiene **validez interna** si proporciona una estimación precisa del contrafactual mediante un grupo de comparación válido.
- Una evaluación tiene **validez externa** si los resultados se pueden generalizar a otras situaciones y/o poblaciones.

Resumen según validez interna y externa

		Credibilidad de la relación causal (validez interna)			
		Alta	Media	Baja	Nula
Validez externa	Efectos promedio para toda la población	RCT (experimental)	Dif-en-Dif	Matching (o propensity score matching)	<ul style="list-style-type: none"> - Comparación Antes/Después - Comparación Tratados/No Tratados (cuando los grupos NO se eligieron aleatoriamente)
	Efectos locales	RD			

Admisión a COAR

- **Requisitos de elegibilidad:**
 - Estudiar los primeros dos años de la secundaria en una escuela pública y edad máxima de 15 años.
 - 3 mejores estudiantes en la escuela de origen en el segundo grado o haber alcanzado uno de los mejores 5 puestos en una competencia nacional.
 - Promedio de al menos 15 sobre 20 al final de la escuela secundaria.
- **Proceso de postulación:** formulario de inscripción, autorización de los padres, examen estandarizado, jornada vivencial y entrevista.
- **Proceso de admisión:**
 - Puntaje final de admisión es la suma del puntaje de un examen estandarizado (40%), ensayo escrito (10%), jornada social (30%) y entrevista (20%)
 - Cuotas de ingreso en cada uno de los 25 departamentos de Perú.
 - Aspirantes con mejores puntajes son admitidos a COAR.

Selección de metodología de eval.

- La regla de asignación genera discontinuidades en la admisión a COAR a nivel departamental que pueden ser explotadas para medir el impacto de haber sido admitido al sistema COAR.

Diseño de Regresión Discontinua

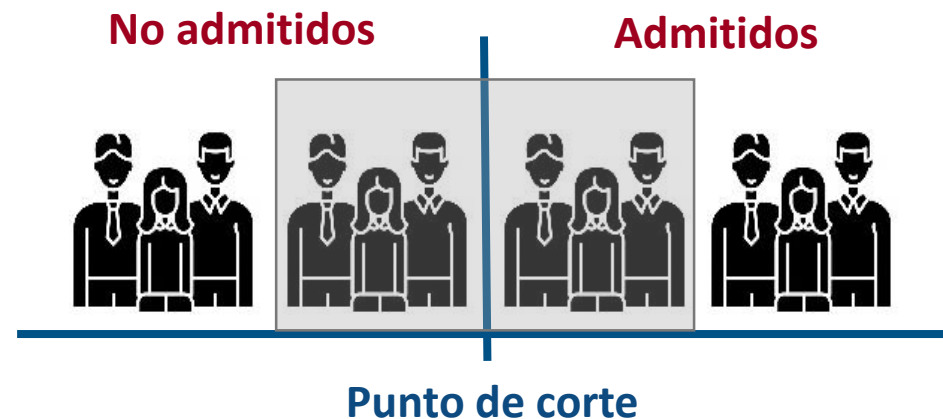
Grupo de control

Alumnos marginalmente no admitidos a COAR y que continuaron sus estudios en la EBR

vs.

Grupo de tratamiento

Alumnos marginalmente admitidos a COAR



Gestión de una evaluación

- 7) En caso de utilizar datos primarios, diseñar los instrumentos de medición y ponerlos a prueba (grupos focales, entrevistas, piloto).
- 8) Planificar el levantamiento de información.
 - Selección de la institución que se encargará de recoger la información.
 - Organización del trabajo de campo
 - Selección del equipo de campo
 - Capacitación del equipo
- 9) Realizar los levantamientos de información planificados.
 - Supervisión y control (monitoreo) del trabajo de campo durante el levantamiento
- 10) Monitorear la implementación para tener información sobre cómo funciona el programa e identificar posibles fallas que limiten el impacto de los resultados.

Gestión de una evaluación

- 11) Monitorear los siguientes factores que pueden **amenazar la validez interna** de la evaluación:
- Externalidades: si el programa afecta de manera indirecta al grupo de control, el grupo de control deja de ser un buen contrafactual.
 - Desgaste de la muestra: no contamos con datos de seguimiento para toda la muestra de estudio.
 - Cumplimiento imperfecto: las unidades del grupo de tratamiento no reciben el programa y/o las unidades del grupo de control sí lo reciben
 - Efectos inducidos por la evaluación y no por el tratamiento

Gestión de una evaluación

Diseño de instrumentos

- Mesas de discusión entre CAF y equipo implementador. Importancia del área implementadora (DEBEDSAR).
- Rondas de revisión de instrumentos.
- Piloto de los instrumentos y ajustes.

Diseño del levantamiento de información

- Diseño del levantamiento de información (COAR vs colegios EBR).
- Diseño de incentivos para evitar desgaste de la muestra: sorteo de 20 tablets.

Trabajo de campo

- Dimensionalidad del trabajo de campo: 1117 escuelas visitadas en todo el país (1095 escuelas de la EBR y 22 COAR).
 - Trabajo logístico previo.
 - Monitoreo del trabajo de campo.
-



PERÚ

Ministerio
de Educación

Estudio diagnóstico a estudiantes de 5° año de secundaria



BANCO DE DESARROLLO
DE AMÉRICA LATINA

Estimado/a estudiante:

En el marco del estudio diagnóstico a estudiantes de 5to año, te haremos algunas preguntas para saber más sobre ti, lo que piensas o sientes. No existen respuestas correctas o incorrectas, solo responde con sinceridad.

