

INTRODUCCIÓN A LOS MÉTODOS MIXTOS DE LA EVALUACIÓN DE IMPACTO

Por Michael Bamberger



Esta es la tercera guía de orientación de una serie de cuatro partes de guías relacionadas con la evaluación de impacto desarrolladas por InterAction y con el apoyo financiero de la Fundación Rockefeller. Las otras guías de esta serie son: *Introducción a la Evaluación de Impacto*; *Relación entre el Monitoreo y la Evaluación con la Evaluación de Impacto* y *Uso de los Resultados de la Evaluación de Impacto*. Puede encontrar la serie completa en el sitio web de InterAction : <http://www.interaction.org/impact-evaluation-notes>.

CONTENIDOS

Introducción	1
Parte I: ¿Por qué métodos mixtos?	3
1.1. ¿Qué es un diseño de evaluación de impacto con métodos mixtos?	3
1.2. Las limitaciones de una dependencia exclusiva en enfoques de evaluación CUANT o CUAL	4
1.3. Los beneficios de un enfoque de métodos mixtos	4
Parte II: El enfoque de métodos mixtos	9
2.1. Cuatro decisiones para diseñar una evaluación de métodos mixtos	9
2.2. Aplicación de enfoques de MM en cada etapa de la evaluación	13
Parte III: Aplicaciones de los diseños de métodos mixtos	19
3.1. Estrategias de muestreo para evaluaciones con MM orientadas CUANT y CUAL	19
3.2. Uso de métodos mixtos para evaluar intervenciones complejas	21
3.3. Evaluación de procesos de cambio de comportamiento	25
Parte IV: Administración de evaluaciones con métodos mixtos	27
4.1. Los diseños de métodos mixtos requieren un enfoque de administración especial	27
4.2. Consejos para que las ONG con recursos limitados movilicen la experiencia profesional y los recursos requeridos para realizar evaluaciones con métodos mixtos	30
Parte V: Estudios de caso que ilustran las diferentes aplicaciones de los diseños de métodos mixtos	32
Referencias	37
TABLAS	
Tabla 1. Les Los métodos mixtos se utilizan de manera diferente para diseños de evaluación en los que predomina una orientación CUANT o CUAL	11
Tabla 2. Métodos CUANT y CUAL de recolección de datos más usados	16
Tabla 3. Tipos diferentes de triangulación utilizados en evaluaciones con métodos mixtos	17
Tabla 4. Ejemplos del análisis de datos con métodos mixtos	18
Tabla 5. Áreas adicionales de experiencia profesional en investigación que se pueden requerir en equipos con una orientación CUANT o CUAL para realizar evaluaciones con MM.	29
Tabla 6. Ejemplos de evaluaciones con MM resumidas en el Anexo 10	36
CUADROS	
Cuadro 1. Cómo se utilizan los términos “impactos” y “evaluación de impacto” en las guías de orientación	3
Cuadro 2. Beneficios operativos por el uso de métodos mixtos	5
Cuadro 3. Análisis de atribución, contribución y sustitución	22
Cuadro 4. Los resultados del programa se ven afectados por la manera en que los beneficiarios responden a los servicios y cómo reacciona el personal ante las respuestas de los beneficiarios. Estudio del cambio de comportamiento en un programa de alimentación escolar en Nicaragua.	25
Cuadro 5. Planificación y presupuesto de tiempo y recursos adicionales que pueden requerirse para una evaluación con métodos mixtos	28
Cuadro 6. Realización de evaluaciones de impacto con métodos mixtos con pocos recursos: Consejos para las ONG con recursos limitados y poca experiencia profesional en evaluaciones	31

FIGURES

Figura 1. Uso de un marco de muestra nacional de hogares para garantizar la representatividad y credibilidad de una evaluación de estudios de caso CUAL	6
Figura 2. Uso de estudios etnográficos CUAL de pueblos y hogares para ayudar a interpretar los hallazgos de una encuesta nacional de muestra CUANT	8
Figura 3. Diseño de método mixto secuencial con un enfoque predominantemente cuantitativo: Estudio de transferencias entre hogares como una estrategia de supervivencia de hogares de bajos ingresos en Cartagena, Colombia	12
Figura 4. Diseño de método mixto secuencial con un enfoque predominantemente cualitativo: Evaluación de la adopción de nuevas variedades de semillas por parte de diferentes tipos de agricultores	12
Figura 5. Diseño de métodos mixtos en varios niveles: Evaluación de los efectos de un programa de alimentación escolar en la asistencia y el desempeño	14
Figura 6. Uso de diseños de métodos mixtos para evaluar intervenciones complejas	23

ANEXOS (disponibles en inglés en: <http://www.interaction.org/impact-evaluation-notes>)

Anexo 1. Ventajas y desventajas de diseños de evaluación cuantitativos
Anexo 2. Ventajas y desventajas de diseños de evaluación cualitativos
Anexo 3. Ejemplos de diseños de evaluación en cada punto del continuo CUANT - CUAL
Anexo 4. Características de enfoques CUANT y CUAL en diferentes etapas de la evaluación
Anexo 5. Cómo los enfoques CUANT y CUAL se complementan entre sí en diferentes etapas de una evaluación
Anexo 6. Comparación entre métodos de muestreo aleatorios y dirigidos
Anexo 7. Variedad de enfoques cuantitativos, cualitativos y basados en la teoría para definir lo contrafactual
Anexo 8. Estrategias para reducir los costos de la recolección y el análisis de datos
Anexo 9. Ejemplo de triangulación: comparación de estimaciones de ingresos de hogares y pobreza de diferentes fuentes
Anexo 10. Estudios de caso de diseños de evaluación con MM equilibrados y con orientación predominantemente CUANT o CUAL
Anexo 11. Cómo los métodos mixtos pueden afianzar los diseños de evaluación CUANT

Introducción

Las evaluaciones con métodos mixtos (MM) procuran integrar las disciplinas de las ciencias sociales con enfoques predominantemente cuantitativos (CUANT) y predominantemente cualitativos (CUAL) a la teoría, la recolección, el análisis y la interpretación de los datos. El objetivo es aumentar la confiabilidad de los datos y la validez de los hallazgos y las recomendaciones, además de ampliar y profundizar la comprensión de los procesos a través de los cuales se logran los resultados e impactos de un programa, y cómo estos se ven afectados por el contexto en el que se implementa tal programa. Aunque ahora los métodos mixtos se utilizan comúnmente en la evaluación de programas, y las solicitudes de propuestas de evaluación con frecuencia exigen el uso de dichos métodos, muchos evaluadores no aprovechan todo el potencial del enfoque de MM.

Esta guía de orientación explica los elementos esenciales del enfoque de MM y cómo se puede utilizar en una evaluación de impacto (EI), además de resaltar los posibles beneficios y aplicaciones para las ONG. La Parte I aborda la pregunta: “¿Por qué métodos mixtos?” Analizamos lo que es un diseño de evaluación de impacto con MM, qué lo diferencia de un diseño de evaluación de impacto CUANT o CUAL y por qué el enfoque es útil para comprender las evaluaciones de desarrollo y las complejidades del mundo real en el que se implementan (Sección 1.1). La creciente popularidad de los MM se debe al reconocimiento de las limitaciones de una dependencia exclusiva en uno de los

dos métodos, CUANT o CUAL, (Sección 1.2) y en los posibles beneficios que se pueden lograr cuando ambos enfoques se combinan adecuadamente (Sección 1.3). Aunque los MM se pueden emplear como parte de una evaluación de impacto extensa y bien financiada, los métodos tienen la flexibilidad de ser igualmente útiles para muchas ONG que requieren evaluaciones creíbles de sus programas, pero cuyos recursos y experiencia profesional para realizar evaluaciones de impacto son limitados.

Después de haber expuesto el caso de los MM, la Parte II describe cuatro decisiones clave que se deben tomar al diseñar una evaluación con

MM (Sección 2.1): en qué etapas de la evaluación se emplearán los MM; si el diseño de MM es secuencial o concurrente; si el diseño será predominantemente CUANT o CUAL, o se utilizará un diseño equilibrado que otorgue igual importancia a ambos enfoques; y si el diseño se implementará en un solo nivel o si será una evaluación en varios niveles. La Sección 2.2 explica cómo se pueden usar los MM en cada etapa del diseño, la implementación y el análisis de una evaluación, no solo como una manera de diversificar los métodos de recolección de datos, como suponen muchos evaluadores.

La Parte III revisa algunas aplicaciones de los diseños de MM. La Sección 3.1 analiza estrategias de muestreo con MM cuando se usan diseños predominantemente CUANT o CUAL y demuestra de qué manera los MM pueden favorecer ambas clases de diseño de muestras. La Sección 3.2 describe el uso de los MM para evaluar intervenciones de desarrollo complejas y la Sección 3.3 expone la manera en que los diseños de MM pueden ayudar a evaluar programas que incluyen procesos de cambio de comportamiento.

La Parte IV aborda temas concernientes a la administración de evaluaciones con MM. La Sección 4.1

explica por qué es necesario un enfoque de administración especial, y la Sección 4.2 analiza cómo las ONG pueden utilizar los MM con eficacia para las evaluaciones que se realizan con pocos recursos (es decir, con restricciones de tiempo y presupuesto, y con una experiencia profesional limitada en lo que concierne a la investigación).

La Sección V presenta tres estudios de caso que muestran cómo se utilizan los MM en evaluaciones predominantemente CUANT o CUAL, y en una evaluación equilibrada que otorga igual importancia a ambos enfoques, CUANT y CUAL. El Anexo 10 presenta 17 ejemplos de evaluaciones con MM que ilustran una gran variedad de enfoques e incluyen tanto evaluaciones extensas y bien financiadas como evaluaciones realizadas con limitaciones de presupuesto, tiempo y datos.

Un desafío en la preparación de esta guía de orientación (GO) fue la gran cantidad de material bibliográfico que existe sobre MM, parte del cual es muy técnico o especializado. Tratamos de mantener el texto accesible para un público amplio y no especializado, a la vez que incorporamos un conjunto de anexos (disponibles en inglés en <http://www.interaction.org/impact-evaluation-notes>) con información más detallada.

Parte I: ¿Por qué métodos mixtos?

1.1. ¿Qué es un diseño de evaluación de impacto con métodos mixtos?

Diferentes evaluadores utilizan los términos “impacto” y “evaluación de impacto” de diferentes maneras. Por lo tanto, a fin de asegurarnos de que todos estemos en la misma página, el Cuadro 1 resume cómo se utilizan estos términos en estas guías de orientación. Los métodos mixtos tienen los mismos objetivos, formulan muchas de las mismas preguntas y emplean todas las herramientas y técnicas de evaluación de impacto descritas en la guía de orientación 1 (GO1) de esta serie, además de basarse en los sistemas de evaluación y monitoreo de proyectos de la manera descrita en la GO2. Al igual que otros diseños de evaluación de impacto, los MM se pueden aplicar en cualquier nivel, desde la evaluación de un proyecto que funciona en un solo pueblo hasta una iniciativa de desarrollo nacional de varios componentes que involucra muchas agencias nacionales e internacionales diferentes.

Es poco frecuente encontrar una única metodología de evaluación que pueda contemplar absolutamente todas las complejidades del funcionamiento de los programas en el mundo real. En consecuencia, los evaluadores deben hallar maneras creativas de combinar diferentes marcos, herramientas y técnicas de evaluación¹; esto explica el creciente

¹ Un tema relacionado importante concierne a la elección del diseño de evaluación apropiado. Debido a las muchas clases de programas diferentes que se evalúan, los contextos variados en los que funcionan y la diversidad de preguntas de evaluación de relevancia para las partes interesadas, no existe un único diseño de evaluación que sea “el mejor” para todas las situaciones. La elección del diseño de evaluación requiere un análisis exhaustivo de la naturaleza del programa, el objetivo y el contexto de la evaluación, y el entorno en el que funciona. Consulte Bamberger, Rugh y Mabry (2012), Capítulos 2 y 11 para obtener un análisis de las estrategias para seleccionar el diseño de evaluación más apropiado. Consulte también Stern et al (2012) para encontrar un análisis detallado de la elección de los métodos.

interés por los enfoques de MM. La característica única de los enfoques de métodos mixtos (MM) es que procuran integrar las disciplinas de las ciencias sociales con enfoques predominantemente CUANT y predominantemente CUAL a la teoría y a la recolección, el análisis y la interpretación de los datos. Aunque, actualmente, muchos evaluadores utilizan una variedad de métodos como rutina, “Lo que distingue una evaluación con método mixto es el uso intencional o planificado de diversos métodos por los fines particulares de los métodos mixtos utilizando diseños particulares de métodos mixtos” (Greene 2005:255). Es muy habitual combinar métodos de recolección de datos para realizar una evaluación con MM, pero también es posible combinar marcos conceptuales, desarrollos de hipótesis, análisis de datos o marcos para la interpretación de los hallazgos de la evaluación.

CUADRO 1. CÓMO SE UTILIZAN LOS TÉRMINOS “IMPACTOS” Y “EVALUACIÓN DE IMPACTO” EN LAS GUÍAS DE ORIENTACIÓN

La guía de orientación 1, “Introducción a la evaluación de impacto” (página 2), define *impactos* como:

los efectos positivos y negativos, deseados y no deseados, directos e indirectos, primarios y secundarios, que produce una intervención. (Definición del Comité de Asistencia para el Desarrollo de la OECD).

Normalmente, se entiende que los impactos ocurren después, y como consecuencia, de resultados intermedios. La diferencia entre resultados e impactos puede ser relativa, y depende de los objetivos establecidos de una intervención. De acuerdo con nuestra definición, una evaluación de impacto incluye:

cualquier evaluación que investigue de manera sistemática y empírica los impactos que produce (y que se pueden atribuir a) una intervención.

1.2. Las limitaciones de una dependencia exclusiva en enfoques de evaluación CUANT o CUAL

Cuando se utilizan de forma aislada, ambos métodos de evaluación, CUANT y CUAL, tienen ventajas y desventajas. El objetivo de los MM es aprovechar las ventajas de ambos enfoques, CUANT y CUAL, e integrarlos para superar sus desventajas. A pesar de los muchos beneficios convincentes de la recolección de datos y métodos de análisis CUANT, también tienen una serie de limitaciones inherentes ([Anexo 1](#)). Muchas de las críticas se refieren a la reducción de datos narrativos en números, y diseños y procedimientos de métodos de recolección de datos inflexibles que son difíciles de adaptar a circunstancias cambiantes. Las categorías estandarizadas en cuestionarios y codificación de datos a menudo no pueden captar los matices de las comunidades o los grupos estudiados, y suele suceder que el análisis carece de la profundidad y del detalle de los métodos CUAL. El riesgo de la evaluación CUANT es la *descontextualización*, es decir, que ignore cómo afectan a los programas las características económicas, políticas, institucionales y socioculturales de las poblaciones estudiadas. Otra crítica frecuente a muchas evaluaciones CUANT es que suponen que los programas funcionan tal como se los planificó y que todos reciben el mismo paquete de servicios (tanto en términos de cantidad como de calidad). Esto es a menudo conocido como enfoque de “caja negra”, ya que la evaluación no mira dentro de la “caja negra” del proyecto.

Los métodos CUAL también son herramientas eficaces para la recolección y el análisis de datos. No obstante, cuando se los utiliza exclusivamente, los diseños de evaluación CUAL también tienen una serie de desventajas posibles ([Anexo 2](#)). Las evaluaciones CUAL a menudo se centran en temas y situaciones particulares, y es más difícil

generalizar a partir de los hallazgos. Muchos evaluadores CUAL también creen que cada evaluación es específica de cada contexto y no es posible o apropiado generalizar. Muchas, pero ciertamente no todas, evaluaciones CUAL utilizan un enfoque holístico; por lo que resulta más difícil aislar elementos y factores particulares, y comprender la contribución específica de diferentes componentes o enfoques del programa. Algunos clientes también se sienten incómodos cuando se depende demasiado de la opinión y la perspectiva del evaluador, sin dejar margen para que el lector revise fácilmente las grandes cantidades de datos escritos y registrados en las que se basó el evaluador. Un último punto es que las evaluaciones CUAL no suministran la clase de documentación detallada respecto de la metodología que, por lo general, presentan los informes de evaluación CUANT y, por lo tanto, es difícil verificar la validez de la recolección de datos y los procedimientos de análisis.²

1.3. Los beneficios de un enfoque de métodos mixtos³

Existen cinco razones principales para utilizar diseños de métodos mixtos (Greene 2005:255–56):

- *Triangulación de los hallazgos de la evaluación*: aumentar la validez o credibilidad de los hallazgos de la evaluación al comparar la información obtenida a partir de diferentes métodos de recolección de datos (por ejemplo, comparar respuestas a preguntas

² Por ejemplo, muchas evaluaciones CUAL no incluyen documentación detallada sobre cómo se seleccionaron los integrantes de un grupo de discusión, y solo algunas pueden proporcionar transcripciones de las entrevistas (para considerar el costo y el tiempo); de modo que, por lo general, no es posible para el lector revisar los datos de manera independiente y evaluar cómo se obtuvieron los hallazgos y las conclusiones.

