



MÓDULO DE EVALUACIÓN DE PLANES, PROGRAMAS Y PROYECTOS DE DESARROLLO



DECISIONES BASADAS EN EVIDENCIA

MARZO, 2018

Diplomado en Monitoreo y Evaluación de Planes, Programas y Proyectos de Desarrollo

Manual de Autoaprendizaje del Módulo de Evaluación de Planes, Programas y Proyectos de Desarrollo

Marzo 2018

PROYECTO EVALUATIONS DE USAID

Teléfono: (51-1) 2258110

Av. del Parque Norte 1160 - Oficina 702 San Borja, Lima - Perú

<http://www.pgrd.org/projects/peru-evaluations/publications/>

Elaboración de contenidos

Alejandro Bardales Hoyos

Luis Soberón Álvarez

Alberto Padilla Trejo

Diseño

Luz Evelyn Remolino Rojas

Reproducido por

Cristhian Arturo Ramírez Carbajal

Dirección: Av. Bolivia N° 148 Int. 3071 – Lima

Lima, publicado y reproducido en marzo 2018

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2018-03372

La información contenida puede ser reproducida total o parcialmente, siempre que se mencione la fuente de origen.

Este manual ha sido posible gracias al apoyo del pueblo de los Estados Unidos de América a través de la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID). Fue elaborada por Partners for Global Research and Development (PGRD) en el marco del proyecto Evaluations.

El contenido no necesariamente refleja la opinión de USAID ni la del Gobierno de los Estados Unidos.

TABLA DE CONTENIDO

ACRÓNIMOS	7
INTRODUCCIÓN	8
Información básica.....	10
A. Secuencia de estudio del módulo	10
B. Requisitos del módulo	11
C. Competencias que lograrás con el módulo	11
D. Mapa de contenidos del manual	12
E. Productos del módulo.....	13
F. Tiempo estimado de estudio	13
SECCIÓN 1	16
Reflexionando mi práctica sobre el diseño de evaluaciones	18
UNIDAD DE APRENDIZAJE 1: CONCEPTOS BÁSICOS SOBRE EVALUACIÓN	19
1.1 Bases conceptuales y metodológicas de la evaluación	19
1.1.1 ¿Qué es evaluación?	19
1.1.2 ¿Cuáles son las funciones o propósitos de la evaluación?	20
1.1.3 ¿En qué momento de la vida del plan, programa o proyecto se evalúa?	21
1.2 Principios éticos y enfoques de equidad, género e interculturalidad en la evaluación	22
1.2.1 ¿Bajo qué principios éticos se debe evaluar?	22
1.2.2 ¿Existen evaluaciones que se guían por o sirven directamente a principios universales?	23
1.3 Niveles de la evaluación: políticas, programas y proyectos	26
1.4 Identificación de la "evaluabilidad" de un plan, programa o proyecto	29
1.5 Enfoques sobre evaluación	30
1.6 Tipos de evaluaciones	32
1.6.1 Según la etapa del plan, programa o proyecto	32
1.6.2 Según los resultados del plan, programa o proyecto	34
1.6.3 Tipos de evaluación del Presupuesto por Resultados (PpR) considerados en el Ministerio de Economía y Finanzas (MEF) del Perú	36
1.6.4 Tipos de evaluación considerados en el Ministerio de Desarrollo e Inclusión Social (MIDIS) del Perú	36

1.6.5	Sobreposición entre los tipos según etapa y resultados.....	37
1.7	Validez y confiabilidad del diseño de evaluación.....	38
1.8	El proceso de diseño de una evaluación	39
ACTIVIDADES APLICATIVAS DE LA UNIDAD 1		41
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1		42
UNIDAD DE APRENDIZAJE 2: DISEÑO DE EVALUACIONES		44
2.1	Diagnóstico del contexto	44
2.1.1	Análisis del entorno	44
2.1.2	¿Cómo hacer un análisis del entorno?	44
2.2	Descripción del plan, programa o proyecto.....	49
2.2.1	¿Cómo realizarlo?	49
2.2.2	Herramientas	49
2.3	Explicitación de la cadena de cambios: lógica de la intervención	50
2.3.1	¿Qué es una Teoría de Cambio?	52
2.3.2	De la Teoría de Cambio a la cadena de cambio o de resultados ..	53
2.4	Identificación de los interesados	58
2.4.1	¿Quiénes son los interesados?	58
2.4.2	¿Cómo realizarlo?	58
2.5	Estudio y verificación de la evaluabilidad del PPP	62
2.6	Definición de las preguntas de evaluación e hipótesis	66
2.6.1	Preguntas guía según dimensiones de evaluación.....	67
2.6.2	Énfasis de las preguntas según tipos de evaluación	71
2.7	Elección de métodos: cuantitativos, cualitativos y mixtos.....	75
2.7.1	Las metodologías cuantitativas	75
2.7.2	Recopilación de información estadística:.....	76
2.7.3	Las metodologías cualitativas	77
2.7.4	Uso de los Métodos cuantitativos en la Evaluación de impacto....	79
2.7.5	Métodos cualitativos	93
2.7.6	Métodos mixtos.....	99
2.8	Identificación de variables e indicadores a medir/verificar y las fuentes de información.....	99
2.8.1	Las variables.....	100
2.8.2	¿Qué es un indicador?	100
2.9	Determinación de la muestra	103
2.9.1	Muestreo cuantitativo	103
2.9.2	Muestreo cualitativo.....	112
2.10	Diseño de los instrumentos de recojo de información	114

