

# Evaluación de Impacto: Transformando Programas Sociales<sup>1</sup>

Tessera LAB: Voces para Decidir

Nº 1

Junio de 2025

## Introducción

En las últimas décadas se ha incrementado notablemente la atención hacia la evaluación de impacto de programas sociales, tanto en el sector público como en el mundo privado y el tercer sector. La evaluación de impacto se refiere, en términos generales, a la medición rigurosa de los cambios atribuibles a una intervención o programa social, diferenciando dichos cambios de lo que habría ocurrido en ausencia del programa (Gertler *et al.*, 2016). Esta práctica se ha vuelto fundamental en un contexto donde los recursos para el desarrollo social son escasos y existe una creciente demanda de evidencia objetiva sobre la efectividad de las intervenciones (Vara-Horna, 2007; EVPA, 2015). Organismos gubernamentales, empresas con programas de responsabilidad social, fundaciones filantrópicas y organizaciones no gubernamentales ahora buscan no solo implementar proyectos sociales, sino también demostrar con evidencia que estos generan los resultados e impactos deseados en las poblaciones objetivo (Banerjee & Duflo, 2011). En este documento se explora la importancia conceptual de la evaluación de impacto, se revisan nuevas tendencias metodológicas y se compilan buenas prácticas internacionales, con un enfoque tanto en programas estatales como en iniciativas impulsadas desde el sector privado o alianzas público-privadas.

---

<sup>1</sup> Evaluación de Impacto: Transformando Programas Sociales © 2025 by María Paz Donoso is licensed under Creative Commons Attribution 4.0 International. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>



## La importancia de la evaluación de impacto en programas sociales

La **evaluación de impacto** cumple un rol crítico en el ciclo de vida de los programas sociales, aportando información valiosa para diversas finalidades: accountability (rendición de cuentas), aprendizaje y mejora continua, y toma de decisiones basada en evidencia. A continuación, se detallan estas dimensiones de importancia:

- **Formulación de políticas basadas en evidencia:** A nivel macro, la institucionalización de la evaluación de impacto fortalece una cultura de toma de decisiones informada por datos. Gobiernos y agencias internacionales han promovido crecientemente la “política social basada en evidencia”, integrando hallazgos de evaluaciones en el diseño de nuevas políticas y programas (Gertler *et al.*, 2016). Por ejemplo, la adopción global de ciertos programas exitosos se ha basado en la evidencia de evaluaciones piloto. Un caso emblemático es el de los programas de transferencias condicionadas, cuyo modelo se expandió a múltiples países en desarrollo tras demostrarse mediante evaluaciones rigurosas su efectividad para reducir la pobreza y mejorar indicadores de salud y educación en las poblaciones beneficiarias (Fiszbein & Schady, 2009). Del lado del sector privado y filantrópico, existe una tendencia similar: las empresas y fundaciones no solo emprenden proyectos sociales por filantropía o valor reputacional, sino que buscan maximizar el retorno social de la inversión. Herramientas como el cálculo del *Social Return on Investment (SROI)* permiten cuantificar en términos monetarios amplios los beneficios sociales generados por cada dólar invertido (Nicholls *et al.*, 2012), lo que facilita comparar iniciativas y escalar las más costo-efectivas. En suma, la evaluación de impacto aporta la evidencia empírica que guía tanto políticas públicas como estrategias de responsabilidad social corporativa, orientándolas hacia intervenciones de alto impacto.
- **Rendición de cuentas y transparencia:** Los financiadores, ya sean organismos públicos, donantes internacionales, inversionistas sociales o directorios corporativos, exigen cada vez más una justificación basada en resultados de las inversiones sociales realizadas. La evaluación de impacto ofrece una forma objetiva de demostrar si un programa *realmente mejora* las condiciones de vida de la población meta. Según la European Venture Philanthropy Association, la medición y gestión del impacto se



considera hoy “el estándar de oro para la transparencia y la rendición de cuentas”, ayudando a identificar qué funciona y qué no en función de impulsar el cambio social (EVPA, 2015). En el ámbito de la inversión de impacto, un estudio global reportó que 94% de los inversionistas considera esencial medir el impacto social de sus inversiones (GIIN, 2020), lo que refleja cómo la evidencia de impacto se ha vuelto central para mantener la confianza de stakeholders y el apoyo sostenido a las iniciativas sociales.