³ Para obtener información reciente sobre los beneficios de los enfoques de métodos mixtos, consulte Adato (2012).

de encuestas con las observaciones directas del entrevistador). Cuando las estimaciones de diferentes fuentes convergen y coinciden, esto aumenta la validez y la credibilidad de los hallazgos o la interpretación. Cuando las diferentes estimaciones se contradicen, el investigador sigue explorando para comprender la razón de las contradicciones (consulte el [Anexo 9](#)).

- **Desarrollo:** utilizar resultados de un método para ayudar a desarrollar la muestra o la instrumentación de otro.
- **Complementariedad:** extender la exhaustividad de los hallazgos de la evaluación a través de resultados de diferentes métodos que amplían y profundizan la comprensión obtenida.
- **Iniciación:** generar nuevas apreciaciones en los hallazgos de la evaluación a través de los resultados de los diferentes métodos

que divergen y, de esta manera, requieren de una conciliación a través de análisis adicionales, reformulaciones o un cambio en la perspectiva.

- **Diversidad de valor:** incorporar una variedad de valores más amplia a través del uso de diferentes métodos que en sí mismos proponen diferentes valores. Esto fomenta una mayor toma de consciencia respecto de las dimensiones de los valores de la evaluación.

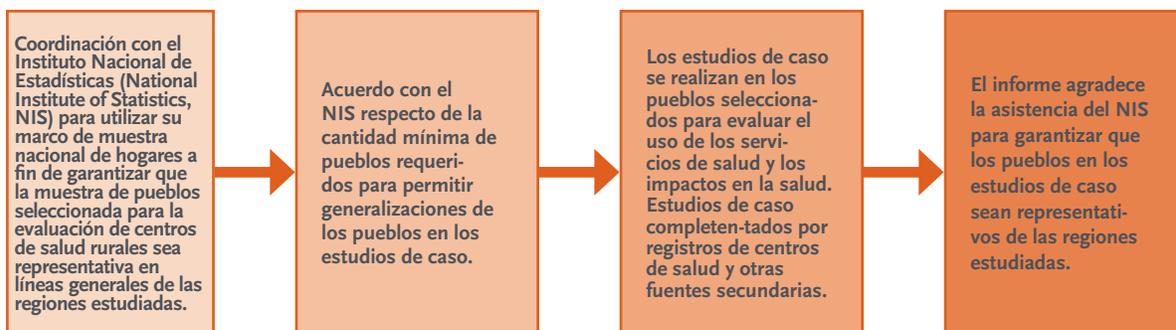
Un beneficio adicional es que un enfoque de MM tiene más probabilidades de aceptación entre evaluadores y usuarios que tienen predilección tanto por enfoques CUANT como por enfoques CUAL.

El Cuadro 2 resume algunos de los beneficios operativos de usar un enfoque de MM.

CUADRO 2. BENEFICIOS OPERATIVOS POR EL USO DE MÉTODOS MIXTOS

- Comprender cómo los factores contextuales locales ayudan a explicar las variaciones en la implementación y los resultados de los programas.
 - Reconstruir datos de base para evaluaciones CUANT cuando no es posible realizar una encuesta de base. Muchas evaluaciones se encargan hacia el final del programa y no tienen información muy confiable sobre las condiciones del proyecto y los grupos de comparación en el momento en que se inició el programa. Esto hace difícil determinar si las diferencias observadas al final del proyecto se pueden atribuir a los efectos del programa o si estas diferencias podrían deberse, al menos en parte, a diferencias preexistentes entre los dos grupos. Por ejemplo, las mujeres que solicitan préstamos para pequeñas empresas pueden provenir de familias que brindan más apoyo a las mujeres propietarias de pequeñas empresas que la mayoría de las familias, o pueden ya tener más experiencia comercial que las mujeres que no solicitan préstamos. Si no se identifican estas diferencias preexistentes, existe el riesgo de sobrestimar los efectos del programa de préstamos. Por lo general, es posible utilizar
- tales técnicas CUAL como entrevistas exhaustivas, entrevistas a informantes clave o grupos de discusión para obtener información acerca de las características de los beneficiarios y no beneficiarios del programa en el momento de inicio del programa. Esta clase de información, que a menudo es muy simple y económica de obtener, puede aumentar en gran medida la validez de los datos de una encuesta exclusivamente CUANT.
 - Fortalecer la representatividad de los estudios CUAL exhaustivos (por ejemplo, al vincular la selección de estudios de caso con el marco de muestreo CUANT) puede facilitar la comparación de los hallazgos con los datos de la encuesta CUANT.
 - Aportar sentido común respecto de la validez y el valor de diferentes clases de datos CUANT y CUAL.
 - Promover una mayor comprensión de las perspectivas de las partes interesadas acerca de la naturaleza de la intervención o cómo se espera lograr sus objetivos. Esto propicia un enfoque más participativo y una mayor coordinación entre partes interesadas y evaluadores.

Figura 1. Uso de un marco de muestra nacional de hogares para garantizar la representatividad y credibilidad de una evaluación de estudios de caso CUAL



Para ilustrar algunos de estos beneficios, examinemos el ejemplo de una evaluación de impacto de centros de salud rurales sobre la salud de la población rural, en especial de mujeres y niños. Una preocupación en particular es la accesibilidad de los servicios de salud a sectores pobres y vulnerables de la población. En este primer ejemplo, la evaluación adopta un **enfoque predominantemente CUAL**. La finalidad de la evaluación es influir en las políticas de salud nacionales al identificar algunas de las razones por las cuales las familias pobres y minoritarias no utilizan los centros de salud. El equipo de evaluación sabe que el Ministerio de Salud ha criticado las evaluaciones CUAL previamente por haberse concentrado intencionalmente en comunidades con problemas particulares consabidos, lo que implica que los hallazgos no son representativos de todo el país y desprestigiarán al Ministerio de Salud. Los evaluadores también saben que el Ministerio algunas veces utilizó la falta de representatividad estadística como una excusa conveniente para ignorar críticas válidas, y desean cerciorarse de que su estudio no sea descartado por ese motivo. Por consiguiente, el equipo de evaluación coordina con el Instituto Nacional de Estadísticas y utiliza su marco de muestra nacional de hogares para garantizar que la muestra de las comunidades que seleccionan sea representativa en líneas generales de todo el país (o la región

donde se realiza el estudio). La evaluación utiliza los mismos métodos CUAL, pero ahora es posible indicar que la muestra de comunidades es representativa en líneas generales de todas las comunidades en las regiones estudiadas (vea la Figura 1).

Ahora bien, supongamos que la misma evaluación será realizada por un equipo diferente que planea usar un **enfoque CUANT** basado en una encuesta de muestra de hogares representativos a nivel nacional. Aunque una encuesta bien diseñada puede obtener estimaciones razonablemente confiables de la proporción de la población que utiliza los centros de salud (no obstante, incluso en ese caso existe la posibilidad de un informe erróneo), los evaluadores son totalmente conscientes de que la encuesta no explicará por qué los hogares utilizan o no los centros de salud. Por lo tanto, invitan a un etnógrafo a incorporarse al equipo y realizar estudios exhaustivos en una cantidad reducida de comunidades. Los estudios etnográficos explorarán las actitudes y creencias relacionadas con la salud de diferentes grupos etnográficos y los factores que influyen en su decisión de utilizar o no los centros de salud. Los estudios también examinarán los factores económicos, políticos, organizacionales, culturales y ecológicos que afectan el funcionamiento de los centros de salud en diferentes comunidades. La primera parte del análisis abordará las diferencias culturales

generales que pueden afectar a todos los centros de salud, y la última parte (el análisis contextual) ayudará a explicar los factores que afectan el funcionamiento de diferentes centros (Figura 2). El director de la evaluación sabe que los diseños de métodos mixtos funcionan bien solo cuando hay respeto y comprensión, y una sensación de igualdad entre los miembros del equipo de diferentes profesiones, de modo que se invitó al etnógrafo a incorporarse al equipo desde el mismo momento en que se realizó la primera reunión de planificación. Las siguientes son algunas de las formas en las que los enfoques CUANT y CUAL se pueden integrar a esta evaluación:

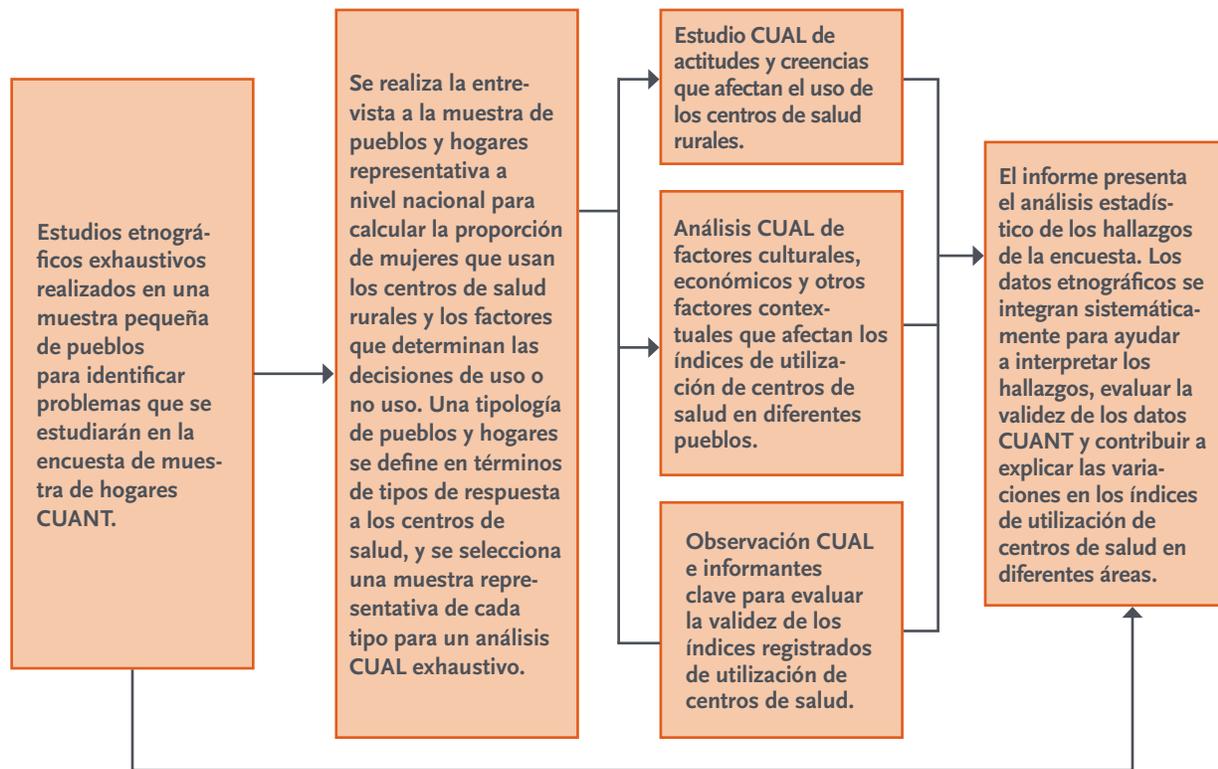
- Los estudios etnográficos rápidos (CUAL) se realizan en comunidades seleccionadas para comprender los temas que se deben tratar en la encuesta y para ayudar a formular las preguntas.
- A continuación, se realiza una encuesta CUANT de hogares utilizando una muestra representativa a nivel nacional.
- El análisis de la encuesta de hogares puede producir una tipología de hogares de acuerdo a su nivel de uso de los centros de salud o sus razones para no utilizarlos. Se seleccionará una muestra de cada tipo para preparar estudios de caso. Los estudios de caso a menudo revelarán que el uso que se alegó o los motivos por no usar los centros de salud no son correctos. Por ejemplo, las mujeres generalmente no mencionan el acoso sexual para responder a la encuesta y, en su lugar, pueden alegar que las horas de atención al público no son convenientes.
- Se utilizará la triangulación con el fin de obtener estimaciones CUANT y CUAL independientes para variables clave (como el uso de instalaciones sanitarias y actitudes hacia estas instalaciones). Una característica clave

de la triangulación es que los procedimientos se integran para identificar contradicciones en diferentes estimaciones y para realizar un seguimiento a fin de comprender la razón de las diferencias. Por ejemplo, se observará cómo se recibe a las personas que ingresan a la clínica o se conversará informalmente con las familias acerca de si utilizan las clínicas y cuándo lo hacen para comparar estos datos con las respuestas a la encuesta.

- Se prepararán borradores de informes CUANT y CUAL independientes, y luego los equipos se reunirán para identificar las áreas en las que hay aparentes diferencias en los datos o la interpretación. En el ejemplo de las contradicciones entre las respuestas a la encuesta sobre la utilización de los centros de salud y los datos de las observaciones y entrevistas exhaustivas, los investigadores CUANT y CUAL se reunirán para analizar las razones de tales contradicciones. Pueden llegar a un acuerdo respecto de que una u otra fuente de información es más confiable. Por ejemplo, las entrevistas exhaustivas con mujeres cuando ningún otro miembro del hogar está presente pueden considerarse más confiables. No obstante, si no está claro qué fuente es más confiable, los investigadores podrían volver al campo a recolectar más datos o se podrían buscar otras fuentes de información, como por ejemplo, una revisión de los registros de los centros de salud sobre visitas de pacientes o consultas a informantes clave, como líderes de la comunidad, enfermeros, etc.

En los dos casos presentados, los MM pueden favorecer la evaluación. Sin embargo, el foco de atención es muy distinto cuando los MM se utilizan para afianzar diseños predominantemente CUAL (Figura 1) que cuando se usan para afianzar diseños predominantemente CUANT (Figura 2).

Figura 2. Uso de estudios etnográficos CUAL de pueblos y hogares para ayudar a interpretar los hallazgos de una encuesta nacional de muestra CUANT



En términos generales, un enfoque de MM es particularmente útil para:

- Examinar las interacciones entre los factores contextuales complejos y cambiantes que pueden influir en la implementación y los impactos del programa.
- Definir y medir indicadores de los factores culturales, históricos, políticos, legales, ambientales y psicosociales que afectan la implementación. Se requieren diferentes metodologías para medirlos.
- Captar procesos complejos de cambios organizacionales y de comportamiento (secciones 3.2 y 3.3).
- Tener en cuenta cómo cambian los programas en respuesta a cómo son percibidos y utilizados por diferentes sectores de la población

objetivo. La experiencia de los primeros usuarios y sus comentarios a vecinos pueden afectar radicalmente la evolución de un programa.⁴

- Muchos procesos y resultados son difíciles de observar o, en algunos casos, incluso es difícil saber si existen. Esto es particularmente importante para evaluar la situación de grupos vulnerables y para programas que afectan actividades ilegales o socialmente desaprobadas, como el uso de drogas, trabajo sexual o la inmigración legal. Todas estas dificultades se multiplican en los programas humanitarios, posteriores a conflictos y otras clases de programas de ayuda ante emergencias.

⁴ *Realist evaluation* (Evaluación realista) (Pawson 2006) proporciona un marco útil para el análisis del cambio de comportamiento y para el análisis de cómo los programas funcionan realmente en el campo.

Parte II: El enfoque de métodos mixtos

2.1. Cuatro decisiones para diseñar una evaluación de métodos mixtos

Cuando se planifica una evaluación con MM, se requieren cuatro decisiones:

1. ¿En qué etapa o etapas de la evaluación se utilizarán los MM?
2. ¿Los métodos CUANT y CUAL se emplearán de manera secuencial o concurrente?
3. ¿Se otorgará igual importancia a los métodos CUANT y CUAL, o predominará una metodología?
4. ¿El diseño será de un solo nivel o de varios niveles?