2.10.1 Instrumentos cuantitativos	114
2.10.2 Instrumentos cualitativos	119
2.11 Presupuesto y cronograma	122
ACTIVIDADES APLICATIVAS DE LA UNIDAD 2	124
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 2	134
SECCIÓN 2.....	140
Reflexionando mi práctica sobre la implementación de evaluaciones	143
UNIDAD DE APRENDIZAJE 1: PAUTAS PARA EL DESARROLLO DE UNA PRUEBA PILOTO	144
1.1. Desarrollo de la prueba piloto.....	144
1.2. Validación de los instrumentos	147
ACTIVIDADES APLICATIVAS DE LA UNIDAD 1	152
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1	155
UNIDAD DE APRENDIZAJE 2: PROCESO DE RECOLECCIÓN DE LA INFORMACIÓN	158
2.1 Plan de recolección de datos.....	159
2.2 Capacitación del personal	161
2.3 Aplicación de técnicas de recolección de datos	162
2.3.1 Aplicación de instrumentos cuantitativos	162
2.3.2 Aplicación de instrumentos cualitativos.....	168
2.4 Control de calidad en la obtención de datos	172
2.5 Criterios para la supervisión de la calidad.....	173
ACTIVIDADES APLICATIVAS DE LA UNIDAD 2	175
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 2	176
UNIDAD DE APRENDIZAJE 3: PROCESAMIENTO DE LA INFORMACIÓN Y CONSISTENCIA DE DATOS.....	180
3.1. Procesamiento de datos cuantitativos	180
3.1.1. Selección del programa.....	181
3.1.2. Diseño de base de datos.....	181
3.1.3. Digitación/captura de datos	183
3.2. Procesamiento de datos cualitativos.....	183
3.2.1. Descubrimiento	184

3.2.2.	Codificación	185
3.2.3.	Programas más utilizados en el análisis cualitativo	187
3.2.4.	Relativización de los datos	188
3.3.	Revisión y consistencia de datos de datos cualitativos y cuantitativos	190
3.3.1.	No respuesta o rechazo	190
3.3.2.	No respuesta o rechazo	191
3.3.3.	Evaluar errores de digitación	192
3.3.4.	Consistencia de la información	192
3.3.5.	Precisión estadística de los estimadores	196
ACTIVIDADES APLICATIVAS DE LA UNIDAD 3		198
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 3		199
UNIDAD DE APRENDIZAJE 4: CATEGORIZACIÓN DE LA INFORMACIÓN E IDENTIFICACIÓN DE HALLAZGOS SIGNIFICATIVOS.....		202
4.1.	Categorización de la información	203
4.1.1.	Análisis cuantitativo.....	203
4.1.2.	Análisis cualitativo.....	228
4.2.	Presentación de datos	231
4.3.	Hallazgos con evidencias, triangulando fuentes y técnicas	232
4.4.	Evaluación de síntesis con interpretaciones.....	233
ACTIVIDADES APLICATIVAS DE LA UNIDAD 4		236
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 4		237
UNIDAD DE APRENDIZAJE 5: PREPARACIÓN Y REDACCIÓN DEL INFORME DE EVALUACIÓN		239
5.1	Cómo preparar y presentar un informe de evaluación	239
5.2	Cómo identificar hallazgos en la evaluación.....	244
5.3	Cómo emitir juicios de valor o mérito en una evaluación	245
5.3.1	Seleccionar un conjunto de criterios (eficacia, eficiencia, etc.) 245	
5.3.2	Establecimiento de estándares para cada uno de los criterios seleccionados	246
5.3.3	Medición del desempeño del programa	249
5.3.4	Síntesis de hallazgos	249
5.4	Cómo informar la atribución de causalidades.....	250
5.5	Cómo elaborar conclusiones	251