- **Asignación eficiente de recursos y mejora de programas:** Dado que los recursos para políticas y programas sociales son limitados, es vital orientarlos hacia intervenciones que *funcionan*. La evaluación de impacto permite aislar el efecto causal del programa sobre resultados de interés (por ejemplo, reducción de la pobreza, mejora educativa o acceso a salud) controlando por otros factores, de modo que los decisores puedan distinguir entre programas efectivos y aquellos que no generan los cambios esperados (Gertler *et al.*, 2016). Esto contribuye a una asignación más racional y eficiente de fondos públicos o privados, priorizando las estrategias con mayor impacto (Vara-Horna, 2007). Además, al identificar impactos no deseados o insuficientes, las evaluaciones proporcionan evidencia para rediseñar o corregir los programas, mejorando su eficacia en ciclos futuros. En otras palabras, la evaluación rigurosa no solo juzga el éxito o fracaso de una intervención, sino que aporta retroalimentación para el aprendizaje organizacional, permitiendo refinar teorías de cambio y modelos de implementación (Patton, 2011).
- **Legitimidad y apoyo a largo plazo:** Un beneficio adicional de las evaluaciones de impacto es que fortalecen la legitimidad de los programas sociales ante la opinión pública y las comunidades atendidas. Cuando los resultados muestran impactos positivos claros (por ejemplo, mejoras significativas en ingreso, salud o seguridad para los beneficiarios), se genera un respaldo político y social que ayuda a institucionalizar o dar continuidad a la iniciativa (Kvam, 2018). Incluso se fomenta la apropiación local: los actores comunitarios pueden percibir que el programa realmente marca una diferencia, lo que aumenta su involucramiento y sostenibilidad en el tiempo (Kvam, 2018). En contraste, la ausencia de evaluaciones independientes puede derivar en escepticismo o en la continuidad de proyectos ineficaces por inercia. Por ello, incorporar evaluaciones de impacto con rigor científico (p. ej., con grupos de control aleatorizados u otros diseños cuasi-experimentales) se ha convertido en parte de las



buenas prácticas de gobernanza en el campo social, garantizando que las decisiones se basen en méritos técnicos y resultados demostrados (GAO, 1978; EVPA, 2015).

## Nuevas tendencias en la evaluación de impacto social

La práctica de la evaluación de impacto de programas sociales se encuentra en constante evolución. En los últimos años han emergido nuevas tendencias y enfoques metodológicos que amplían las herramientas disponibles para evaluadores y gestores, respondiendo tanto a desafíos clásicos (como atribuir causalidad) como a nuevas prioridades (por ejemplo, incorporar consideraciones de equidad o aprovechar grandes volúmenes de datos). A continuación, se describen algunas tendencias destacadas:

- **Enfoques experimentales y “randomistas” en auge, con matices:** Desde comienzos del siglo XXI se consolidó la evaluación experimental (especialmente mediante ensayos aleatorizados controlados o RCTs, por sus siglas en inglés) como “estándar de oro” para medir impacto en programas sociales. Organizaciones como J-PAL y 3ie han impulsado cientos de evaluaciones experimentales en diversos países, generando evidencia sobre qué intervenciones funcionan en educación, salud, inclusión financiera, etc. (Banerjee & Duflo, 2011). Esta tendencia ha ganado tal prominencia que en 2019 el Premio Nobel de Economía fue otorgado a tres economistas por su trabajo en la aplicación de RCTs a la lucha contra la pobreza, demostrando la efectividad de micro-intervenciones como tutorías escolares o tratamientos de salud preventiva (Banerjee *et al.*, 2019). No obstante, junto al auge de los RCTs, ha crecido también una apreciación de sus límites y condiciones de uso. Nuevas corrientes abogan por enfoques más contextualizados: por ejemplo, la evaluación por comparaciones sucesivas o estudios de caso comparativos como alternativas cuando los RCT no son factibles o éticos (BetterEvaluation, 2014). En resumen, la comunidad evaluativa reconoce la importancia de métodos rigurosos de identificación causal, pero complementados con otros diseños (diseños cuasi-experimentales, análisis longitudinales, métodos cualitativos) que permitan entender *cómo* y *por qué* un programa produce (o no) ciertos impactos, más allá de la simple comprobación de *si* los produce.