Decisión 1: ¿En qué etapas de la evaluación se utilizarán los métodos mixtos?

La mayoría de las evaluaciones con MM solo combinan métodos CUANT y CUAL en una o quizás dos etapas de la evaluación, por lo general, en la recolección de datos. No obstante, un diseño de MM es mucho más eficaz si se integran los enfoques CUANT y CUAL en varias etapas de la evaluación (o idealmente en todas). La Sección 2.2 explica cómo se pueden emplear los MM para afianzar cada etapa de una evaluación, el [Anexo 4](#) compara enfoques CUANT y CUAL en cada etapa de una evaluación y el [Anexo 5](#) presenta ejemplos de cómo se complementan los enfoques CUANT y CUAL para que cada etapa de la evaluación sea más efectiva.

Decisión 2: ¿El diseño de MM es secuencial o concurrente?

Diseños de métodos mixtos secuenciales

En los diseños secuenciales, los métodos CUANT

y CUAL se utilizan en fases. Por ejemplo, la evaluación puede comenzar con un estudio exploratorio CUAL para tratar de comprender los problemas clave y cómo los perciben las poblaciones afectadas. Esto sirve para diseñar una encuesta CUANT, que luego se administra en una muestra seleccionada al azar. A continuación, los datos podrían analizarse utilizando métodos de análisis CUANT o CUAL, o ambos. En otro ejemplo, podría utilizarse una encuesta CUANT rápida para identificar y cuantificar los principales tipos de granjas y actividades agrícolas. Esta información luego se usaría para seleccionar una muestra representativa de granjas para la preparación de estudios de caso CUAL exhaustivos. Los estudios de caso probablemente se analizarían utilizando métodos CUAL y la encuesta de muestra se analizaría utilizando técnicas CUANT. La Figura 3 es un ejemplo de un diseño secuencial para evaluar transferencias entre hogares como una estrategia de supervivencia de familias pobres. Esta evaluación comenzó con un estudio etnográfico (CUAL) para comprender las características de las comunidades, seguido de una encuesta CUANT a hogares y un análisis econométrico de los hallazgos.

Diseños concurrentes

En los diseños concurrentes, los enfoques CUANT y CUAL se utilizan al mismo tiempo. Un ejemplo de un diseño concurrente es uno en el que los datos CUANT y CUAL se recolectan de forma simultánea, utilizando la triangulación para comparar la información de los resultados, los impactos y otros indicadores clave de distintas fuentes independientes. Otro ejemplo es cuando se utilizan métodos CUAL para realizar un análisis contextual del lugar de un proyecto (o las áreas circundantes) al mismo tiempo que se lleva a cabo una encuesta

de muestra CUANT a hogares o personas. Esto brinda la posibilidad de realizar un análisis muy profundo pero más complicado en el que se analizan las interacciones entre el entorno (contexto) y el proceso de implementación del proyecto.

Consideraciones operativas para decidir entre un diseño secuencial y un diseño concurrente

Una ventaja de los diseños secuenciales es que la logística es generalmente más fácil de organizar. La recolección de datos que hace uso de cuestionarios estructurados a menudo requiere un equipo grande de entrevistadores en el campo que sigan un cronograma de selección de hogares definido con precisión y una cantidad de entrevistas que se deberán realizar cada día. Los supervisores de campo deben saber dónde está trabajando cada enumerador porque el control de calidad por lo general incluye visitas de seguimiento a una submuestra de hogares. El supervisor también debe estar disponible para responder las preguntas de los enumeradores. En contrapartida, los métodos etnográficos y muchas otras clases de métodos CUAL tienen un cronograma mucho más flexible en términos de duración y en el que los investigadores estarán disponibles en cualquier momento dado. Por este y otros motivos, con frecuencia los diseños de MM concurrentes pueden ser difíciles de administrar, en particular para equipos de evaluación con unos pocos supervisores experimentados. Los diseños concurrentes pueden ser un problema en particular en áreas donde la planificación logística (por ejemplo, traslados a los sitios, lugares donde alojarse, la seguridad) puede tornarse difícil de coordinar; también son

más complicados en cuanto al manejo de opiniones y comentarios, ya que se tendrían que realizar ajustes con mayor rapidez que en el caso de los diseños secuenciales. Por otro lado, los diseños concurrentes tienen la ventaja de que la recolección y el análisis de los datos se pueden realizar más rápidamente.

Decisión 3: ¿El diseño de MM será predominantemente CUANT o CUAL, o se utilizará un diseño equilibrado?

Resulta útil pensar en una representación de los diseños de evaluación como un continuo de enfoques que va desde enfoques exclusivamente CUANT pasando por enfoques que otorgan igual importancia a los métodos CUANT y CUAL hasta enfoques exclusivamente CUAL (Bamberger et al 2012 pp. 324-34; Greene y Caracelli 2003). Esto es importante ya que diferentes evaluadores, quienes quizás comenzaron sus carreras con una orientación predominantemente CUANT o CUAL, pueden tener expectativas muy diferentes respecto de lo que implica una evaluación con MM. También reviste importancia porque, debido a la orientación profesional de los evaluadores, en la mayoría de las evaluaciones con MM predomina un enfoque CUANT o CUAL (en lugar de uno equilibrado).⁵

La Tabla 1 ilustra cómo se utilizan los MM en evaluaciones en las que el enfoque predominante es CUANT o CUAL, y el [Anexo 3](#) ofrece ejemplos de diseños de evaluación en diferentes puntos de este continuo.

⁵ Está emergiendo una nueva generación de evaluadores que se capacitaron en MM como un enfoque de evaluación integrado, y comienzan a aparecer algunos estudios con un enfoque más equilibrado sin una orientación predominante, aunque aún son la minoría. El *Journal of Mixed Method Research* es una buena fuente de ejemplos de diseños equilibrados.

Tabla 1. Los métodos mixtos se utilizan de manera diferente para diseños de evaluación en los que predomina una orientación CUANT o CUAL

¿Qué enfoque predomina?	¿Cómo funciona el enfoque predominante?	¿Cómo se utiliza la otra orientación para afianzar el diseño?
CUANT	Por lo general, la evaluación administra un cuestionario estructurado a una muestra seleccionada al azar de individuos, grupos, instituciones o comunidades, y el análisis depende principalmente de métodos econométricos u otros métodos cuantitativos.	Se utilizan entrevistas exhaustivas, observaciones y entrevistas grupales para diseñar el cuestionario. Muestras reducidas de casos seleccionados a partir de la muestra principal también pueden brindar una comprensión más profunda de las relaciones estadísticas halladas en el análisis CUANT. Los casos pueden ser representativos de cada categoría principal identificada en el análisis o utilizados para estudiar observaciones atípicas u otros grupos seleccionados a propósito.*
Se otorga igual importancia a enfoques CUANT y CUAL	Las encuestas CUANT se combinan con una variedad de técnicas CUAL diferentes. A veces, estas últimas se centran en el proceso y el análisis contextual; en otros casos, el foco de atención está en la misma unidad de análisis que las encuestas (por ejemplo, individuos, hogares, comunidades, organizaciones), pero se utilizan diferentes métodos de recolección de datos.	
CUAL	Estudios de caso, entrevistas exhaustivas y otras técnicas CUAL se aplican a muestras relativamente pequeñas de individuos, hogares, comunidades o grupos.	Una encuesta CUANT rápida se usa para identificar los problemas o grupos que se abordarán en estudios CUAL exhaustivos o bien, para demostrar que una muestra CUAL es razonablemente representativa de la población total.
* Consulte el Anexo 6 y Bamberger et al 2012 pp. 360-61 para obtener una descripción de muestreo dirigido.		

La Figura 3 describe un diseño secuencial con un enfoque predominantemente CUANT. Este es un estudio de transferencias de dinero y bienes entre hogares como estrategia de supervivencia de hogares urbanos pobres en Colombia (Wansbrough, Jones y Kappaz 2000). El objetivo del estudio fue describir los patrones de transferencias y estimar si eran lo suficientemente importantes para actuar como una red de seguridad social informal que brinda ayuda a los sectores más pobres de la comunidad en épocas de necesidad. Estas transferencias entre hogares son difíciles de identificar y d medir. Por tal motivo, un antropólogo vivió

en la comunidad durante un mes para estudiar los patrones de transferencias y colaborar con el diseño del cuestionario para una encuesta CUANT que luego se administró en cientos de hogares. Los datos se analizaron utilizando análisis econométricos CUANT.

La Figura 4 ilustra un diseño secuencial con un enfoque predominantemente CUAL. Describe una evaluación hipotética para analizar la adopción de nuevas variedades de semillas por parte de diferentes tipos de familias rurales. Los métodos de recolección de datos principales son cualitativos:

Figura 3. Diseño de método mixto secuencial con un enfoque predominantemente cuantitativo: Estudio de transferencias entre hogares como una estrategia de supervivencia de hogares de bajos ingresos en Cartagena, Colombia

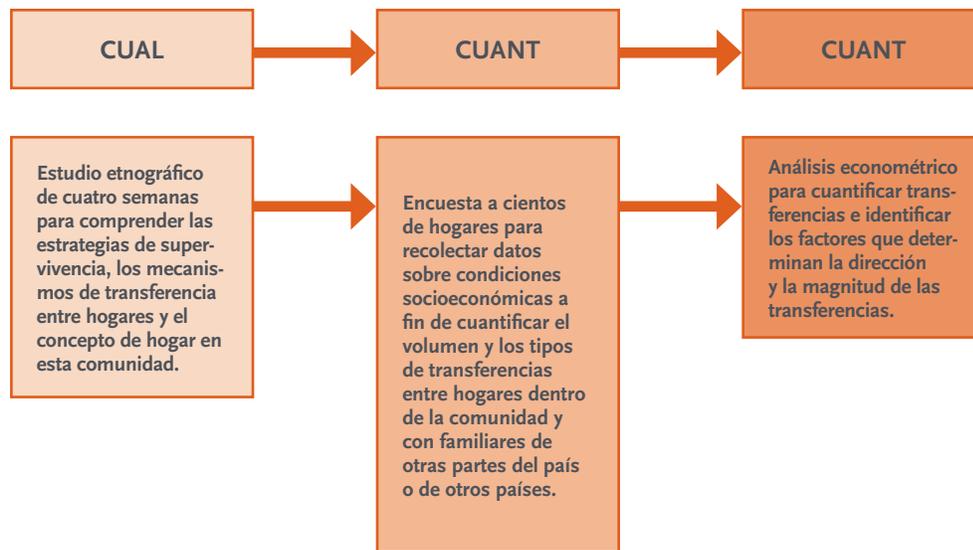
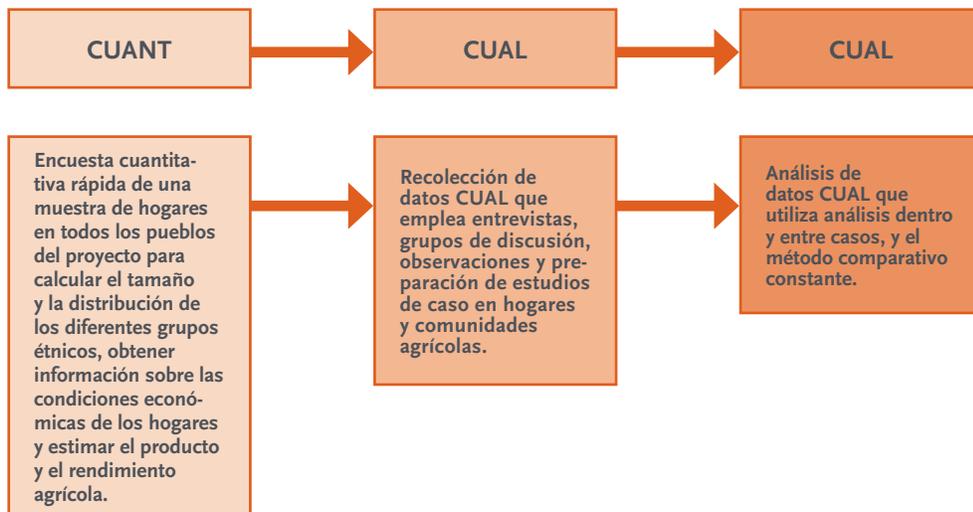


Figura 4. Diseño de método mixto secuencial con un enfoque predominantemente cualitativo: Evaluación de la adopción de nuevas variedades de semillas por parte de diferentes tipos de agricultores



entrevistas, grupos de discusión, observaciones y estudios de caso de hogares individuales y pequeñas comunidades agrícolas. Los métodos de análisis principales también son cualitativos: análisis dentro y entre casos, y el método comparativo constante. Sin embargo, para obtener información

sobre la distribución étnica de los hogares, las condiciones económicas de los hogares y la producción agrícola, la evaluación comienza con una encuesta CUANT rápida que cubre una muestra de hogares en todos los pueblos incluidos en el proyecto de extensión agrícola. Los hallazgos de

este estudio se utilizaron para identificar los tipos de hogares que se deben estudiar en más profundidad a través de métodos de recolección de datos CUAL, y para garantizar que los casos seleccionados sean representativos en líneas generales del total de la población de la encuesta.

Cualquiera de los dos diseños de evaluación anteriores se podrían haber modificado para otorgar igual importancia al enfoque CUANT y al enfoque CUAL. En el caso del estudio de transferencias entre hogares, la encuesta a hogares se pudo haber complementado con estudios de caso CUAL sobre familias o redes de transferencias informales. Estos se pudieron haber integrado en el análisis para comparar la descripción y la interpretación de las funciones y la operación de las redes de transferencias obtenidas a partir de los estudios CUAL con los hallazgos del análisis econométrico. En el segundo ejemplo, se pudo haber realizado un estudio CUAL o CUANT de puntos de marketing para calcular los cambios en las ventas del producto agrícola de las áreas del proyecto y, posiblemente, los cambios en la compra de bienes de consumo por parte de las familias de las áreas del proyecto.

El ejemplo 3 de la Parte V describe un diseño equilibrado (integrado) de evaluación con MM utilizado para evaluar un extenso programa de desarrollo comunitario en la India. El diseño otorga igual importancia a los enfoques CUANT y CUAL en todas las etapas de la evaluación.

Decisión 4: ¿El diseño de MM será en un único nivel o se utilizará un diseño de varios niveles?

Los diseños que presentamos hasta ahora funcionan en un único nivel, como la granja o el hogar. No obstante, los MM también son una herramienta eficaz para la evaluación de sistemas de entrega de servicios (por ejemplo, departamentos de

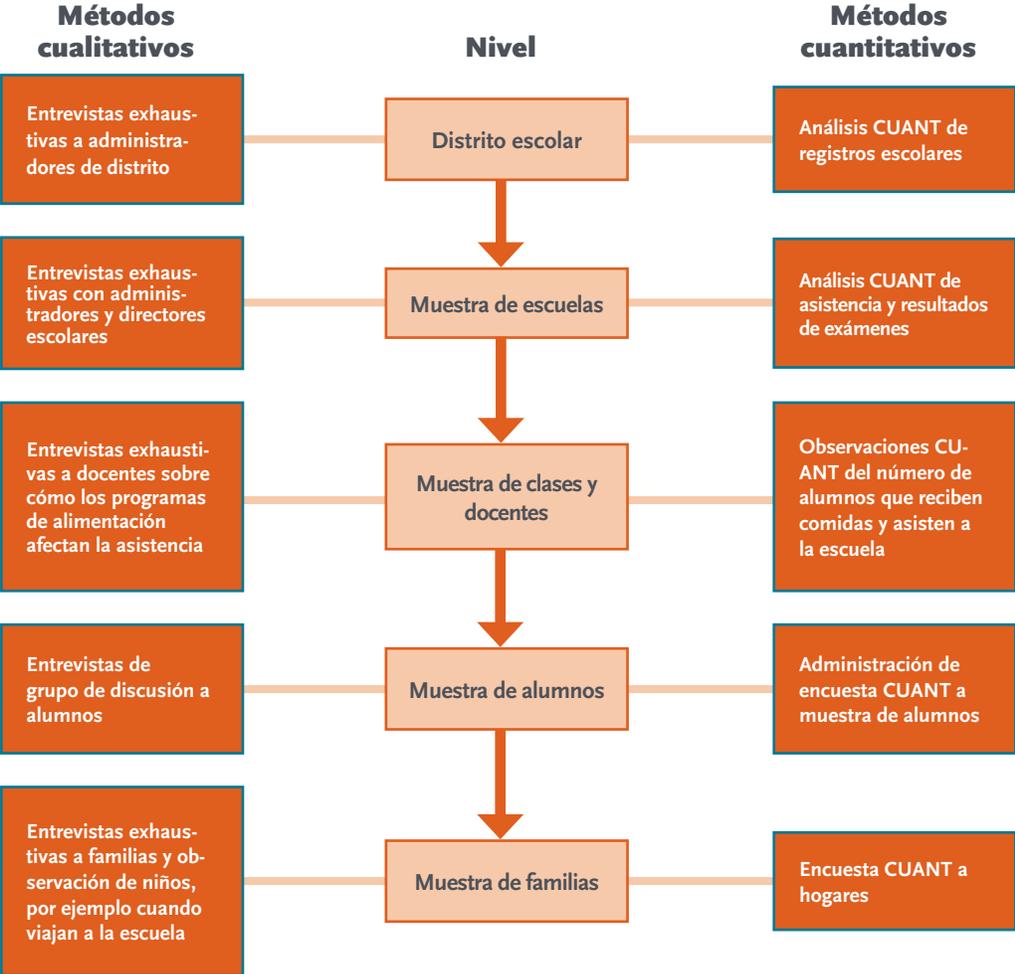
educación distritales, servicios de salud estatales o un programa nacional para fortalecer los gobiernos municipales) que requieren la descripción y el análisis de los vínculos entre los diferentes niveles. Estas evaluaciones se pueden tornar muy complejas y costosas. Los diseños de métodos mixtos que combinan datos CUANT y CUAL en cada nivel pueden a menudo proporcionar hallazgos válidos y creíbles sobre la base de muestras más pequeñas y más económicas.