5.6	Cómo formular recomendaciones	252
5.6.1	Cuándo se debe formular recomendaciones	252
5.6.2	Cómo son las recomendaciones efectivas	253
5.6.3	Cómo hacer recomendaciones	255
5.7	Cómo elaborar lecciones aprendidas	258
ACTIVIDADES APLICATIVAS DE LA UNIDAD 5		259
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 5		260
UNIDAD DE APRENDIZAJE 6: ¿COMÓ EVALUAR EL INFORME FINAL DE UNA EVALUACIÓN?		262
6.1	Criterios para evaluar el informe de una evaluación	262
ACTIVIDADES APLICATIVAS DE LA UNIDAD 6		265
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 6		270
SECCIÓN 3		273
Reflexionando mi práctica sobre la elaboración de Términos de Referencia y planes de evaluación		274
UNIDAD DE APRENDIZAJE 1: ¿CÓMO ELABORAR TÉRMINOS DE REFERENCIA PARA UNA EVALUACIÓN?		275
1.1	¿Por qué es necesario elaborar Términos de Referencia (TDR)?	275
1.2	¿Cuál es el proceso de elaboración de los TDR para una evaluación?	275
1.2.1	Identificar las necesidades de la contratación:	275
1.2.2	Identificar el diseño básico de la evaluación:	276
1.2.3	El análisis de mercado:	277
1.3	Contenidos de los Términos de Referencia para una evaluación	278
1.4	Diseño de la supervisión y monitoreo de una evaluación	289
ACTIVIDADES APLICATIVAS DE LA UNIDAD 2		291
ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1		292
BIBLIOGRAFÍA DEL MÓDULO		293
ANEXOS		298
Anexo 1: Respuestas de preguntas de autoevaluación		298
Anexo 2: Glosario de términos unificado de la OECD & Normatividad vigente del sector público del Perú		308

ACRÓNIMOS

CERLALC	Centro Regional para el Fomento del Libro en América Latina y el Caribe
CMR	Mecanismos, Contexto y Resultado
CONEVAL Social	Consejo Nacional de Evaluación de la Política de Desarrollo Social
DD.HH.	Derechos Humanos
EPS	Entidades Prestadoras de Salud
EuropeAid	Cooperación y Desarrollo de la Comisión Europea
INEI	Instituto Nacional de Estadística e Informática
MIDIS	Ministerio de Desarrollo e Inclusión Social del Perú
MVCS	Ministerio de Vivienda, Construcción y Saneamiento
OECD	Organización de Cooperación y Desarrollo Económico
ONU	Organización de las Naciones Unidas
PAV	Plataforma de Aprendizaje Virtual
PNSR	Programa Nacional de Saneamiento Rural
PNUD	Programa de las Naciones Unidas para el Desarrollo
PPP	Planes, Programas y Proyectos de Desarrollo
SUNASS	Superintendencia Nacional de Servicios de Saneamiento
UNEG	United Nations Evaluation Group
UNICEF	Fondo de las Naciones Unidas para la Infancia
USAID	United States Agency for International Development

INTRODUCCIÓN

Este manual tiene como finalidad que los y las participantes del módulo logren o mejoren un conjunto de competencias para diseñar y valorar evaluaciones de planes, programas y proyectos (PPP). El manual emplea el enfoque educativo basado en lograr competencias, combina la exposición de conceptos con la especificación de procedimientos para lograr productos típicos de la evaluación: diseños, planes, Términos de Referencia e informes. Además, se incorporan actividades aplicativas y ejercicios para la autoevaluación en cada una de las unidades de aprendizaje.

El manual está hecho para "practicantes" de la evaluación, a quienes suministra información para conducir evaluaciones o para diseñarlas y supervisarlas cuando estas son ejecutadas por terceros. Trata de dar un enfoque integral de la evaluación desde el diseño hasta el uso de sus resultados. Desde otro punto de vista, es pragmático en cuanto trata de proporcionar métodos y técnicas cuantitativas y cualitativas. Sin embargo, es evidente que aún falta incorporar enfoques y métodos de evaluación provenientes de paradigmas como el constructivismo.