- **Mayor integración de metodologías mixtas (mix de cuantitativo y cualitativo):**  
Relacionado con lo anterior, existe consenso en que las evaluaciones de impacto más útiles son aquellas que integran diferentes fuentes y tipos de datos (*mixed methods*) para ofrecer una visión integral. Por un lado, los métodos cuantitativos proporcionan mediciones objetivas del cambio en indicadores clave (ingresos, tasas de escolaridad, mortalidad, etc.) atribuible al programa. Por otro lado, los métodos cualitativos (entrevistas, grupos focales, estudios de caso etnográficos) aportan matices y comprensión profunda de los procesos de cambio social, las percepciones de los beneficiarios y las razones detrás del éxito o fracaso de la intervención (Gertler *et al.*, 2016). La tendencia reciente enfatiza que no se debe concebir la evaluación de impacto solo como un ejercicio estadístico, sino como un proceso de investigación social holística. De hecho, muchos marcos contemporáneos hablan de combinar evaluaciones de impacto cuantitativas con evaluaciones de proceso o de implementación, de modo que los resultados numéricos puedan interpretarse adecuadamente en su contexto (Weiss, 1998; Rogers, 2014). La triangulación de metodologías aumenta la validez de las conclusiones y produce recomendaciones más accionables. Por ejemplo, un estudio de impacto puede revelar que un programa de capacitación laboral no mejoró significativamente los ingresos de los participantes; los datos cualitativos podrían explicar que, si bien la capacitación fue útil, barreras estructurales (p.ej., falta de transporte o discriminación) impidieron a algunos beneficiarios acceder a mejores empleos. Tales insights cualitativos son esenciales para ajustar el diseño del programa o complementarlo con otras políticas.
- **Enfoques teóricos y de “causalidad contribucional”:** Otra tendencia en el ámbito evaluativo es el desarrollo de enfoques basados en teorías de cambio y análisis de contribución, en contraste con la visión puramente experimental. Dado que muchos programas sociales operan en contextos complejos donde múltiples factores influyen en los resultados, se promueven métodos que documenten la lógica causal completa del programa. Herramientas como el mapeo de teoría de cambio y el análisis de contribución (Mayne, 2012) permiten atribuir resultados a la intervención incluso sin un contrafactual estadístico perfecto, mediante la construcción de una narrativa causal sustentada en evidencia. Estas aproximaciones teóricas son especialmente útiles cuando no es viable o ético establecer grupos de control aleatorios. Asimismo, ha cobrado fuerza la evaluación realista (Pawson & Tilley, 1997), la cual se pregunta “¿qué funciona, para quién, en qué contexto y por qué?”, reconociendo que el impacto de un



programa puede variar según características locales o poblacionales. La adopción de estos enfoques responde a la necesidad de comprender mecanismos subyacentes y condiciones habilitantes de los impactos, más allá de la simple constatación de efectos promedio. En la práctica, muchas evaluaciones actuales combinan un *núcleo cuantitativo* (que estima el efecto neto) con un *análisis teórico cualitativo* (que examina el proceso causal), proporcionando así recomendaciones más finas sobre cómo mejorar o escalar el programa.

- **Uso de tecnología, *big data* e inteligencia artificial:** La revolución digital ha alcanzado también al campo de la evaluación. Una tendencia emergente es aprovechar nuevas fuentes de datos masivos (lo que se denomina *Big Data*) y herramientas analíticas avanzadas para medir impactos de forma más eficiente o captar dimensiones antes inaccesibles (Bamberger, 2016). Por ejemplo, el análisis de datos de telefonía móvil, transacciones digitales o imágenes satelitales puede revelar cambios en bienestar económico, movilidad o condiciones ambientales ligados a una intervención, complementando los datos tradicionales de encuestas (York *et al.*, 2020). La inteligencia artificial y el *machine learning* se están utilizando para procesar grandes conjuntos de datos y detectar patrones de impacto en tiempo real (Hejnowicz & Chaplowe, 2021). Estas técnicas permiten, por ejemplo, monitorear continuamente la ejecución de un programa mediante tableros de visualización en tiempo real, o predecir qué beneficiarios tienen mayor probabilidad de experimentar cierto resultado, posibilitando intervenciones preventivas. Otra innovación tecnológica es el uso de encuestas móviles y plataformas digitales para recopilar datos de beneficiarios de manera más ágil y económica (p. ej., encuestas por SMS o aplicaciones). Esto facilita realizar evaluaciones de alcance amplio o en geografías dispersas a un costo menor. Si bien el potencial del big data es enorme, también plantea retos en términos de privacidad, calidad de datos y capacidad analítica. No obstante, la tendencia general es que los evaluadores incorporen progresivamente habilidades de ciencia de datos y colaboren con especialistas en datos para enriquecer los enfoques tradicionales (Bamberger, 2016). En resumen, la tecnología está transformando la forma de recolectar, analizar y presentar la evidencia de impacto, haciéndola más rápida, granular e interactiva.
- **Énfasis en la participación, la equidad y los enfoques centrados en el sujeto:** Paralelamente a las innovaciones técnicas, en la última década ha cobrado fuerza una