La Figura 5 ilustra un diseño de métodos mixtos en varios niveles para evaluar los efectos de un programa de alimentación escolar en el desempeño y la asistencia. La evaluación debe recolectar datos a nivel del distrito escolar, una muestra de escuelas, una muestra de clases y docentes en cada escuela y una muestra de alumnos y familias. En cada nivel, se recolectan datos cuantitativos y cualitativos, y se comparan. Los métodos CUAL (como las observaciones, los grupos de discusión y las entrevistas con informantes clave) también pueden ayudar a examinar las conexiones entre los diferentes niveles (por ejemplo, interacciones entre funcionarios del distrito y docentes y administradores escolares).

2.2. Aplicación de enfoques de MM en cada etapa de la evaluación

Esta sección explica las diferentes maneras en las que generalmente se aplican los enfoques CUANT y CUAL en cada etapa de una evaluación, y cómo los dos enfoques se pueden combinar en un diseño de MM. Los [Anexos 4](#) y [5](#) proporcionan más detalles. Al leer esta sección, se debe comprender que para una evaluación con MM extensa y bien financiada, el equipo podría incluir la contratación de uno o más integrantes por sus conocimientos CUANT o CUAL específicos (por ejemplo, muestreo estadístico para un estudio a gran escala

Figura 5. Diseño de métodos mixtos en varios niveles: Evaluación de los efectos de un programa de alimentación escolar en la asistencia y el desempeño



sobre mala nutrición o entrevistas CUAL exhaustivas sobre temas sensibles como la violencia doméstica). Sin embargo, muchas evaluaciones (quizás la mayoría) no pueden darse este lujo y los integrantes del equipo deberán aplicar ambos enfoques, CUANT y CUAL, según se requiera. Es obvio que, incluso para las evaluaciones extensas, es deseable que todos los integrantes de un equipo estén familiarizados con todos los métodos de evaluación que se usarán. Los párrafos siguientes ilustran cómo los métodos CUANT y CUAL se pueden integrar en diferentes etapas de la evaluación.

- *Formulación de hipótesis.* Las evaluaciones CUANT derivan hipótesis *deductivamente* de teorías existentes o publicaciones bibliográficas, mientras que las evaluaciones CUAL elaboran hipótesis *inductivamente* a medida que progresa el estudio. Los MM combinan ambos enfoques. Por ejemplo, una hipótesis desarrollada deductivamente utilizando un enfoque CUANT se puede estudiar y perfeccionar a través de enfoques CUAL, como entrevistas u observaciones. Por otro lado, las etapas iniciales de la recolección de datos CUAL pueden describir procesos y problemas

que un enfoque CUANT puede probar a través de datos recolectados en una encuesta de muestra.

- **Muestreo.** Por lo general, las evaluaciones CUAL utilizan una cantidad relativamente pequeña de temas seleccionados *deliberadamente* (muestreo teórico) para garantizar que se incluyan todos los grupos importantes. Por el contrario, las evaluaciones CUANT habitualmente usan una muestra seleccionada al azar, lo que permite la generalización en poblaciones más grandes y la comparación estadística de diferentes grupos (por ejemplo, los grupos del proyecto y de comparación). El muestreo de MM utiliza el mismo marco de muestreo para generar una muestra de encuesta CUANT extensa y seleccionar una muestra pequeña pero representativa para un análisis CUAL exhaustivo. *Garantizar que las muestras CUAL sean razonablemente representativas del total de la población de muestra es una de las contribuciones más importantes de los diseños de MM.* El ejemplo de la evaluación de los centros de salud ilustra cómo se pueden complementar los enfoques CUANT y CUAL aplicados al muestreo. El [Anexo 6](#) resume las diferencias entre las estrategias de muestreo CUANT y CUAL.
- **Diseño de evaluación.** La mayoría de las evaluaciones CUANT emplean uno de una pequeña cantidad de diseños aleatorios o cuasiexperimentales. Siempre que sea posible, se entrevista a las muestras representativas del proyecto y a los grupos de comparación en dos o más puntos durante el transcurso del proyecto para comparar los cambios en resultados o indicadores de impacto. Por el contrario, las evaluaciones CUAL tratan de describir procesos constantes de cambio que a menudo son afectados por diferentes factores y que afectan a diferentes individuos o grupos en diferentes

maneras. Algunas evaluaciones CUAL procuran entender el programa a través del análisis de relaciones entre muchos elementos diferentes de la comunidad u otro entorno en el que funciona el programa, mientras otros adoptan un foco de atención más profundo en temas individuales sin necesariamente centrarse en el contexto más amplio. En algunas evaluaciones bien financiadas, el evaluador puede vivir en la comunidad o visitarla con frecuencia durante un tiempo, pero en la mayoría de los casos esto no es posible, y se pone el énfasis en los grupos de discusión, entrevistas exhaustivas, informantes clave, etc. Por lo general, las evaluaciones CUAL no tratan de establecer una relación de causa y efecto directa entre las intervenciones del proyecto y los resultados. Una de las muchas maneras en las que se pueden combinar dos enfoques es emplear métodos CUAL para estudiar el proceso de implementación del proyecto y la influencia de variables contextuales en el rendimiento del proyecto en algunas de las comunidades donde se está realizando una encuesta CUANT a los participantes del proyecto.⁶

- **Métodos de registro y recolección de datos.** La Tabla 2 enumera algunas de las técnicas CUANT y CUAL de recolección de datos más usadas. Mientras que las evaluaciones CUANT recolectan datos numéricos estandarizados, las evaluaciones CUAL a menudo utilizan métodos de recolección de datos menos estructurados que proporcionan mayor flexibilidad y procuran entender las complejidades de una situación. Aunque la ventaja de los métodos CUANT de recolección de datos es que producen datos estandarizados que miden cambios en el tiempo o entre grupos, estos métodos

⁶ Existe una cantidad de técnicas para transformar descripciones CUAL de factores contextuales en variables CUANT (por ejemplo, variables ficticias [Dummy]) que se pueden incorporar al análisis de regresión (consulte la Tabla 4).

Tabla 2. Métodos CUANT y CUAL de recolección de datos más usados

CUANT	CUAL
<ul style="list-style-type: none"> • Encuestas estructuradas de hogares, granjas, usuarios de servicios públicos, etc. • Guías de observación estructuradas • Medidas antropométricas de altura y peso • Análisis de HIS y anemia mediante exámenes y recolección de muestras de sangre • Recuentos automáticos (por ejemplo, personas que ingresan a un edificio) • Análisis sociométrico^{±**} • GIS (generación y análisis de mapas de GPS)^{**} • Programa Sistema de Información para la Gestión sobre datos de entradas y salidas • Revisión de datos institucionales: historias clínicas, registros escolares, etc. ^{**} 	<ul style="list-style-type: none"> • Entrevistas exhaustivas • Informantes clave • Observaciones de participantes • Observaciones de no participantes^{**} • Estudios de caso • Entrevistas al salir los cliente^{**} • Estudios de pacientes ficticios • Grabación de videos o audio^{**} • Fotografía • Análisis de documentos^{**} • Artefactos • Entrevistas grupales (p. ej., grupos de discusión, reuniones de la comunidad)^{**} • Técnicas grupales participativas (p. ej., evaluación rural significativa, cambios más importantes) • Entrevistas por Internet
<p>[±] Técnicas de encuesta para estudiar la formación de grupos, cómo se difunde la información, identificación de líderes de opinión y otros patrones de organización social en un grupo o en una comunidad.</p> <p>^{**} Indica que estas técnicas se pueden utilizar tanto cuantitativa como cualitativamente. Están ubicadas en la columna correspondiente al método que las utiliza más comúnmente.</p>	

no son adecuados para obtener información sobre temas sensibles o para entrevistar grupos que son difíciles de abordar. La recolección de datos con MM saca provecho de las ventajas de los datos CUANT, a la vez que examina con mayor profundidad, obtiene datos sensibles y estudia procesos y el cambio de comportamiento.

- **Triangulación.** Una característica clave de los MM es el uso sistemático de la triangulación (Anexo 9). Mientras que tanto los evaluadores CUAL como CUANT utilizan la triangulación para obtener dos o más estimaciones independientes de variables de resultados clave, los MM tienden a utilizarla más sistemáticamente y como parte integral del diseño de la evaluación.

Las evaluaciones CUANT utilizan la triangulación para incluir verificaciones de coincidencias en los instrumentos de encuesta o para comparar

fuentes de datos secundarias con la información provista por los encuestados. Las evaluaciones CUAL emplean la triangulación de una manera más general, pero generalmente con un énfasis en profundizar y ampliar la comprensión a través de las diversas perspectivas obtenidas de diferentes fuentes de información, en lugar de una verificación de coincidencias. Los diseños de MM triangulan estimaciones CUANT y CUAL (consulte la Tabla 3). Los MM utilizan la información obtenida a través de la triangulación para: aumentar la confiabilidad y la validez de las estimaciones de indicadores clave al comparar información de diferentes fuentes; profundizar la comprensión del significado de las relaciones estadísticas identificadas en el análisis cuantitativo y garantizar que se capten y se comparen las perspectivas de todas las partes interesadas clave, con particular énfasis en grupos pobres y vulnerables. Si las estimaciones obtenidas de diferentes fuentes coinciden, esto aumenta la validez y la credibilidad de los datos,

Tabla 3. Tipos diferentes de triangulación utilizados en evaluaciones con métodos mixtos

Método	Ejemplos
Uso de diferentes marcos conceptuales	Comparar marcos de análisis feministas, de derechos humanos, de exclusión social o económicos (p. ej., costo-beneficio)
Diferentes métodos de recolección de datos	Comparar encuestas estructuradas, observaciones directas, datos secundarios, artefactos
Diferentes entrevistadores	Comparar sexo, edad, origen étnico, condición económica, vestimenta, lenguaje del entrevistador, etc., en las respuestas
Diferentes momentos	Comparar respuestas u observaciones en diferentes momentos del día, días de la semana, épocas del año
Diferentes lugares y contextos	Comparar respuestas y observaciones cuando las entrevistas se realizan en el hogar en presencia de otras personas, en lugares donde el entrevistado puede hablar con mayor libertad, en la calle y otros lugares públicos, en el trabajo, en el aula

particularmente de las estimaciones basadas en muestras pequeñas, y produce estimaciones más confiables que si todos los demás recursos se hubieran invertido en una técnica particular, como la encuesta a los hogares.

- *Análisis de datos.* Los evaluadores CUAL emplean una amplia variedad de métodos de análisis de datos para identificar patrones y relaciones generales a fin de obtener una visión holística de las complejas interacciones entre un proyecto y el entorno en el cual está insertado. El objetivo del análisis CUANT, por otro lado, es describir las características estadísticas de variables clave para determinar

la importancia estadística de las diferencias entre el grupo del proyecto y el grupo de comparación, e identificar factores que contribuyen a la magnitud y dirección del cambio. El análisis de datos con métodos mixtos emplea análisis CUAL para tratar de comprender el significado que otorgan los diferentes sujetos o grupos a las asociaciones estadísticas halladas en el análisis CUANT y para proporcionar casos y ejemplos a fin de ilustrar los hallazgos. Por otro lado, se puede usar el análisis CUANT para evaluar cómo los casos incluidos en los estudios CUAL representan al total de la población de interés y qué sectores no se incluyeron.

Tabla 4. Ejemplos del análisis de datos con métodos mixtos

Enfoque	Descripción	Ejemplo
A. Análisis paralelo de datos con métodos mixtos	Esto incluye dos procesos de análisis independientes: los datos CUANT se analizan utilizando métodos CUANT convencionales (como tablas de frecuencia, tablas cruzadas, análisis de regresión, etc.) mientras que se realiza otro análisis independiente de datos CUAL utilizando métodos CUAL, como el análisis del contenido. Luego se comparan los hallazgos de los dos conjuntos de análisis.	En la Evaluación de la pobreza de 2003 del Banco Mundial en Guatemala, equipos independientes fueron los responsables de recolectar datos CUAL y CUANT. El análisis CUAL se realizó en 5 pares de pueblos, que representaban los principales grupos étnicos. Los datos CUANT del mismo conjunto de pueblos se analizaron por separado y los dos conjuntos de datos se integraron solo en la etapa final del análisis. La combinación de los dos análisis independientes proporcionó un contexto político e histórico más amplio para comprender los impactos del funcionamiento del programa (Teddlie y Tashakkori [2009] Cuadro 11.5).
B. Análisis de datos con métodos mixtos de conversión	<p>a. Los datos CUAL se convierten en indicadores CUANT (“se cuantitativizan”) utilizando calificación, puntaje y escalas* para que se puedan usar técnicas de análisis CUANT.</p> <p>b. Los datos CUANT se convierten en indicadores CUAL (“se cualitativizan”) para que se puedan usar procedimientos de análisis CUAL.</p>	<p>c. a. Los datos sobre el contexto político, económico, social, ambiental, legal y administrativo en el cual funciona un proyecto a menudo se presentan en forma cualitativa y redactada. Los indicadores se pueden “cuantitativizar” a través de una conversión a variables ficticias (Dummy). Por ejemplo: “la economía está creciendo” = 1, “la economía no está creciendo” = 0. Estas variables ficticias se pueden incorporar al análisis de regresión.</p> <p>d. b. En la Figura 4, una tipología cuantitativa de agricultores se podría “cualitativizar” al elaborar descripciones redactadas de las diferentes actitudes hacia la adopción de nuevas variedades de semillas.</p>
C. Análisis de datos secuencial con métodos mixtos	<p>a. Después del análisis de datos CUAL sigue el análisis CUANT.</p> <p>b. Después del análisis de datos CUANT sigue el análisis CUAL.</p> <p>c. Diseños de MM iterativos. El análisis incluye pasos CUANT y CUAL secuenciales.</p>	<p>a. a. En la Figura 3, el estudio de las estrategias de supervivencia comienza con un análisis cualitativo de informes redactados sobre los patrones de transferencias entre hogares que funcionan como soporte de hogares vulnerables. El análisis CUAL colabora con el diseño de la encuesta CUANT de transferencias entre hogares, que luego se analiza mediante técnicas econométricas.</p>
D. Análisis con métodos mixtos en varios niveles	Las técnicas de análisis CUANT y CUAL se utilizan en diferentes niveles de un diseño de evaluación de varios niveles.	La Figura 5 ilustra el uso de un análisis con MM en varios niveles para evaluar los impactos de un programa de alimentación escolar en el desempeño y la asistencia. Tanto el análisis CUANT como el análisis CUAL se realizaron de manera secuencial en el nivel del distrito, la escuela, el aula, el alumno y la familia. Esto permitió un análisis de las conexiones entre los diferentes niveles.**

* Por ejemplo, los informes redactados acerca de las actitudes de los grupos políticos locales hacia un programa de desarrollo social se podrían convertir en una escala numérica donde: 3 = “el grupo político tiene una actitud favorable hacia el programa”; 2 = “el grupo político no está a favor ni se opone” y 1 = “el grupo se opone al programa”. La calificación, el puntaje o las escalas son maneras un tanto diferentes de realizar la conversión.