Recuerda que las respuestas son confidenciales y en ningún caso será revelada tu identidad. La información recopilada será solo utilizada en nuestro proyecto con fines de investigación.

Escribe con MAYÚSCULAS en los espacios asignados a cada pregunta y **marca con X** donde corresponda. Si tienes alguna pregunta, levanta la mano y con gusto te ayudaremos.

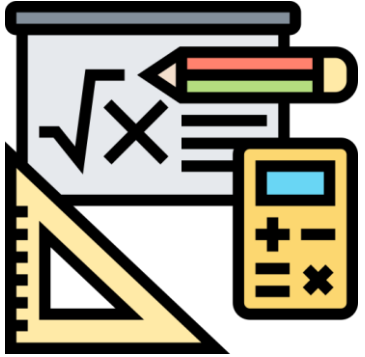
¡Gracias por ayudarnos a contestar esta encuesta!

Te recordamos que estaremos sorteando 20 tablets entre los alumnos que completen la encuesta. Por ello te pedimos que, al final de la misma, nos proporciones tus datos de contacto y correo electrónico

Principales resultados COAR

- Impactos nulos en desempeño académico (matemática y comprensión lectora) así como tampoco en habilidades socioemocionales (GRIT, autoeficacia, autosuficiencia, liderazgo, estrés académico).
 - Efectos positivos (y grandes) en **la expectativa de continuar estudiando en la universidad**
 - En el corto plazo (6 meses después del egreso de COAR), **mayor postulación a universidad**, pero no aun en matriculación. Los alumnos tienden a postularse más a carreras STEM.
 - Menor agresión verbal y menor discriminación.
-

Difusión de los resultados: desafíos



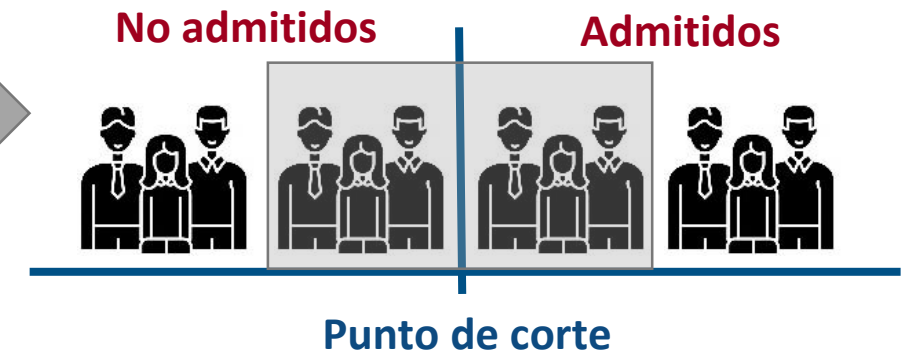
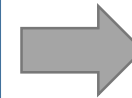
- **Resultados menos favorables de lo esperado:** No se identificaron impactos en la mayoría de indicadores relevantes de los COAR (aprendizajes, habilidades blandas, entre otros).
 - ✓ Diversos indicadores evaluados y análisis implementados
 - ✓ Diseño e implementación participativo: retroalimentación continua
 - ✓ Importancia del estudio cualitativo para complementar e interpretar los resultados
 - ✓ Presentación a diversos actores del MINEDU y académicos adaptando el contenido según la audiencia

Difusión de los resultados: desafíos

- Importancia de transmitir el alcance de los resultados.

El experimento es local

Los resultados son válidos para los alumnos con las calificaciones más bajas entre los aspirantes admitidos a COAR y los alumnos con las calificaciones más altas entre los no admitidos.



Si se quisiera evaluar el impacto de COAR para el total de alumnos, debería aplicarse un proceso de selección aleatoria.

Uso de los resultados: desafíos

- Seguimiento por parte de la USE a las recomendaciones que surgieron de la evaluación y análisis complementarios propuestos
- Análisis de toda la evidencia disponible sobre los COAR
- Ciclo de diseño de política y presupuestal
- Refuerzo de la cultura de evaluación y generación de evidencia para la toma de decisiones

¡Muchas gracias!

ahatrack@caf.com



www.caf.com
[@AgendaCAF](https://twitter.com/AgendaCAF)

Teoría del cambio

La teoría de cambio de un programa se podría definir como la *guía* que nos indica hacia **dónde** vamos (resultados) y **cómo** llegamos a nuestra meta (procesos).



Pregunta: ¿Cómo espero que se produzcan los impactos?

Respuesta: *Si* [insumos] y [actividades] *producen* [productos], esto debería *conducir a* [resultados] que al final *contribuirán al* [impacto].

Teoría del cambio

La teoría del cambio permite:

- Mapear la cadena de causalidad, partiendo de los insumos de la intervención y finalizando con el impacto.
- Establecer los supuestos que deben cumplirse para que el programa logre su objetivo.
- Definir indicadores para medir cada paso de la cadena causal.
 - Importancia del monitoreo de la intervención
- Si el programa no produce los impactos esperados, determinar en qué parte de la cadena causal se produjo el fallo y por qué.