orientación normativa: las evaluaciones deben ser más inclusivas y sensibles a contextos socioculturales. Esto se refleja en varias sub-tendencias: por un lado, las evaluaciones participativas, donde se involucra a los beneficiarios y comunidades en el diseño de la evaluación, la identificación de indicadores relevantes y la interpretación de los hallazgos (Cousins & Whitmore, 1998). Esto promueve la pertinencia de la evaluación y empodera a los actores locales, asegurando que los resultados respondan a sus preocupaciones y generen aprendizaje local. Por otro lado, está el auge de las evaluaciones con enfoque de equidad y derechos humanos (Bamberger & Segone, 2011). UNICEF y otros organismos han impulsado guías para evaluaciones focalizadas en equidad, que examinan específicamente cómo un programa beneficia (o podría perjudicar) a los grupos más vulnerables y marginados, y si se están cerrando brechas de desigualdad. Este enfoque amplía la definición de “éxito” del programa: no basta con generar un impacto promedio positivo, sino que importa quiénes se benefician y si la intervención está alcanzando a los más necesitados. Asimismo, la agenda de género ha permeado el campo evaluativo, dando lugar a evaluaciones de impacto con perspectiva de género, que analizan diferenciadamente los efectos en mujeres, hombres y diversidades, e incluyen indicadores de empoderamiento, violencia de género, etc. (UN Women, 2015). En síntesis, las nuevas tendencias enfatizan que la evaluación de impacto debe incorporar criterios de justicia social, no solo medir “qué tanto cambió un indicador”, sino cómo se distribuyen esos cambios entre distintos grupos poblacionales y si se respetan valores éticos en el proceso evaluativo (p. ej., consentimiento informado, no maleficencia, uso responsable de datos personales).

- **Evaluación para la mejora continua: aprendizaje en tiempo real y *developmental evaluation*:** Tradicionalmente, la evaluación se concebía como un ejercicio *ex post* a la conclusión del programa, para juzgar su éxito. Hoy se observa un cambio hacia prácticas de evaluación formativa continua, donde la línea divisoria entre implementación y evaluación se difumina. El concepto de evaluación en tiempo real se refiere a sistemas de monitoreo-evaluación que van entregando resultados y análisis conforme el programa avanza, permitiendo a los gestores ajustar la estrategia sobre la marcha (Rogers, 2014). Un enfoque innovador alineado con esto es la evaluación evolutiva o *developmental evaluation* (Patton, 2011). Este tipo de evaluación se emplea en intervenciones complejas, innovadoras o de entornos muy dinámicos, donde el evaluador se integra al equipo de programa para proveer retroalimentación constante, ayudando a adaptar el proyecto a condiciones cambiantes y fomentando la innovación



continua. La evaluación evolutiva no busca tanto la atribución causal precisa, sino apoyar el desarrollo del programa para maximizar su impacto conforme aprende qué funciona. Este cambio de paradigma –de la evaluación como auditoría *post hoc* a la evaluación como componente inherente de la gestión adaptativa– es cada vez más notorio en organizaciones de desarrollo y filantropía que operan en contextos de incertidumbre (Rogers, 2014; Patton, 2011). La consecuencia es que los evaluadores asumen nuevos roles como *facilitadores de aprendizaje* y se requieren estructuras flexibles para permitir iteraciones rápidas basadas en evidencia emergente.

En conjunto, estas tendencias reflejan un campo de la evaluación de impacto en expansión, que adopta herramientas de otras disciplinas (estadística avanzada, ciencia de datos, ciencias sociales críticas) y se adapta a las demandas contemporáneas de pertinencia y velocidad. Para las empresas, fundaciones y entes públicos, mantenerse al día en estas tendencias es clave para diseñar evaluaciones más sólidas y útiles. No se trata solo de aplicar las técnicas más novedosas por moda, sino de contar con un arsenal metodológico amplio para abordar diferentes preguntas evaluativas en distintos contextos, siempre al servicio de mejorar la efectividad de las intervenciones sociales (Reynolds *et al.*, 2012). La innovación en evaluación, bien aplicada, permite superar limitaciones de enfoques tradicionales y descubrir maneras más efectivas de generar y medir el cambio social (Rogers, 2014).

## Buenas prácticas en la evaluación de impacto: lecciones y recomendaciones

A partir de la experiencia acumulada internacionalmente, han emergido una serie de buenas prácticas que guían la planificación y ejecución de evaluaciones de impacto de programas sociales. Estas prácticas ayudan a garantizar la calidad, fiabilidad y utilidad de las evaluaciones, aumentando la probabilidad de que sus resultados influyan positivamente en la toma de decisiones. A continuación se presentan algunas de las mejores prácticas ampliamente reconocidas, acompañadas de ejemplos donde corresponde:

1. **Integrar la evaluación desde el diseño del programa:** Una máxima moderna es "*pensar en la evaluación desde el día uno*". Las intervenciones sociales deben concebirse con una teoría de cambio explícita y un marco de resultados claro, lo que





facilita más adelante la evaluación (Weiss, 1998). Esto implica definir desde el inicio objetivos específicos y medibles, así como indicadores clave de desempeño asociados a esos objetivos (Wadhwani Foundation, 2022). Por ejemplo, un programa de capacitación laboral podría fijar como meta “aumentar en 50% la tasa de empleo formal de los participantes a los 12 meses”, y delinear indicadores como tasa de colocación, retención en el empleo y aumento salarial. Contar con estas metas SMART (específicas, medibles, alcanzables, relevantes y con plazo) proporciona un norte claro tanto para implementadores como evaluadores. Asimismo, es crucial establecer una línea de base antes de iniciar las acciones del programa (Wadhwani Foundation, 2022). Levantar datos iniciales de los indicadores pertinentes (por ejemplo, ingresos promedio de los beneficiarios antes de una intervención económica) permitirá luego comparar y atribuir cambios con mayor confianza. Integrar la evaluación en el diseño también supone prever recursos presupuestarios para las actividades de evaluación y, de ser posible, planificar un diseño evaluativo (experimental o cuasi-experimental) desde el arranque –por ejemplo, determinando aleatoriamente qué comunidades recibirán primero la intervención y cuáles después (grupos control), u otras estrategias que establezcan comparabilidad.

2. **Garantizar independencia, rigor técnico y ética en la evaluación:** La credibilidad de una evaluación de impacto depende de la objetividad e independencia con que se lleve a cabo. Una buena práctica es involucrar evaluadores externos o terceros independientes, o al menos, separar claramente el rol de quienes implementan y quienes evalúan, para evitar sesgos o conflictos de interés (OECD DAC, 2019). En cuanto al rigor técnico, se recomienda emplear los métodos de análisis más robustos disponibles dados los datos y contexto: esto puede significar usar técnicas econométricas apropiadas (regresión multivariante, análisis de sensibilidad, estimación de errores estándar robustos, etc.) y reportar no solo resultados puntuales sino también intervalos de confianza y tamaños de efecto. Asimismo, se deben documentar supuestos y limitaciones con transparencia (Gertler *et al.*, 2016). Por ejemplo, si se realiza una evaluación cuasi-experimental mediante pareamiento (*propensity score matching*), explicar claramente que la validez causal depende del supuesto de no haber sesgo no observado, etc. La ética es otra dimensión fundamental: toda evaluación que involucre personas debe cumplir estándares éticos, como obtener consentimiento informado de participantes encuestados, resguardar la confidencialidad de datos personales y evitar cualquier daño o negación injusta de



beneficios (Gertler *et al.*, 2016). Los comités de ética y revisión institucional (IRB) desempeñan un rol clave cuando se hacen experimentos con sujetos humanos, asegurando que los tratamientos evaluados no vulneren derechos (por ejemplo, si se retrasa un beneficio a un grupo control, justificar que esto no pone en riesgo su bienestar básico). En síntesis, calidad científica y ética deben ir de la mano; una evaluación técnicamente sofisticada pero conducida sin respeto a los participantes o con datos dudosos carecerá de legitimidad.

3. **Utilizar enfoques participativos e inclusivos:** Involucrar a los diversos stakeholders en el proceso evaluativo suele mejorar tanto la relevancia de la evaluación como la posterior adopción de sus recomendaciones. Una buena práctica es desarrollar un plan de participación donde se identifique quiénes son los interesados (beneficiarios, personal de campo, socios locales, financiadores) y cómo contribuirán. Por ejemplo, se pueden realizar *talleres con la comunidad beneficiaria* para definir qué resultados les importan (quizá valoran resultados cualitativos que los técnicos no habían previsto), o para validar y contextualizar los hallazgos antes de su difusión final (Cousins & Whitmore, 1998). La participación ayuda a que los indicadores seleccionados sean culturalmente pertinentes y a que los criterios de éxito integren la perspectiva de quienes viven la realidad evaluada. Esto enlaza con la equidad: asegurarse de incluir voces de mujeres, minorías u otros grupos vulnerables en la evaluación es esencial para no perpetuar sesgos. Un ejemplo de buenas prácticas viene de evaluaciones de proyectos de desarrollo rural, donde los equipos evaluadores convocan a asambleas comunitarias para discutir los cambios observados: este diálogo no solo legitima los resultados, sino que puede aportar evidencia cualitativa adicional (testimonios, historias de vida) que enriquece la comprensión del impacto (Esteves *et al.*, 2012). En cuanto a inclusión, otra práctica creciente es aplicar metodologías de evaluación sensible al género (involucrando a mujeres evaluadoras, asegurando espacios seguros para que beneficiarias den su opinión) y accesibles (p. ej., usar encuestas en el idioma local, o formatos adecuados para personas con bajos niveles de alfabetización). Todo ello redundará en una evaluación más justa y útil para todos.
4. **Definir indicadores y métricas significativas, equilibrando cuantitativo y cualitativo:** Seleccionar adecuadamente qué medir es crítico. Las buenas prácticas indican que los indicadores de impacto deben derivarse lógicamente de la teoría de cambio del programa, ser factibles de medir con precisión y realmente capturar los