** Por ejemplo, el análisis de los registros en el nivel del distrito puede identificar las escuelas que están por encima o por debajo de los promedios de asistencia o puntajes de desempeño. Esta información se puede utilizar para seleccionar las escuelas por encima y por debajo del promedio que se incluirán en la muestra. De manera similar, se podrían utilizar entrevistas exhaustivas con docentes para seleccionar una muestra de alumnos con características particulares de interés para incluirlos en los grupos de discusión.

Fuente: Adaptado de C. Teddlie y A. Tashakkori 2009, *Foundations of Mixed Methods Research (Fundamentos de la investigación con métodos mixtos)*. Capítulo 11, Sage Publications (con autorización). La mayoría de los ejemplos fueron desarrollados por este autor.

Parte III: Aplicaciones de los diseños de métodos mixtos

3.1. Estrategias de muestreo para evaluaciones con MM orientadas CUANT y CUAL

Modelo 1: Uso de métodos mixtos para afianzar una evaluación principalmente CUANT

Con la mayoría de los diseños de evaluación CUANT, las encuestas de muestras se administran antes y después de la intervención del proyecto con un grupo de control o de comparación. El tamaño de muestra requerido para las encuestas CUANT se calcula a través de la dimensión del efecto y la capacidad estadística.⁷ Los métodos CUAL se pueden emplear para afianzar el diseño en uno o más de los siguientes puntos:

- **Estudios exploratorios o diagnósticos para comprender el contexto y los problemas antes de desarrollar los instrumentos de la encuesta.** Estos pueden ser estudios cualitativos rápidos con una duración de solo unos días o estudios más prolongados en los que un antropólogo o sociólogo se queda a vivir en una comunidad durante un período de semanas o meses. En un proyecto grande, que funciona en diferentes regiones geográficas o ecológicas, podrían necesitarse estudios diagnósticos en una cantidad de regiones o comunidades diferentes. A veces, los estudios serán llevados a cabo por un investigador solo

⁷ La dimensión del efecto se refiere a la dimensión del cambio o del impacto que se calcula. Cuanto mayor es el cambio que se estima, menor es el tamaño requerido para la muestra. La capacidad estadística se refiere a la probabilidad de que la prueba estadística identifique correctamente que existe un impacto real. Cuanto más alto es el nivel de confianza requerido, mayor deberá ser el tamaño requerido para la muestra. Consulte Bamberger et al 2012, Capítulo 15, Sección 4 para obtener un análisis de la estimación del tamaño de la muestra.

(un investigador por comunidad o región) mientras que en otros casos el investigador principal podría recibir ayuda de un equipo de asistentes para realizar encuestas rápidas o estudios de observación de participantes o no participantes sobre, por ejemplo, sistemas de transporte de la comunidad, el uso del tiempo por parte de las mujeres, o sistemas de producción y marketing. Aunque los estudios pueden tardar varios días o semanas, la unidad de muestreo primaria, por lo general, será una comunidad o un grupo, y normalmente solo se estudiarán algunos grupos o comunidades. No obstante, es posible entrevistar a grandes cantidades de personas utilizando técnicas de recolección de datos no estructurados o semiestructurados y, en algunos casos, también se pueden realizar encuestas de muestras rápidas.

- **Grupos de discusión formados por diferentes segmentos de la población objetivo.** Estos pueden llevarse a cabo durante la etapa preparatoria de la evaluación o después de que se hayan analizado las encuestas cuantitativas y se hayan identificado los grupos de interés principales. Idealmente, se deberán realizar tres o cuatro grupos de discusión con cada grupo económico o demográfico de interés para la evaluación (Teddlie y Tashkkori 2009, Tabla 8.5), aunque la cantidad de grupos, por lo general, es menor cuando se trabaja con limitaciones de tiempo y de presupuesto.
- **Se pueden agregar módulos semiestructurados y especializados a una encuesta de muestra, y administrar a una submuestra de encuestados.** Por ejemplo, se puede

administrar la encuesta principal al jefe/a del hogar (quien en muchas culturas suele ser un hombre), pero en una submuestra de hogares se puede entrevistar al cónyuge. A veces, el mismo entrevistador puede administrar el módulo especial al cónyuge (u otro integrante del hogar), pero en muchos casos será necesario preparar una entrevista aparte, a menudo en un lugar o en un momento cuando el marido u otros integrantes del hogar no estén presentes. Por lo general, los módulos se administran del 10 al 25 por ciento de la muestra original.

- **Preparación de estudios de caso para complementar la encuesta.** A menudo, resulta útil preparar estudios de caso en una muestra pequeña de encuestados para proporcionar una comprensión más acabada de los temas de interés para la evaluación. Por ejemplo, en la evaluación de un programa de agricultura, se podrían elaborar estudios de caso con el fin de ilustrar diferentes clases de sistemas agrícolas. Para un proyecto educativo, los casos podrían comprender familias con ingresos más altos y más bajos, familias que viven cerca y lejos de la escuela o familias de diferentes grupos étnicos o religiosos. Una vez más, el número de casos normalmente será bastante reducido, aunque la duración puede ser bastante prolongada. Cuando se preparan estudios de caso en organizaciones (como escuelas o cooperativas agrícolas) o comprenden comunidades enteras (por ejemplo, para ilustrar los efectos de sistemas de transporte mejorados), el estudio será más complicado y, por lo general, se entrevistará a una cantidad significativa de personas para cada caso.

Modelo 2: Uso de diseños de métodos mixtos para afianzar una evaluación CUAL

Los diseños de métodos mixtos también se

pueden usar para afianzar una evaluación CUAL que utiliza grupos de discusión, observaciones de participantes, observaciones de no participantes y preparación de estudios de caso. Un desafío para muchos de estos diseños es el peligro de la parcialidad si las muestras de individuos o grupos no son representativas. Por ejemplo, a menudo las personas que asisten a los grupos de discusión son aquellas que están muy a favor o en contra de un proyecto, aquellas que tienen el tiempo y los recursos para participar (pueden tener que hacer arreglos de transporte), o bien, en algunos casos (por lo general, sin que lo sepa el evaluador), algunos participantes pueden ser enviados por el gobierno local u otro grupo con un interés particular. En consecuencia, es difícil incorporar mucha información valiosa y esclarecedora al informe de evaluación de una manera creíble. De modo similar, muchos investigadores se sienten más cómodos al hablar con ciertos grupos que con otros; es así que puede haber cierta parcialidad en la selección de los estudios de caso. Por lo general, la muestra es relativamente pequeña pero la cantidad de entrevistas varía según el tamaño de la población estudiada y el nivel de precisión requerido para los hallazgos.⁸

⁸ Consulte Bamberger et al. 2012, Tabla 15.6, p. 389 si desea obtener reglas generales para calcular el tamaño de las muestras para diferentes clases de métodos de recolección de datos CUAL. No obstante, el tamaño de las muestras varía según el tamaño y la complejidad del programa que se está evaluando y el nivel de precisión requerido para las estimaciones. Si se deben realizar generalizaciones a partir de los estudios de caso, grupos de discusión u otros métodos, es importante asegurar que se seleccionen casos que sean razonablemente representativos y también que se incluyan suficientes casos para que los hallazgos se consideren creíbles. De modo que, aunque uno o dos casos puedan ser valiosos para ilustrar procesos o comportamientos, rara vez resulta apropiado utilizar una muestra tan pequeña para realizar afirmaciones como “La mayoría de los agricultores sintieron que...” o “La mayoría de las madres creían que los centros de salud...”. No es posible usar procedimientos estadísticos convencionales para calcular el tamaño de muestra apropiado con una cantidad tan reducida de casos, de modo que, inevitablemente, la opinión se debe combinar con la consulta a las partes interesadas respecto de qué tamaño de muestra se consideraría razonable o creíble.

El muestreo con métodos mixtos garantiza que se seleccionen casos CUAL e informantes que sean representativos en líneas generales de la población total. Esto afianza la validez de los hallazgos generales. Las técnicas cuantitativas, como una encuesta de muestra rápida, pueden ser una forma útil de comparar las características socioeconómicas de los individuos o grupos incluidos en los estudios cualitativos con las características de la población total. Por lo general, el tamaño de muestra CUANT es relativamente pequeño, ya que la encuesta solo se utiliza para garantizar que los estudios de caso u otros datos CUAL sean razonablemente representativos. Pero si se requieren estimaciones más precisas, se pueden emplear los procedimientos de cálculo del tamaño de la muestra analizados en la sección anterior.

Modelo 3: Uso de un diseño de método mixto equilibrado (integrado)

Aunque en la mayoría de los casos, los diseños de método mixto se utilizan para complementar un diseño predominantemente CUANT o CUAL, hay casos en los se podría utilizar un diseño de método mixto integrado que otorgue igual importancia a ambos enfoques. Estos diseños pueden consistir en la combinación de técnicas cuantitativas y cualitativas, además de técnicas de métodos mixtos específicos, en diferentes etapas de la evaluación. La estimación de los tamaños de muestra requeridos debe combinar los procedimientos CUANT para las encuestas de muestra y la reglas generales analizadas en la sección anterior para los componentes CUAL. La combinación de ambos requiere un criterio basado en la experiencia. El ejemplo 3 de la Parte V describe un diseño de MM equilibrado (integrado) utilizado para evaluar un programa en la India a gran escala cuya finalidad es aumentar la capacidad de las comunidades para administrar subvenciones de los gobiernos estatales. Las encuestas de muestra, utilizadas para obtener estimaciones cuantitativas de los resultados de

un programa, se combinaron con la recolección de datos cualitativos a través de visitas periódicas a una muestra de comunidades del proyecto y de control para observar los procesos de cambio y para realizar entrevistas en una pequeña muestra de hogares, así como a funcionarios locales y a informantes clave. Se empleó el mismo marco de muestreo tanto para las muestras CUANT como las CUAL, y se seleccionaron estas últimas para garantizar que fueran representativas.

3.2. Uso de métodos mixtos para evaluar intervenciones complejas

Por lo general, se define a las intervenciones de desarrollo complejas⁹ como intervenciones que tienen algunas de las siguientes características:

- Planificación y evaluación dirigidas por el país.
- El programa progresa de una manera no lineal.
- Muchos componentes o servicios y varios objetivos.
- Hay resultados cuantitativos y cualitativos.
- Las poblaciones objetivo son difíciles de identificar o controlar.
- Se brinda apoyo presupuestario general sin una definición clara de los servicios que se financiarán.
- Varios donantes y agencias.
- Se emplean procesos de implementación y planificación participativa y, por consiguiente, los resultados e impactos son difíciles de definir de manera anticipada y pueden cambiar constantemente.
- El contexto es complejo.
- La situación cambia de manera constante y la intervención trata de adaptarse a estos cambios.

⁹ Para obtener más información sobre intervenciones complejas y su evaluación, consulte Funnell y Rogers (2011), Patton (2011) y Bamberger et al, (2012), Capítulo 16.

CUADRO 3. ANÁLISIS DE ATRIBUCIÓN, CONTRIBUCIÓN Y SUSTITUCIÓN

Aunque la evaluación del proyecto a veces puede aplicar diseños experimentales o cuasiexperimentales que puede controlar para brindar explicaciones alternativas de los cambios observados (resultados o impactos) en el grupo del proyecto, rara vez es posible lograr el mismo nivel de rigor en la evaluación de los efectos de intervenciones complejas. En consecuencia, los evaluadores de los efectos de intervenciones complejas deben decidir cuál de los siguientes niveles de análisis se pueden aplicar:

Análisis de atribución: se establece correspondencias entre el grupo del proyecto y el grupo de comparación de modo que se pueden controlar (eliminar) las explicaciones alternativas de los cambios observados. Si hay diferencias estadísticamente significativas entre el grupo del proyecto y el grupo de comparación, esto se toma como evidencia de que la intervención del proyecto contribuyó a estos cambios.

Análisis de contribución: el análisis evalúa la contribución de una agencia de desarrollo particular al logro de los cambios generales que son el resultado de las intervenciones financieras y técnicas de una cantidad de agencias de desarrollo diferentes que ofrecen su colaboración. A veces, el análisis también incluirá los programas de agencias nacionales (gubernamentales y de la sociedad civil).

Análisis de sustitución: se realiza una evaluación del *aumento neto en recursos* en el programa o sector objetivo que resulta de la contribución de una agencia de desarrollo en particular. Esto tiene en cuenta las desviaciones de recursos a otras actividades por parte del gobierno nacional (reduciendo el aumento neto en recursos) o cualquier efecto multiplicador (incrementando el aumento en recursos).

Fuente: adaptado de Bamberger et al 2012, Cuadro 16.3 pp. 403-4.

- Existen procesos complejos de cambio de comportamiento que abarcan a la población objetivo, los proveedores de servicios y otras agencias, por ejemplo, la policía, el ejército y el poder judicial (Sección 3.3).
- Mientras que algunas intervenciones complejas son extensas, otras son relativamente reducidas pero comprenden procesos complejos de cambio de comportamiento o poblaciones que son difíciles de identificar o estudiar.

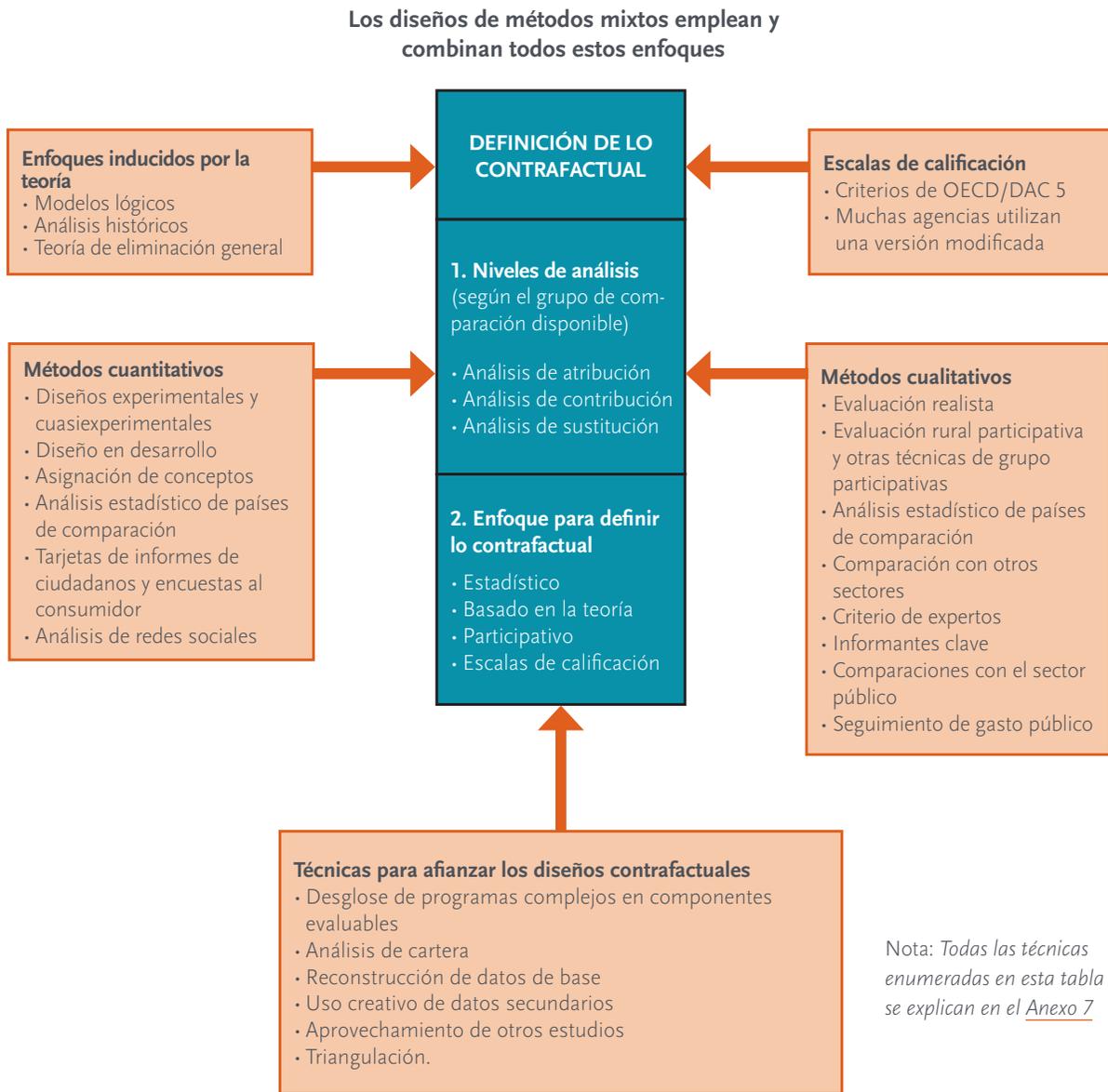
Los diseños de métodos mixtos son aptos para evaluar las intervenciones complejas, ya que se puede combinar el diseño estandarizado, la recolección de datos y el análisis con herramientas que pueden captar las complejidades del entorno del programa, la naturaleza cambiante de este y los resultados y procesos de cambio de comportamiento deseados que están teniendo lugar. También es posible combinar las perspectivas de diferentes partes interesadas con indicadores cuantitativos “objetivos”, y existe

una variedad de herramientas CUAL para recolectar datos sensibles de poblaciones difíciles de abordar. Al mismo tiempo, los MM ofrecen maneras de calcular los impactos cuantitativos de la forma más creíble posible en circunstancias difíciles y de utilizar la triangulación para aumentar la validez del constructo de los resultados calculados y de la interpretación de lo que realmente está sucediendo en el terreno.