Se proporciona información detallada para diseñar, implementar, informar y hacer útil una evaluación, pero además se sugiere material de lectura complementario para profundizar en determinados aspectos que debe conocer un evaluador, como incorporar el género en una evaluación, por ejemplo.

El manual se estructura en secciones y unidades. La primera sección se refiere al diseño de una evaluación, la segunda da las pautas básicas para su implementación y la tercera está dedicada a cómo hacer un plan de evaluaciones y elaborar Términos de Referencia.

Cada sección se compone de unidades de aprendizaje, cada una con actividades, ejercicios y lecturas complementarias para reforzar los conocimientos.

La actividad aplicativa principal consiste en diseñar la evaluación de un PPP, seleccionado entre los que están en ejecución o recientemente terminados en la institución de procedencia del participante, y que no hayan sido previamente evaluados. El participante deberá especificar su contribución o valor agregado a lo ya avanzado en su institución.

Complementariamente se incluyen actividades aplicativas sobre la evaluación de un informe final de evaluación y referidas a la elaboración de términos de referencia para la evaluación de un PPP.

El equipo docente espera que logres y mejores las competencias previstas y que puedas aplicarlas en la mejora de los procesos de evaluación de tu institución.

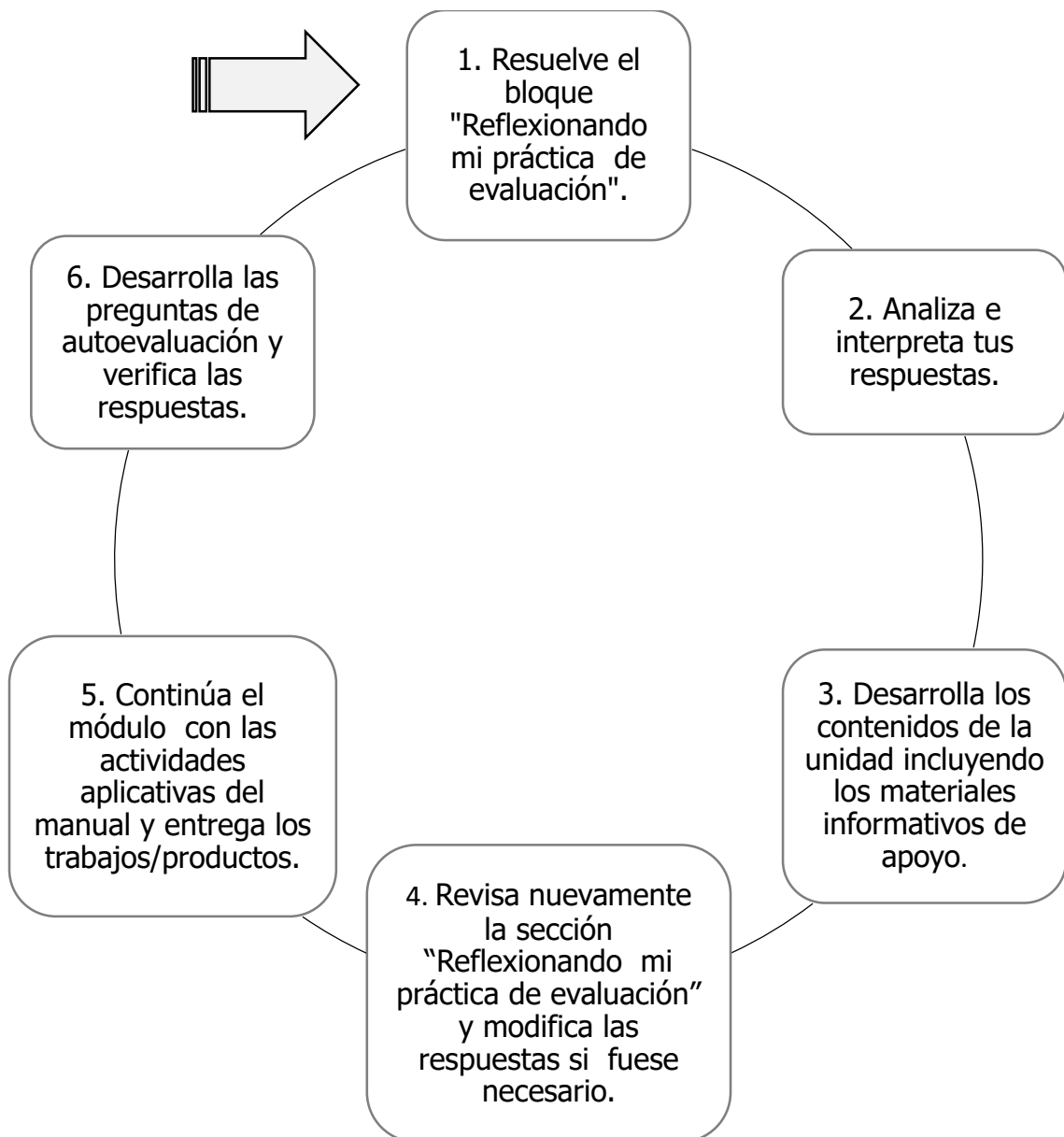
El equipo docente

Información básica

En esta parte encontrarás la información relacionada con la secuencia de estudio del módulo, los requisitos, las competencias que lograrás, el mapa de contenidos, el tiempo estimado de estudio y la duración.

A. Secuencia de estudio del módulo

A continuación te presentamos la secuencia que debes seguir para el desarrollo del módulo:



B. Requisitos del módulo

Para participar en el módulo es necesario:

- ✓ Manejar información básica de planificación (planes, programas y proyectos de desarrollo).
- ✓ Preferentemente cumplir funciones de evaluación.
- ✓ Estar motivado para cumplir funciones de evaluación.

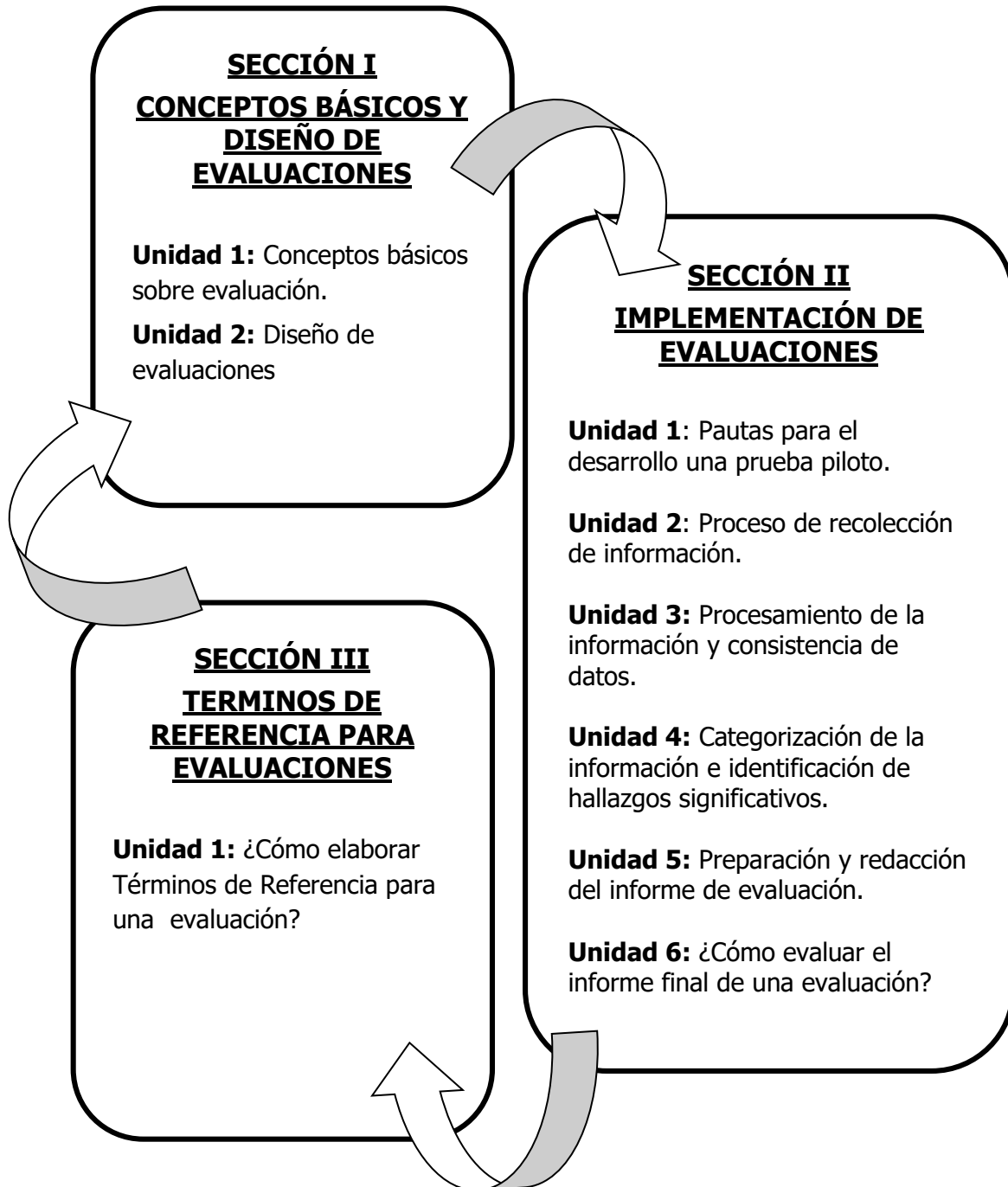
C. Competencias que lograrás con el módulo