cambios sustantivos esperados en la vida de las personas (Gertler *et al.*, 2016). Se recomienda evitar tanto los indicadores demasiado vagos o generales (que dificultan la medición y atribución) como aquellos demasiado estrechos que pueden *distorsionar* las prioridades del programa. Un balance razonable incluye indicadores de resultado inmediato (outputs), resultado intermedio (outcomes) e impacto final. Por ejemplo, en un programa educativo: *output* = número de docentes capacitados; *outcome* = mejora en prácticas pedagógicas en aula; *impacto* = aumento en el aprendizaje de los estudiantes medido en pruebas estandarizadas. Adicionalmente, combinar indicadores cuantitativos (p.ej., % de incremento en ingreso, puntaje promedio en test) con indicadores cualitativos o de *percepción* (p.ej., grado de satisfacción de beneficiarios, testimonios de empoderamiento) brinda un panorama completo (Wadhwani Foundation, 2022). Las encuestas de satisfacción y las historias de cambio significativo (*most significant change*) son herramientas válidas para capturar dimensiones menos tangibles pero relevantes del impacto. Otro aspecto de buenas prácticas es asegurarse de medir posibles efectos negativos o no esperados del programa, no limitarse solo a los objetivos positivos. Por ejemplo, si se evalúa un programa de microcrédito, además de medir el aumento en emprendimientos o ingresos, conviene indagar si hubo endeudamiento excesivo o estrés en los hogares beneficiarios (Vanclay, 2002). Una evaluación integral considera el impacto *neto*, tanto beneficios como costos sociales.

5. **Fortalecer los sistemas de datos y la calidad de la información:** La solidez de cualquier evaluación descansa en gran medida en la calidad de los datos recopilados. Una buena práctica es invertir en sistemas de monitoreo y evaluación (M&E) robustos dentro del programa, que alimenten la evaluación de impacto. Esto incluye capacitación al personal en recolección de datos, uso de herramientas digitales para minimizar errores (p.ej., encuestas con tablets y validación en tiempo real), y protocolos claros de control de calidad (revisiones, auditorías de datos, validación externa) (Wadhwani Foundation, 2022). Siempre que sea posible, se deben utilizar múltiples fuentes de datos para contrastar información. Por ejemplo, en la evaluación de un programa de capacitación laboral, los datos de encuestas a participantes pueden complementarse con registros administrativos (colocación laboral reportada por oficinas de empleo) y con información de terceros (opinión de empleadores sobre el desempeño de egresados, etc.) (Wadhwani Foundation, 2022). La convergencia de distintas fuentes aumenta la confiabilidad de los hallazgos. Otra recomendación es



aprovechar bases de datos existentes (*data mining* de censos, encuestas nacionales, registros de salud, etc.) que puedan servir de línea de base o grupo de comparación. En la actualidad, con la disponibilidad de datos abiertos y big data, los evaluadores tienen más oportunidades de enriquecer sus análisis, pero deben también asegurarse de la pertinencia y limpieza de esos datos externos (Bamberger, 2016). Finalmente, documentar cuidadosamente todos los pasos de manejo de datos (guías de código, diccionarios de variables, etc.) y abrir los datos una vez concluida la evaluación (salvaguardando privacidad) se está volviendo una buena práctica para promover transparencia y replicabilidad de la evaluación (Gertler *et al.*, 2016). Esto permite que la comunidad evaluadora o académica valide los hallazgos o realice análisis secundarios que extraigan todavía más lecciones del programa evaluado.

6. **Comunicación efectiva de resultados y uso de la evidencia:** Una evaluación de impacto solo genera valor si sus resultados se utilizan para la mejora de políticas o programas. Por ello, las buenas prácticas enfatizan la importancia de elaborar una estrategia de comunicación de los hallazgos adaptada a distintos públicos (Patton, 2011). Esto implica que, además del informe técnico extenso, se preparen *resúmenes ejecutivos claros*, con visualizaciones comprensibles, mensajes clave destacados y recomendaciones concretas para la acción. En contextos gubernamentales, suele ser útil presentar los resultados en reuniones de alto nivel con tomadores de decisión, enfocando la discusión en implicaciones prácticas (por ejemplo: “el programa A es más costo-efectivo que el B para lograr el objetivo X, por lo que se recomienda reasignar presupuesto en el próximo ciclo”). En el sector privado o fundaciones, puede significar presentar el impacto social logrado a directorios o donantes, evidenciando el “valor por dinero” y cómo se alinea con la misión institucional. Una técnica recomendada es contar historias de cambio ilustrativas junto con las estadísticas, para humanizar los resultados y mostrar casos de éxito o desafíos que resuenen con los responsables (BetterEvaluation, 2014). También es importante socializar los resultados con las comunidades beneficiarias, cerrando el ciclo participativo: por ejemplo, realizar asambleas locales donde se devuelvan los hallazgos en lenguaje sencillo, para rendir cuentas a la comunidad y discutir colaborativamente cómo mejorar el programa. En cuanto a la promoción del uso, se sugiere involucrar a los “usuarios primarios” de la evaluación desde el inicio (Patton, 2011). Si gestores y decisores sienten *propios* los resultados (por haber contribuido a las preguntas evaluativas y seguido el proceso), es más probable que implementen las recomendaciones. Crear un plan de acción post-