La Figura 6 resume la amplia variedad de herramientas y técnicas a las que puede recurrir un diseño de MM al seleccionar el diseño apropiado para evaluar una intervención compleja (todas las técnicas enumeradas en esta figura se describen en el [Anexo 7](#). Un elemento central del enfoque de métodos mixtos es ampliar las maneras en las que se puede definir y calcular lo contrafactual.¹⁰ Lo anterior se hace en dos pasos.

¹⁰ La GO1 (p. 1) define lo contrafáctico como “un cálculo de lo que habría pasado si la intervención no hubiese ocurrido [...] por ejemplo, comparaciones con un grupo que no recibió la intervención”.

Figura 6. Uso de diseños de métodos mixtos para evaluar intervenciones complejas



Paso 1: Según el tipo de grupo de comparación que está disponible, la evaluación realizará el análisis en uno de los siguientes niveles¹¹ (consulte el Cuadro 3):

- Análisis de atribución
- Análisis de contribución
- Análisis de sustitución

Luego, el Paso 2 selecciona el enfoque, o la combinación de enfoques, que se utilizará para definir y calcular lo contrafactual:

- *Grupo de comparación estadística*: se utiliza un grupo de comparación o control estadísticamente correspondiente.
- *Basado en la teoría*: el proceso de implementación y los resultados definidos en la teoría de cambio se comparan con lo que se observa en el terreno. Cuanto más estrecha es la correspondencia entre la realidad y la teoría de cambio, más fuerza cobra el supuesto de que la intervención contribuyó a los resultados. No obstante, también es importante definir y comprobar explicaciones alternativas de los cambios observados (hipótesis rival).¹²
- *Basado en la participación*: los efectos del programa se calculan a través de la evaluación por parte de los beneficiarios a quienes está dirigido el programa y otras partes interesadas. Un ejemplo de este enfoque es el Cambio

¹¹ Consulte la GO₁, sección 9, para obtener más información sobre una amplia variedad de enfoques para evaluar la atribución y la contribución. Bamberger et al (2012) pp. 403-405 proporciona un debate más extenso que incluye un análisis de sustitución.

¹² Un punto débil de muchas teorías de cambio es que no identifican explicaciones alternativas de los resultados observados. Si los cambios coinciden con la teoría de cambio, esto se toma como evidencia de que los cambios se pueden atribuir a los efectos del proyecto, lo que no es una conclusión válida. Una teoría de cambio debería identificar explicaciones alternativas (hipótesis rivales) y crear en el modelo formas de comprobar estas explicaciones rivales de las causas de los cambios observados.

más importante, y otro es el uso de la evaluación rural participativa, además de otras técnicas de análisis de grupo.

- *Escalas de calificación*: se les pide a expertos o a partes interesadas que califican el funcionamiento del programa o el grado de cambio producido por la intervención según un conjunto de escalas de calificación. Las escalas del DAC y de la OECD para evaluar la relevancia, la eficiencia, la efectividad, el impacto y la sostenibilidad son un ejemplo de escalas de calificación de uso común.

Los diseños de métodos mixtos pueden combinar, según corresponda: enfoques inducidos por la teoría, enfoques cualitativos y estadísticos, y escalas de calificación. El [Anexo 7](#) enumera la variedad de opciones disponibles para cada uno de estos enfoques. Se puede emplear una cantidad de técnicas para afianzar el diseño contrafactual (desglose de intervenciones complejas, análisis de cartera, reconstrucción de datos de base, uso creativo de datos secundarios y aprovechamiento de otros estudios). La validez del constructo de las estimaciones de todas estas fuentes se refuerza por medio de la triangulación.

No existe un único enfoque que sea el más indicado para una evaluación con métodos mixtos, ya que los evaluadores deben seleccionar un conjunto de herramientas y técnicas que sean las más adecuadas según las limitaciones políticas, de presupuesto, de tiempo y de datos, y los objetivos de cada evaluación. Los estudios de caso enumerados en la Parte V ilustran la variedad de enfoques de métodos mixtos que se utilizaron para diferentes evaluaciones. La creatividad, el conocimiento de una amplia variedad de enfoques CUANT, CUAL y basados en la teoría, y la predisposición para recurrir a diferentes disciplinas son requisitos esenciales de las evaluaciones con métodos mixtos.

CUADRO 4. LOS RESULTADOS DEL PROGRAMA SE VEN AFECTADOS POR LA MANERA EN QUE LOS BENEFICIARIOS RESPONDEN A LOS SERVICIOS Y CÓMO REACCIONA EL PERSONAL ANTE LAS RESPUESTAS DE LOS BENEFICIARIOS. ESTUDIO DEL CAMBIO DE COMPORTAMIENTO EN UN PROGRAMA DE ALIMENTACIÓN ESCOLAR EN NICARAGUA.

Se organizó un programa de desayuno escolar en una zona rural inseguro respecto de los alimentos en Nicaragua para aumentar la matrícula escolar. En algunas escuelas, los docentes también alimentaban a los hermanos más pequeños que llegaban con las madres que llevaban a los alumnos a la escuela. Se corría la voz entre las madres y a medida que la aumentaba la cantidad de hermanos, algunos docentes les daban el desayuno a todos los hermanos y otros docentes no lo hacían. La decisión dependía en gran medida no solo de la actitud de cada docente, sino también de la predisposición de las madres a preparar los desayunos y de los agricultores locales, a donar más alimentos. En algunas comunidades, estas respuestas de comporta-

miento transformó el programa de desayuno escolar en un programa de nutrición importante que implicó la movilización de alimentos de los agricultores locales, hombres que ayudaban a transportar grandes cantidades de alimentos, la construcción de cocinas escolares y, en algunos casos, la instalación de energía eléctrica en las aulas, mientras que las madres organizaban los desayunos y, a veces, hacían huertas escolares. Los resultados finales en cada escuela y las lecciones aprendidas solo se pudieron determinar mediante la observación exhaustiva del comportamiento de los docentes, los padres y las comunidades, y pareció variar según las actitudes individuales, las condiciones económicas locales y, en ocasiones, el origen étnico.

3.3. Evaluación de procesos de cambio de comportamiento

Existen muchos proyectos en los que el proceso de implementación está menos definido y es menos lineal de lo que parecía al principio. Esto significa que la evaluación de impacto debe tener en cuenta estos procesos de cambio de comportamiento, ya que a menudo generan programas que tienen una cantidad de resultados e impactos no deseados. Estos son algunos de los motivos por los que esto ocurre:

- En la mayoría de los casos, los supuestos beneficiarios deciden activamente si participarán o no en el programa y, a menudo, eligen qué servicios utilizarán y cuáles no.
- Por lo general, los programas se modifican según la manera en que los beneficiarios iniciales perciben y responden a los servicios y las opiniones que brindan a otros. Muchos diseños de programas se modifican como resultado de estas interacciones dentro de la población objetivo y entre la población y los prestadores de servicios (consulte el Cuadro 4).
- El personal del proyecto tiene una actitud propia hacia diferentes grupos, creencias sobre

quiénes se beneficiarán y quiénes no, y una manera propia de interactuar con la comunidad.¹³ Aunque el personal reciba una capacitación rigurosa, cada persona reaccionará de una manera levemente distinta.

- Cada proyecto funciona en un entorno único donde la implementación y los resultados son afectados por un conjunto único de factores sociales, económicos, políticos, ambientales y legales, entre otros.¹⁴

13 Una revisión de los factores que afectan el éxito de los sistemas educativos estatales brasileños en aumentar el acceso de las familias de bajos ingresos determinó que muchos docentes creían que los niños de hogares pobres tenían una alta probabilidad de abandonar la escuela, e incluso si continuaban en la escuela, tendrían malas notas y mala disciplina. Por consiguiente, muchos docentes no sentían que valiera la pena el esfuerzo de ayudar a alentar a niños de hogares pobres. Bamberger y Segone (2011) alegan que muchas causas de inequidad se determinan socialmente ya que los gobiernos y la opinión pública pueden no desear alentar el acceso de diferentes grupos de inmigrantes o minorías étnicas a los servicios públicos.

14 El análisis de sistemas proporciona marcos útiles para comprender el funcionamiento de estos actores contextuales y también para entender cómo la historia afecta las actitudes y expectativas de las partes interesadas y las comunidades con respecto a las nuevas intervenciones del proyecto. El espacio no permite tratar el análisis de los sistemas. Si desea obtener una breve introducción al análisis de sistemas y cómo puede utilizarse en la evaluación de impacto, consulte Bamberger (2011), *How to design and manage equity focused evaluations* (Cómo diseñar y administrar evaluaciones orientadas al capital), Sección 5.2.C.

Las siguientes son algunas de las maneras en que los diseños de métodos mixtos pueden evaluar estos programas:

- Un estudio de diagnóstico rápido realizado al inicio de la evaluación puede contribuir a aclarar el entorno en el cual se implementa el programa y los factores contextuales que probablemente afectarán la implementación. El estudio también puede describir la estratificación social e identificar grupos marginales y vulnerables que podrían estar excluidos del acceso a los beneficios del proyecto. El análisis CUAL inicial se puede combinar con una encuesta CUANT rápida para calcular la magnitud y la distribución de, por ejemplo, grupos vulnerables.
- El alcance del monitoreo del programa convencional se puede extender para proporcionar más información CUANT detallada acerca de las características de los grupos sociales que usan y no usan los servicios del proyecto.
- El marco de análisis de cuello de botella desarrollado por UNICEF puede ofrecer un marco más riguroso para el análisis de los factores que determinan qué sectores tienen y no tienen acceso al proyecto, y los factores que actúan sobre la oferta y la demanda, y condicionan el acceso.¹⁵
- Hay una serie de técnicas CUAL disponibles para observar lo que realmente sucede durante la implementación de un proyecto, cómo responden los diferentes grupos al proyecto y cómo el diseño y el plan de implementación originales se ven modificados por la interacción entre diferentes sectores de la población objetivo y el personal del proyecto. Estas técnicas incluyen, pero no se limitan a: observación de participantes; estudios de panel en los que se visita periódicamente a una pequeña muestra de individuos, hogares o comunidades durante todo el proyecto; grupos de discusión, técnicas de evaluación rural participativa o autoinformes.¹⁶

15 Consulte Bamberger y Segone (2011), *How to design and manage equity-focused evaluations* (Cómo diseñar y administrar evaluaciones orientadas al capital), pp. 45-50. UNICEF. Disponible en http://mymande.org/?q=equity_focused_evaluations_intro.

16 Los ejemplos de las técnicas de autoinforme incluyen solicitar a una pequeña cantidad de participantes que lleven diarios personales donde registren experiencias y actividades de relevancia para el proyecto, o pedirles a las familias que lleven un registro de ingresos y gastos. En algunos estudios, se distribuyen teléfonos celulares entre los participantes para que llamen e informen, por ejemplo, sus patrones de viaje, lugares donde se sienten inseguros o tentados a comprar drogas, etc.

Parte IV: Administración de evaluaciones con métodos mixtos

4.1. Los diseños de métodos mixtos requieren un enfoque de administración especial

Aunque algunas aplicaciones de diseños de métodos mixtos implican solo incorporar métodos de recolección de datos adicionales a un diseño predominantemente CUANT o CUAL, una estrategia de evaluación con MM totalmente integrada implica mucho más que esto. Para disfrutar de todos los beneficios de combinar enfoques y métodos CUANT y CUAL, es necesario planificar una estrategia de evaluación con MM desde el primer día. Esto exige un compromiso de la gerencia para garantizar el tiempo y los recursos adicionales requeridos en una implementación eficaz del enfoque, y para defenderlos de las presiones administrativas de recortar presupuestos y tiempo. Esto, por lo general, requiere de una campaña energética por parte del gerente de la evaluación para asegurar la aceptación de la gerencia sénior y las agencias financiadoras. Las instrucciones en las reuniones de administración, los informes breves o capacitaciones sobre métodos, los boletines informativos, videos breves, etc., pueden ser útiles herramientas de defensa.

El equipo de evaluación también debe poder producir evidencia convincente para demostrar que los MM generan beneficios adicionales. Por ejemplo, se podría seleccionar una cantidad reducida de evaluaciones típicas con tiempo y recursos adicionales aprobados para evaluar el enfoque de MM y para comparar los resultados con enfoques de evaluación estándar utilizados en programas similares. ¿Es posible demostrar que el tiempo y los recursos adicionales requeridos para un enfoque de MM agregan valor?

Además de asegurar suficiente tiempo y recursos, la administración de una evaluación con MM necesita atención especial a las siguientes áreas (consulte el Cuadro 5):

Composición del equipo de investigación. Un equipo de investigación ideal debería incluir investigadores principales de dos o más disciplinas (por ejemplo, antropología, medicina, derecho, sociología, economía). No obstante, para la mayoría de las evaluaciones, las limitaciones de recursos no permitirán esto y se deben encontrar maneras de asegurar que los integrantes del equipo de evaluación puedan manejar todas las herramientas y técnicas CUANT y CUAL básicas. Un desafío para el gerente de la evaluación es contribuir a que se adquieran estos conocimientos básicos, por lo general, con el apoyo de consultores u otras agencias. Cuando los integrantes del equipo tienen aptitudes y orientaciones profesionales diferentes, es importante permitir cierto tiempo y brindar oportunidades para que cada investigador se familiarice con la metodología de los demás y que los integrantes del equipo desarrollen confianza y respeto mutuo. Esto es aún más importante cuando el equipo de evaluación está formado por especialistas de diferentes países, además de diferentes disciplinas. La Tabla 5 indica algunas de las áreas adicionales de experiencia profesional en investigación que se pueden requerir en equipos con una orientación CUANT o CUAL.

Administración de enfoques de métodos mixtos durante el diseño de la evaluación. El apoyo de la gerencia puede ser necesario para asegurar que el marco de la evaluación emplee todas las disciplinas incluidas y que las preguntas y los temas

CUADRO 5. PLANIFICACIÓN Y PRESUPUESTO DE TIEMPO Y RECURSOS ADICIONALES QUE PUEDEN REQUERIRSE PARA UNA EVALUACIÓN CON MÉTODOS MIXTOS

Las evaluaciones con MM suelen necesitar tiempo (p. ej., para la formación de equipos) y dinero adicionales. El gerente debe decidir si los recursos extras están justificados y, si es así, garantizar que se incluyan en el presupuesto.

Composición e integración del equipo de investigación:
El gerente de la evaluación desempeña un papel fundamental en asegurar que los integrantes del equipo de diferentes disciplinas trabajen juntos con eficacia. Esto requiere tiempo y esfuerzo extras, y será necesario.

- Dar tiempo a los investigadores para que desarrollen la comprensión y el respeto por las disciplinas y el trabajo de los demás. Asegurar que todos los integrantes del equipo conozcan el material bibliográfico básico y los debates actuales del otro campo.
- Garantizar formación de equipo y conexiones similares para los investigadores locales.

Enfoques integrados durante el diseño de la evaluación:
El gerente de la evaluación debe alentar activamente a todos los integrantes del equipo a aprovechar al máximo los nuevos marcos, la recolección de datos y las herramientas analíticas. Para lograrlo, deberá:

- Asegurar que se convoque a investigadores de diferentes disciplinas en la etapa del diseño de la evaluación o aproximadamente al mismo tiempo, de modo que todos sientan que están haciendo una contribución importante a la evaluación general y no crean que son simplemente la “frutilla del postre”. Esto es especialmente importante cuando predomina una disciplina.
- Garantizar que la evaluación utilice las teorías y enfoques de todas las disciplinas comprendidas en la evaluación (p. ej., antropología, medicina, derecho, sociología, economía, demografía), cada una usada para enriquecer y ampliar las demás.
- Asegurar que los conceptos y métodos no se saquen de contexto, sino que hagan uso de los debates intelectuales y los enfoques de las disciplinas respectivas.

Recolección de datos y uso de la triangulación:
Muchas propuestas de evaluación se refieren a la triangulación, pero no suele utilizarse de manera sistemática. La triangulación es un componente clave de un enfoque de métodos mixtos y es responsabilidad del gerente garantizar su pleno uso. El gerente deberá:

- Seleccionar métodos de recolección de datos CUANT y CUAL que se complementen entre sí, y especificar cómo se combinarán en el trabajo de campo y el análisis.
- Elegir al menos dos métodos de cálculo independientes para hipótesis e indicadores clave.
- Asegurar la documentación completa de todos los métodos de selección de muestra, recolección de datos y análisis.

Análisis de datos y posible seguimiento en el campo:
El gerente debe garantizar que se haya elaborado un plan de análisis de datos con MM desde el inicio de la evaluación. Esto incluye:

- Presentar análisis separados de hallazgos CUANT y CUAL para destacar diferentes interpretaciones y hallazgos, y preparar un informe integrado a partir de todos los datos.
- Utilizar un procedimiento de triangulación sistemática para verificar si existen contradicciones o interpretaciones diferentes.
- Presupuestar los recursos y el tiempo para realizar visitas de seguimiento en el campo.
- Destacar las diferentes interpretaciones y hallazgos de los distintos métodos, y analizar cómo estos enriquecen el estudio. Los resultados diferentes, y en apariencia contradictorios, se deben considerar como una ventaja importante del enfoque integrado, en lugar de ser un obstáculo.
- Presentar casos y material CUAL para ilustrar o comprobar los hallazgos CUANT.

Presentación y difusión de los hallazgos:
El gerente debe alentar al equipo a ampliar la variedad de métodos de presentación y difusión para garantizar que se capte la totalidad de lo que los datos de MM pueden ofrecer. Esto incluye:

- Combinar los informes escritos convencionales y las presentaciones en PowerPoint con más métodos de presentación participativa. Desarrollar más informes innovadores y fáciles de utilizar para evitar extensos informes técnicos para públicos sin conocimientos técnicos.
- Ampliar la heterogeneidad de partes interesadas invitadas a las presentaciones de hallazgos para incluir grupos de la comunidad y de la sociedad civil que no suelen ser consultados en muchas evaluaciones CUANT.

Tabla 5. Áreas adicionales de experiencia profesional en investigación que se pueden requerir en equipos con una orientación CUANT o CUAL para realizar evaluaciones con MM.

EVALUACIONES CON ORIENTACIÓN CUANT	EVALUACIONES CON ORIENTACIÓN CUAL
<ul style="list-style-type: none"> • Conocimiento de métodos de recolección de datos CUAL básicos (p. ej., entrevistas exhaustivas, técnicas de entrevistas grupales, técnicas de observación) • Conocimiento de la cultura local • Experiencia con programas de desarrollo organizacionales y comunitarios, y evaluación de estos • Uso sistemático de la triangulación 	<ul style="list-style-type: none"> • Métodos de muestreo estadístico • Administración de la recolección de datos CUANT, en particular, la gestión de cuestionarios estructurados • Análisis de datos estadísticos

de investigación incorporen cada una de estas metodologías. Esto puede ser importante para organizaciones que tradicionalmente han dependido de métodos principalmente CUANT o CUAL y donde el equipo de evaluación quizás tuvo que ser estimulado para garantizar que se otorgara igual importancia a los marcos y métodos de disciplinas recientemente introducidas.¹⁷

Garantizar el uso de MM durante la recolección de datos. La decisión acerca de utilizar una recolección de datos secuencial o concurrente es, en parte, un tema técnico, pero incluye consideraciones de administración. La recolección de datos CUANT y CUAL consecutiva puede ahorrar una importante cantidad de tiempo, lo que también puede ahorrar dinero, pero esto necesitará sistemas de gestión eficientes para coordinar equipos de recolección de datos CUAL y encuesta CUANT que pueden tener diferentes métodos de operación. Cuando la recolección de datos está restringida por dificultades logísticas y de seguridad, el uso de la recolección de datos consecutiva puede requerir de una administración eficiente del transporte para llevar y traer a los investigadores de diferentes áreas, además de una coordinación con

la policía y el ejército en situaciones donde se debe obtener permiso para visitar comunidades y preparar y aprobar itinerarios precisos.

Administración de análisis de datos con métodos mixtos. Con un diseño predominantemente CUANT, por lo general, el análisis de datos no comienza hasta que todos los datos, o su mayoría, se hayan recolectado, ingresado en la base de datos y limpiado. Sin embargo, el análisis de datos CUAL puede ser un proceso constante que comienza apenas los investigadores ingresan al campo. La administración de datos se torna más difícil porque se trata de un proceso interactivo más que lineal. Los hallazgos iniciales del análisis de datos CUAL se suelen utilizar para diseñar preguntas para encuestas CUANT, y el análisis de datos CUANT inicial se suele utilizar a fin de seleccionar las muestras para el análisis CUAL exhaustivo. Esto presenta una serie de desafíos de administración especiales: (a) la recolección de datos se debe realizar con mayor rapidez para no demorar el inicio de la siguiente etapa del diseño de evaluación y recolección de datos; (b) los procedimientos de calidad de los datos son más complejos, ya que deben evaluar la calidad tanto de la recolección de datos CUANT como de la recolección de datos CUAL y (c) los procedimientos de triangulación se deben usar sistemáticamente para verificar la coincidencia de las estimaciones de indicadores de resultados clave obtenidas a partir de diferentes métodos de recolección de datos, y debe haber procedimientos para realizar un seguimiento

¹⁷ Suele suceder que no se convoca a profesionales de disciplinas nuevas hasta que el diseño de evaluación ya se haya definido, y solo se les pedirá que recolecten datos que se ajusten a diseños de evaluación ya establecidos. La experiencia muestra que las evaluaciones orientadas CUANT suelen emplear investigadores con una orientación CUAL solo para que dirijan algunos grupos de discusión o realicen estudios de caso a fin de demostrar que la evaluación está utilizando métodos mixtos, pero sin integrarlos realmente en el diseño de evaluación general.

y determinar los motivos de las contradicciones entre datos de diferentes fuentes. Todos estos desafíos requieren sistemas de administración de datos más eficientes y complejos.

4.2. Consejos para que las ONG con recursos limitados movilicen la experiencia profesional y los recursos requeridos para realizar evaluaciones con métodos mixtos

Las evaluaciones con métodos mixtos ofrecen una gran ventaja para las ONG que necesitan informes de evaluación de calidad y creíbles pero deben realizar las evaluaciones con pocos recursos, solo con una limitada experiencia profesional de los integrantes del equipo en lo que respecta a evaluaciones o con pocos recursos para contratar a expertos externos. Para las muchas ONG que realizan evaluaciones con estas limitaciones del mundo real, con frecuencia, se necesitará creatividad para obtener la experiencia profesional requerida. Esto se puede lograr a través de capacitación adicional para el equipo de investigación actual o mediante la búsqueda de la experiencia profesional necesaria a través de la colaboración con otras agencias. No existen reglas rígidas respecto de los niveles mínimos de requisitos de experiencia profesional, tiempo y recursos para realizar una evaluación con MM metodológicamente correcta, pero idealmente un nivel mínimo de experiencia debería estar a disposición de todas las áreas identificadas en la Tabla 5. El Cuadro 6 ofrece consejos sobre cómo lograr la experiencia profesional esencial y movilizar recursos adicionales cuando se trabaja con un presupuesto limitado.

Todas estas limitaciones afectan la credibilidad de los hallazgos de la evaluación. ¿Qué tan creíble es una evaluación que se tuvo que realizar con un presupuesto ajustado, poco tiempo en el campo, acceso limitado a los datos y, posiblemente, con problemas de seguridad? Aunque no hay una solución mágica para resolver estas dificultades, existen numerosas

maneras en las que los MM pueden lograr un rigor metodológico aceptable y hallazgos creíbles. Las siguientes son algunas de las maneras útiles de lograr hallazgos creíbles al trabajar con estas limitaciones:

a. Basar la evaluación en una teoría de cambio bien articulada (tanto la GO1 como la GO2 incluyen extensos estudios sobre el desarrollo y el uso de la teoría de cambio en el análisis y el diseño de la evaluación de impacto). Una teoría de cambio que se desarrolla a través de consultas participativas a las partes interesadas puede definir pasos y procesos a través de los cuales se lograrán los rendimientos, los resultados y los impactos, además de identificar las suposiciones críticas que se evaluarán. La teoría de cambio debe incluir hitos acordados, de modo que si se debe realizar una evaluación cuando aún es demasiado pronto para medir resultados, los hitos pueden contribuir a mejorar la credibilidad de la evidencia de que el proyecto está bien encaminado hacia el logro de sus objetivos. Los MM pueden afianzar la teoría de cambio al incorporar indicadores CUANT y CUAL, estudiar lo que sucede durante la implementación del proyecto y describir procesos importantes de cambio de comportamiento.

b. Consultar a las partes interesadas para asegurar que consideren los métodos de evaluación y los indicadores clave creíbles, y para tener conocimiento de los problemas que puedan surgir y abordararlos.

c. Utilizar la triangulación para maximizar la validez de las estimaciones basadas en muestras pequeñas de diferentes fuentes (consulte la Sección 2.2).

d. Emplear el muestreo con métodos mixtos para garantizar la representatividad de los datos CUAL de estudios de caso, entrevistas exhaustivas y observaciones (consulte la Sección 3.1, Modelo 2).

CUADRO 6. REALIZACIÓN DE EVALUACIONES DE IMPACTO CON MÉTODOS MIXTOS CON POCOS RECURSOS: CONSEJOS PARA LAS ONG CON RECURSOS LIMITADOS Y POCA EXPERIENCIA PROFESIONAL EN EVALUACIONES

Aunque los métodos mixtos se pueden utilizar como parte de una evaluación de impacto compleja, costosa y rigurosa, la mayoría de las ONG tienen tiempo y recursos limitados para invertir en evaluaciones de impacto y muy poca experiencia profesional en evaluaciones. No obstante, necesitan evaluar los resultados e impactos de sus programas de una manera creíble. La buena noticia es que las herramientas y técnicas con MM son muy flexibles y pueden contribuir a mejorar la calidad de las evaluaciones de impacto para las ONG que realizan evaluaciones con pocos recursos. A continuación, se enumeran algunos consejos prácticos analizados en esta guía de orientación.

Comience gradualmente y extienda el alcance a medida que obtiene experiencia.

- Por lo general, es más simple y más económico comenzar utilizando los MM en una sola etapa de la evaluación; por ejemplo, usar al menos dos métodos independientes para recolectar datos o realizar un estudio diagnóstico rápido para diseñar un cuestionario. Aproveche los datos que ya se están generando a través del sistema de control del programa. A medida que aumenta su experiencia, considere extender el alcance.
- Habitualmente, es más fácil empezar usando diseños de evaluación de impacto secuenciales, en lugar de concurrentes (consulte la Sección 2.2).
- Aunque algunos enfoques de métodos mixtos son bastante complicados de usar, existen muchas otras técnicas que son simples y económicas, y que el personal con experiencia limitada en investigación puede aprender y usar fácilmente.
- Integre el diseño de MM con el sistema de control y evaluación (M&E) y aproveche al máximo los datos recolectados a través de estos (consulte la GO2).

Asegúrese de que los hallazgos sean metodológicamente correctos y creíbles para las partes interesadas.

- Para comenzar, concéntrese en las clases de datos que las partes interesadas consideran creíbles. No gaste todos sus recursos en encuestas de muestras extensas si sus clientes y las partes interesadas prefieren estudios de caso y descripciones exhaustivas de cómo funciona realmente el programa.
- Al mismo tiempo, trate de equilibrar las preferencias de las partes interesadas con una combinación criteriosa de datos CUANT y CUAL, y con el uso de la triangulación y otras técnicas

para aumentar la validez. Los estudios de MM recolectan una amplia variedad de datos cuantitativos y cualitativos, de modo que incluso a partir de una pequeña evaluación, es posible seleccionar las clases de evidencia que son más convincentes para las partes interesadas.

- El muestreo de MM puede asegurar que se pueda seleccionar una pequeña cantidad de estudios de caso para garantizar que sea representativa en líneas generales del total de la población objetivo, y así, aumentar la validez/credibilidad de los hallazgos (consulte la Sección 2.3).
- Al combinar la evidencia de diferentes fuentes, la triangulación puede aumentar la credibilidad de las estimaciones a partir de muestras pequeñas (consulte la Sección 2.2 y el Anexo 9).

Manténgase dentro de su presupuesto y nivel de experiencia profesional.

- Existen muchas maneras de reducir los costos de la recolección de datos y, a la vez, asegurar que los datos sean confiables (consulte el Anexo 8).
- Muchos departamentos universitarios enseñan sobre métodos mixtos y quizás sea posible colaborar con un miembro del cuerpo docente o estudiante de forma gratuita o por un costo simbólico para una de sus asignaciones de campo. Muchos estudiantes de posgrado podrían considerar la preparación de un estudio de caso como una de las asignaciones de la carrera.

Obtenga ayuda.

- Existen muchos seminarios web y otros recursos en Internet sobre el diseño y el uso de los MM (<http://www.interaction.org/resources/training> es un buen sitio para empezar).
- Muchas de las grandes ONG internacionales tienen departamentos de evaluación permanente que pueden estar dispuestos a brindar orientación gratuita por teléfono o en persona cuando visitan su país.

Ejemplos del mundo real de diseños de métodos mixtos económicos.

- El Estudio de caso n.º 7, Evaluación del Proyecto educativo de UNICEF en Timor L'Este, y el Estudio de caso n.º 11, Evaluación del Fondo de Desarrollo Comunitario de Eritrea, ilustran cómo se pueden utilizar los métodos mixtos cuando es difícil acceder a los datos y los recursos son limitados.

Parte V: Estudios de caso que ilustran las diferentes aplicaciones de los diseños de métodos mixtos

Las tres evaluaciones descritas en esta sección ilustran las metodologías utilizadas en: una evaluación predominantemente CUANT, una evaluación predominantemente CUAL y una evaluación en la que los métodos CUANT y CUAL tienen igual importancia. Juntos, estos tres ejemplos ilustran una amplia variedad de enfoques de evaluación con MM que se pueden emplear. La Tabla 6 (al final de esta sección) enumera 17 ejemplos de diseños de evaluación con MM que se resumen en el [Anexo 10](#). Las evaluaciones tratan proyectos respaldados por las ONG, agencias de las Naciones Unidas y el Banco Mundial.

Ejemplo 1: Una evaluación con orientación CUANT: Evaluación de un programa de reconstrucción postconflicto en Liberia (DFID y el Comité Internacional de Rescate)

La evaluación, finalizada en diciembre de 2008, evaluó el impacto de un programa de reconstrucción impulsado por la comunidad, financiado por el DFID e implementado por el Comité Internacional de Rescate en el norte de Liberia en 2006-7 después del conflicto. En 42 comunidades, el programa de reconstrucción impulsado por la comunidad utilizó subvenciones globales, instituciones de toma de decisión elegidas por el pueblo, proyectos de desarrollo comunitario y planificación participativa para mejorar el bienestar socioeconómico, la gestión local y la cohesión comunitaria. La evaluación examinó si el programa logró sus objetivos.