evaluación es una buena práctica: enumerar qué decisiones o ajustes se tomarán basados en la evidencia y quién será responsable de darles seguimiento. En resumen, difundir estratégicamente la evidencia y traducirla en decisiones cierra el círculo virtuoso de la evaluación, consolidando la cultura de aprendizaje.

7. **Calidad y estándares internacionales:** Por último, cabe mencionar que existen marcos y estándares de calidad que pueden guiar a las instituciones en la realización de evaluaciones de impacto. Organismos como la OCDE han establecido criterios de evaluación (relevancia, eficacia, eficiencia, impacto, sostenibilidad) que ayudan a enmarcar tanto la evaluación misma como la valoración de un programa (OECD DAC, 2019). Asimismo, la Red de Evaluación de la ONU (UNEG) y sociedades profesionales han desarrollado estándares de competencia y códigos de conducta para evaluadores que cubren integridad, competencia técnica, respeto por la dignidad de las personas, etc. (UNEG, 2020). Adherirse a estas buenas prácticas profesionales significa, por ejemplo, que los informes deben ser precisos, completos, equilibrados en presentar tanto éxitos como fracasos, y emitidos a tiempo para incidir en las decisiones (GAO, 1978). Varias agencias públicas en el mundo han institucionalizado sistemas de revisión de la calidad de las evaluaciones (peer review internas o externas) antes de su publicación, para garantizar que se cumplan tales estándares. La aplicación consistente de estas normas eleva la calidad global de las evaluaciones y facilita el reconocimiento mutuo de resultados entre organizaciones (pues todas confían en métodos rigurosos comparables). Para empresas y ONG que se inician en la cultura evaluativa, familiarizarse con guías y manuales publicados por el Banco Mundial, el BID, la OCDE, Naciones Unidas, etc., puede ser un buen punto de partida para internalizar qué se espera de una *buena* evaluación de impacto (Gertler *et al.*, 2016; Kvam, 2018).

En suma, las buenas prácticas descritas proveen un mapa para ejecutar evaluaciones de impacto que sean útiles, confiables y éticas. No todas las evaluaciones podrán implementar cada recomendación al pie de la letra debido a restricciones de contexto o recursos, pero tenerlas como referencia ayuda a elevar el estándar y evitar errores comunes. Para el sector privado, adoptar estas prácticas en la evaluación de programas sociales (por ejemplo, de inversión comunitaria o proyectos de sostenibilidad) puede mejorar significativamente la gestión del impacto corporativo, permitiendo a la empresa maximizar el valor social creado y comunicarlo mejor a sus grupos de interés. Para las



ONG y sector público, estas prácticas fortalecen la eficacia institucional y la capacidad de aprendizaje, factores clave para abordar problemas sociales complejos de manera adaptativa y basada en evidencia.

## Conclusiones

La evaluación de impacto de programas sociales se ha convertido en una pieza indispensable para asegurar la efectividad y el aprendizaje en las intervenciones orientadas al bien público. En un entorno donde tanto gobiernos como empresas y organizaciones de la sociedad civil enfrentan presiones para demostrar resultados tangibles, la capacidad de medir rigurosamente el cambio generado por sus acciones es crítica. Este documento ha argumentado que la evaluación de impacto no es solo un ejercicio técnico, sino un componente esencial de la buena gestión y la responsabilidad social. Su importancia radica en proveer evidencia para rendir cuentas, mejorar la asignación de recursos, escalar programas exitosos y corregir rumbos cuando sea necesario, todo ello contribuyendo en última instancia a maximizar el impacto positivo en las comunidades atendidas.