Se empleó un diseño de evaluación con métodos mixtos. La motivación para utilizar un enfoque de métodos mixtos fue el reconocimiento de las

dificultades de evaluar, únicamente sobre la base de datos de encuesta CUANT, si los cambios en las actitudes planteadas reflejaban cambios reales. Se esperaba que la exposición al proyecto pudiera enseñarles a los participantes cuáles eran las maneras esperadas de responder a las preguntas sobre actitudes hacia la cooperación y la democracia, sin afectar su capacidad real o su propensión a participar en acciones comunitarias colectivas. Se diseñó la evaluación para abordar estos desafíos mediante la combinación de datos de encuesta CUANT sobre cambios en las actitudes y el comportamiento declarado con datos experimentales de campo que pudieran medir los cambios de comportamiento reales. Tanto el método de la encuesta como el método experimental de campo se basaron en una prueba controlada aleatoria; al inicio, se eligieron comunidades al azar para que recibieran tratamientos de centros de investigación cooperativa (que desarrollan comités comunitarios y brindan apoyo para programas de impacto rápido), mientras que una cantidad casi igual de comunidades de control no recibieron tales programas. Otros métodos CUANT utilizados incluyeron un cuestionario de capital social para evaluar cómo los resultados eran afectados por el capital social existente y cómo el programa afectaba el capital social, y la recolección de datos de encuesta sobre organización social y cohesión social.

Los métodos CUAL y experimentales de campo incluyeron la recolección de datos sobre organización comunitaria y cohesión social a través de entrevistas exhaustivas y observaciones sobre el comportamiento individual en un juego de bienes

públicos. Seis meses después de terminado el programa de reconstrucción impulsado por la comunidad, se les dio a todas las comunidades de tratamiento y de control la oportunidad de recaudar fondos de una ONG liberiana para implementar un proyecto de bienes públicos. Los únicos requisitos eran que las comunidades eligieran un proyecto de antemano y nombraran a tres personas para administrar el dinero. También se les dijo a las comunidades que la cantidad de dinero que recibirían (de \$0 a, aproximadamente, \$500) dependería de los resultados de un juego de bienes públicos en todo el pueblo, en el que 24 personas elegidas al azar podían elegir privadamente conservar una suma de dinero para uso propio o aportar dicha suma a un fondo comunitario (con una contribución equivalente adicional provista por los investigadores) a fin de utilizarse para el bien público. Luego, los investigadores recolectaron los datos sobre cómo las comunidades eligieron los proyectos y a los representantes, y observaron los patrones de juego en la actividad sobre bienes públicos. Se consideró que los juegos lograron arrojar una estimación más precisa sobre las actitudes y el comportamiento ya que existía un costo monetario real en actuar para complacer a terceros.

Ejemplo 2: Una evaluación con MM de orientación CUAL: Evaluación de un programa para prevenir la violencia contra la mujer en El Salvador (Oxfam América)

El programa de 10 años para combatir la violencia contra la mujer se inició en 2008. Funciona a nivel nacional, municipal y comunitario, y combina campañas de concienciación pública con la organización de las mujeres y la sociedad civil, además de promover una coordinación más eficaz entre las autoridades públicas y las mujeres organizadas. Uno de los principales desafíos concierne a la atribución de cambios en las actitudes y el comportamiento respecto de las

estrategias de campaña y defensa nacional y municipal, en particular, entre las agencias del sector público. La cuestión de lo contrafactual es difícil. Otra dificultad fue la disponibilidad limitada de buenos datos CUANT sobre niveles y cambios en la violencia contra la mujer. Los datos policiales, judiciales y sanitarios no eran confiables ni exactos en cuanto a la cantidad de denuncias. Como siempre, las víctimas se resistían a denunciar el abuso. Por lo tanto, el equipo de evaluación decidió que la manera más precisa de medir la influencia de la campaña era a través de casos de comparación exhaustivos, orientados a objetivos intermedios y complementados, cuando fuera posible, por los datos CUANT disponibles.

El diseño de evaluación con MM se basó en una teoría de cambio. Esta identificaba cómo varios resultados promovidos en coalición podrían secuenciarse y combinarse para producir los impactos deseados: coordinación e intercambios entre agencias regionales; nuevas políticas y normas (legislación, programa de estudios escolar, políticas municipales de prevención); cambios en las relaciones sociales y el comportamiento a través de la toma de conciencia y una mejor aplicación de la ley; y bienestar y agencia (mayor conocimiento; mujeres con más confianza en sí mismas para tomar medidas; mujeres que se benefician con servicios mejorados; mujeres, jóvenes y aliados influyen sobre quienes toman decisiones). Se definió un conjunto de 11 indicadores a largo plazo para evaluar resultados e impactos en un período de 10 años, y se determinaron “objetivos de referencia” cada tres años.

La evaluación se basó en un conjunto de principios de proceso (enfoque de métodos mixtos y equipo multidisciplinario, institución de investigación regional, diseño de colaboración y consulta a todos los colaboradores, y uso de triangulación y

consulta para validar los hallazgos). Los elementos clave del enfoque metodológico incluyeron:

- Estudios de caso comparativos y exhaustivos en dos municipalidades.
- Síntesis de documentos entre partes interesadas importantes.
- Entrevistas sobre la efectividad de las elecciones de defensa a nivel nacional.
- Atención a las conexiones entre diferentes niveles de intervención (nacional, municipal y local).
- Horizontes temporales diferentes para evaluar resultados diferentes.

Los caminos de la investigación y las fuentes de datos clave incluyeron:

- Proporcionar un contexto sobre tendencias generales a través de la recolección de indicadores a nivel nacional y municipal (estadísticas gubernamentales, estadísticas sobre violencia contra la mujer, promulgación e implementación de leyes sobre violencia contra la mujer, compromisos presupuestarios).
- Dos estudios de caso comparativos y exhaustivos centrados en los resultados en municipalidades específicas, las percepciones de seguridad de las mujeres, la evaluación de la efectividad de las acciones de prevención, confianza de las mujeres en sí mismas para tratar la violencia ejercida contra ellas. Las fuentes de datos incluyeron entrevistas exhaustivas, observaciones, tendencias y estadísticas municipales, y grupos de discusión con mujeres que informaban sobre los cambios en las condiciones.
- Evaluación de defensa: trazar el camino de influencia de la campaña a través de la revisión de documentos y entrevistas con los creadores de políticas y otros actores sobre eventos clave

y puntos de decisión, y preguntar si hubiera habido avances sin las campañas de defensa. Evaluar si la campaña influyó en la versión final de la legislación y la efectividad de los mecanismos de la sociedad civil para influir sobre las políticas de prevención. Utilizar datos secundarios para comparar tendencias en las áreas donde operaron y no operaron las campañas.

El análisis y la interpretación de los hallazgos combinaron:

- Análisis de contribución en el nivel más general del programa que utilizó datos de componentes del proyecto, estudios de caso y trazado de los caminos de influencia (según el marco de la teoría de cambio). La causalidad y la atribución se pueden estudiar para proyectos de componentes de la coalición del programa, pero el foco de atención en el nivel del programa sirve para evaluar suposiciones básicas de un cuerpo de trabajo sumamente complejo.
- La teoría de cambio fue un elemento clave en la definición de lo contrafactual y proporcionó un marco para comparar los procesos y los cambios observados con el proceso de cambio previsto.

Ejemplo 3: Un diseño de MM equilibrado: Evaluación de los impactos del programa de Reforma de desarrollo comunitario de Gram Panchayats en la India (Banco Mundial)

El objetivo del programa era delegar poder a los gobiernos elegidos democráticamente en los pueblos y promover una mayor participación de las mujeres y las castas con reconocimiento expreso en la Constitución de la India. También intentaba mejorar la efectividad del uso de fondos a través de una mayor participación de la comunidad. El gobierno otorgó subvenciones que

serían administradas por la comunidad local para la implementación de proyectos económicos y de infraestructura social elegidos por la comunidad. El programa en evaluación era un programa de capacitación de dos semanas para ciudadanos sobre toma de decisión y planificación participativa, diseñado para mejorar la eficiencia con la que utilizarían las subvenciones. El diseño del programa incluyó la asignación al azar de las comunidades a los grupos de tratamiento y de control (que recibían subvenciones, pero no capacitación).

La evaluación utilizó un diseño equilibrado de métodos mixtos que comprendió los siguientes pasos:

- Selección de 200 pueblos (Gram Panchayats) con asignaciones al azar a grupos del proyecto y de control (CUANT).
- Investigación exploratoria sobre tenencia de tierras, propiedad de bienes públicos, participación y redes sociales (QUAL).
- Encuesta de base anterior a los programas de capacitación (CUANT).
- Análisis de procesos exhaustivo en cinco áreas del proyecto y cinco áreas de control (QUAL) para observar los cambios en cómo se organizan las comunidades, planifican y administran proyectos, y los efectos en la participación y la integración de las mujeres y las castas con reconocimiento constitucional. Esto implicó visitas de uno o dos días cada semana durante un período de dos años. Las visitas combinaron observaciones en el nivel comunitario, entrevistas con políticos y líderes de la comunidad, y entrevistas exhaustivas a 20 familias de cada pueblo. Las observaciones abarcaron: dinámica política y social, corrupción, cambio económico y afiliación a redes.
- Repetición de la encuesta de base después de dos años (CUANT).

- Se utilizó la triangulación en todo momento para comparar las estimaciones CUANT y QUAL de los cambios e impactos.

Algunas de las ventajas del enfoque de MM incluyeron:

- El recordar los procesos y debates en las reuniones de los pueblos no resulta fidedigno; por eso, asistían observadores a las reuniones y transcribían las actas y la dinámica de las reuniones.
- Estudios previos han hallado que las encuestas producen estimaciones muy variables de inequidad (la variación es demasiado grande para poder usar los hallazgos en análisis estadísticos), de modo que se emplearon técnicas de evaluación rural participativa, en las que se seleccionaron participantes que fueran representativos de la población total, con el fin de complementar los datos de la encuesta.
- El desarrollo comunitario, la participación, la corrupción y otros procesos clave son difíciles de captar en las encuestas; por lo tanto, los datos de la encuesta se compararon con las observaciones, informantes clave y entrevistas grupales participativas.
- Al mismo tiempo, fue esencial obtener estimaciones cuantitativas de la relación costo-efectividad de los programas de capacitación, ya que este era un programa piloto para informar decisiones respecto de si el programa se debería reproducir en otros estados.¹⁸ Por consiguiente, las estrategias de muestreo aseguraron que los estudios de caso cualitativos y otros datos se pudieran vincular a la encuesta de muestra CUANT para garantizar que los hallazgos de la evaluación se pudieran generalizar.

¹⁸ La evaluación incluyó una comparación con otros cuatro estados (no incluidos en este resumen).

Tabla 6. Ejemplos de evaluaciones con MM resumidas en el [Anexo 10](#).

Título	Agencias que apoyan el programa y realizan o encargan la evaluación
A. Diseños de evaluación con orientación cuantitativa	
1. Reconstrucción posconflicto en Liberia.	DFID y Comité Internacional de Rescate
2. Evaluación a largo plazo del programa de Tostan para reducir la circuncisión femenina en aldeas de Senegal.	UNICEF
3. Evaluación de un programa de transferencia de efectivo condicional en Kazajstán.	Save the Children
4. Evaluación de impacto del trabajo de rehabilitación y emergencia de la FAO en el desarrollo de comunidades rurales.	FAO
B. Diseños de evaluación con orientación cualitativa	
5. Evaluación de los impactos de un programa de prevención de la violencia de género en El Salvador.	Oxfam América
6. Programa de seguridad alimentaria para la vida y la subsistencia en Bangladesh.	USAID, Save the Children y TANGO International
7. Evaluación del Programa de educación de UNICEF en Timor-L'Este.	UNICEF
8. Evaluación de resultados de capital del proyecto Educación para todos en Nepal.	UNICEF, NORAD y otros colaboradores
9. Evaluación de resultados de capital del proyecto de Sanidad total impulsado por la comunidad en Camboya.	UNICEF
10. Evaluación entre agencias en tiempo real de la respuesta humanitaria a la crisis de desplazamiento de 2009 en Pakistán.	UNICEF y diversos colaboradores
11. Evaluación del componente de construcción de carreteras del Fondo de Desarrollo Comunitario de Eritrea.	Banco Mundial
12. Evaluación del proyecto de Escuelas comunitarias de Egipto.	UNICEF
13. Evaluación del proyecto de Facilitación judicial comunitaria de Tanzania.	UNICEF
14. Evaluación de la Respuesta de UNICEF, en el área de Protección infantil en Indonesia, al Tsunami del Océano Índico de 2004.	UNICEF
C. Diseño de evaluación equilibrado que otorga igual importancia a enfoques CUANT y CUAL	
15. Evaluación del Proyecto de desarrollo a nivel de Kekamatan en Indonesia.	Banco Mundial y Gobierno de Indonesia
16. Evaluación del Programa de reforma de Panchayats en la India.	Banco Mundial
D. Meta-análisis (análisis secundario de una cantidad de evaluaciones para identificar hallazgos generales)	
17. CARE International. Informe de evaluación de impacto 2005-2010. Latinoamérica y el Caribe.	CARE International

Referencias

- Adato, M (2012) "Combining quantitative and qualitative methods for program monitoring and evaluation: Why are mixed methods designs best?" in Lopez-Acevedo, G; Krause, P & Mackay, K (editors) Building better policies: The nuts and bolts of monitoring and evaluation systems. World Bank.
- Bamberger, M. (2010). *Reconstructing baseline data for impact evaluation and results measurement. No. 4. The Nuts and Bolts of M&E Series. Poverty Reduction and Equity Department. The World Bank.* <http://site-resources.worldbank.org/INTPOVERTY/Resources/335642-1276521901256/premnote-ME4.pdf>.
- Bamberger, M., Rao, V., & Woolcock, M. (2010). *Using mixed-methods in monitoring and evaluation: Experiences from international development.* In Tashakkori A. and Teddlie C. Sage Handbook of Mixed- Methods in Social and Behavioral Research. Sage Publications. pp 613-642.
- Bamberger, M., Rugh, J. & Mabry, L. (2012). *RealWorld Evaluation Second Edition.* Sage Publications.
- Chen, H. (2005) *Practical Program Evaluation: Assessing and improving planning, implementation, and effectiveness.* Sage Publications.
- Donaldson, S., Christie, C. & Mark, M. (2009). *What counts as credible evidence in applied research and evaluation practice?* Sage Publications.
- Funnell, S. & Rogers, P. (2011). *Purposive programme theory: effective use of theories of change and logic models.* Jossey-Bass Publications.
- Mayne, J. (2008) "Contribution analysis: An approach to exploring cause and effect." Rome: Institutional learning and change initiative , ILAC Brief No. 16. May 2008. http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf
- Organization for Economic Cooperation and Development. 2010 (Second Edition). *Evaluating Development Cooperation: Summary of Key Norms and Standards.* Paris: OECD. <http://www.oecd.org/dataoecd/12/56/41612905.pdf>
- Patton, M.P. (2011). *Developmental Evaluation: applying complexity concepts to enhance innovation and use.* Guilford Press.
- Proudlock, K & Ramalingam, B (2009) *Improving humanitarian impact assessment: Bridging theory and practice.* ALNAP
- Rieper, O Leeuw, F & Ling, T (2010). *The evidence book: concepts, generation and use of evidence.* Transaction Publishers.
- Salmen, L. (1987). *Listen to the People: Evaluation of Development Projects.* New York. Oxford University Press.
- Stern, E; Stame, N; Mayne, J; Forss, K; Davis, R & Befani, B (2012) *Broadening the range of designs and methods for impact evaluation.* DFID

UNICEF (2009a). Children and the 2004 Indian Ocean Tsunami: Evaluation of UNICEF's response in Indonesia, Sri Lanka and Maldives, 2005-08. Overall synthesis report.

Weiss 2001. Theory based evaluation: Theories of change for poverty reduction programs pp. 103-14 in Evaluation and Poverty Reduction edited by O. Feinstein and R. Piccitto. New Brunswick. Transaction

Williams B. (2005). Systems and Systems Thinking in Mathison, S (editor) Sage Encyclopedia of Evaluation pp. 405-412.

Williams, B. & Imam I. (Eds.), (2007). Systems Concepts in Evaluation: An expert anthology. American Evaluation Association.