Asimismo, hemos visto que el campo de la evaluación de impacto está en constante renovación. Las nuevas tendencias —desde el uso de metodologías experimentales combinadas con enfoques participativos, hasta la incorporación de Big Data e inteligencia artificial, pasando por una mayor sensibilidad a la equidad y el aprendizaje en tiempo real— ofrecen oportunidades para llevar las evaluaciones a un siguiente nivel de calidad y relevancia. Lejos de ser modas pasajeras, estas tendencias responden a desafíos reales (p.ej., cómo evaluar en contextos complejos o cómo aprovechar flujos masivos de información) y, bien aplicadas, pueden potenciar la capacidad de generar conocimiento útil para la toma de decisiones sociales. Para las organizaciones comprometidas con programas sociales, estar al tanto de estas innovaciones evaluativas significa mantenerse a la vanguardia en prácticas de impacto.

Por último, el levantamiento de buenas prácticas a nivel internacional nos deja valiosas lecciones: planificar la evaluación desde el inicio, garantizar independencia y rigor, involucrar a las partes interesadas, cuidar la calidad de los datos, comunicar efectivamente los hallazgos y promover el uso de la evidencia. Estas directrices conforman un marco ético-técnico que aumenta la probabilidad de que las evaluaciones



cumplan su propósito de informar y mejorar. En un sentido más amplio, fomentar una cultura de evaluación dentro de las instituciones —ya sean ministerios, empresas o fundaciones— es clave para transitar hacia modelos de gestión basados en el conocimiento y la mejora continua.

En conclusión, la evaluación de impacto de programas sociales no debe verse como un trámite o una imposición de donantes, sino como una herramienta estratégica al servicio de quienes diseñan e implementan intervenciones para el cambio social. Cuando se realiza con calidad y con apertura al aprendizaje, la evaluación se vuelve un aliado para *amplificar* aquello que funciona y *rectificar* aquello que no, en la búsqueda permanente de soluciones más efectivas, equitativas y sostenibles a los problemas sociales. Empresas, ONG, fundaciones y sector público comparten la responsabilidad de adoptar estas prácticas evaluativas de vanguardia, colaborando incluso entre sí (mediante alianzas de aprendizaje, intercambio de datos y experiencias) para elevar el impacto colectivo de sus esfuerzos. Tal como el lema de la comunidad global de evaluadores indica, “*Better evidence, better policies, better lives*”, contar con mejor evidencia conducirá a mejores políticas y en definitiva a mejores vidas para las poblaciones a las que servimos.



## Bibliografía

- Bamberger, M.** (2016). *Integrating Big Data into the Monitoring and Evaluation of Development Programmes*. United Nations Global Pulse, New York.
- Banerjee, A. V., & Duflo, E.** (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.
- Banerjee, A. V., Duflo, E., & Kremer, M.** (2019). Research on improving the effectiveness of social programs (Premio Nobel de Economía 2019). *American Economic Review*, **109**(5), 4047–4070.
- Cousins, J. B., & Whitmore, E.** (1998). Framing participatory evaluation. *New Directions for Evaluation*, **1998**(80), 5-23.
- Esteves, A. M., Franks, D., & Vanclay, F.** (2012). Social impact assessment: the state of the art. *Impact Assessment and Project Appraisal*, **30**(1), 34-42.
- European Venture Philanthropy Association (EVPA).** (2015). *Measuring and Managing Impact: A Practical Guide*. Brussels: EVPA.
- Fiszbein, A., & Schady, N.** (2009). *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M.** (2016). *Impact Evaluation in Practice* (2nd ed.). Washington, DC: World Bank.
- Global Impact Investing Network (GIIN).** (2020). *Annual Impact Investor Survey 2020*. New York: GIIN.
- Kvam, R.** (2018). *Evaluación del impacto social: Integrando los aspectos sociales en los proyectos de desarrollo*. Washington, DC: Banco Interamericano de Desarrollo.
- Mayne, J.** (2012). Contribution analysis: Coming of age. *Evaluation*, **18**(3), 270-280.
- Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T.** (2012). *A Guide to Social Return on Investment*. Social Value UK (SROI Network).
- Patton, M. Q.** (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.
- Pawson, R., & Tilley, N.** (1997). *Realistic Evaluation*. London: SAGE Publications.





**Reynolds, M., Forss, K., Hummelbrunner, R., Marra, M., & Perrin, B.** (2012). Complexity, systems thinking and evaluation – an emerging relationship? *ILAC Brief*, **26**, 1-8.

**Rogers, P.** (2014). *Week 7: Innovation in Evaluation* (Blog series). BetterEvaluation.

**Wadhwani Foundation.** (2022). *Buenas prácticas para medir el impacto de los programas de cualificación de la mano de obra*. Wadhwani Foundation (Blog).

**Weiss, C. H.** (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

**York, P., Bamberger, M., & Bolanos, C.** (2020). Evaluation and the big data challenge. *American Journal of Evaluation*, **41**(2), 155-170.