Los participantes que cumplen con las actividades del módulo, logran:

Elaborar el diseño de evaluaciones según tipo y técnicas de evaluación, con énfasis en evaluaciones de desempeño.
Seleccionar técnicas y elaborar instrumentos de recolección de datos considerando con estándares de control y calidad.
Elaborar instrumentos de recojo de datos según el diseño de la evaluación.
Identificar los procedimientos y herramientas para procesar los datos, verificando su calidad.
Identificar los criterios para categorizar la información de la evaluación de acuerdo a las preguntas, objetivos y tipo de evaluación.
Aplicar criterios para la identificación de hallazgos significativos en función a la pregunta y objetivos de la evaluación.
Redactar las conclusiones aplicando criterios técnicos.
Formular recomendaciones de acuerdo al contexto institucional, a los objetivos de evaluación e incorporando la perspectiva de los usuarios.
Formular lecciones aprendidas de acuerdo a los resultados de la evaluación.
Evaluar el informe final de una evaluación aplicando estándares.
Elaborar Términos de Referencia acordes al diseño de evaluación requerida.

D. Mapa de contenidos del manual

El módulo está organizado en tres secciones:



E. Productos del módulo

Diseño de una evaluación (hasta la formulación de preguntas)

Instrumento de recolección de información

Informe del análisis del informe final de una evaluación

Diseño de una evaluación (versión completa final) con los términos de referencia

F. Tiempo estimado de estudio

El módulo comprende 156 horas distribuidas de la siguiente manera.

Módulo	Créditos	Distribución de horas			
		Presenciales	Virtuales	No Presenciales	Total
Evaluación de planes, programas y proyectos de desarrollo	7.6	48	36	72	156

Se desarrolla en doce (12) semanas, con actividades presenciales y no presenciales, de acuerdo al siguiente calendario:

Semanas	I		II	III	IV	V	VI	
Horas	16	3	10	10	10	10	10	
Modalidad*	P	NP	NP	NP	NP	NP	NP	
Temas	Rev***	Sec. I Unid 1	Sec. I Unid 2					
Semanas	VII		VIII	IX	X	XI	XII	
Horas	16	9	9	9	10	10	8	16
Modalidad*	P	NP	NP	NP	NP	NP	NP	P
Temas	Rev***	Sec. I Unid 2	Sec. II Unid 1 - 3	Sec. II Unid 4 -5	Sec. II Unid 6	Sec. II Unid 6	Sec. III Unid 1	Rev* **

P* : Presencial

NP** : No presencial (abarca actividades virtuales)

Rev*** : Revisión

SECCIÓN 1

SECCIÓN 1

CONCEPTOS BÁSICOS Y DISEÑO DE EVALUACIONES

CAPACIDADES

Diseñar una evaluación

Elaborar instrumentos de recojo de datos según el diseño de la evaluación

PRODUCTOS

Diseño de una evaluación

Instrumentos de recolección de información

CONTENIDOS

UNIDAD 1: Conceptos básicos sobre evaluación

UNIDAD 2: Diseño de evaluaciones

MATERIALES INFORMATIVOS DE APOYO

UNIDAD 1:

- Metodología de evaluación de la ayuda exterior de la Comisión Europea, Volumen 1, 2006. p.11-22.
- Integración de los derechos humanos y la igualdad de género en la evaluación – Hacia una Guía del UNEG. 2011. p. 13–16.
- Independent Evaluation Group. World Bank Group Impact Evaluations, Relevance and Effectiveness. The World Bank Group. 2013.
<https://openknowledge.worldbank.org/bitstream/handle/10986/13100/757230PUB0EPI00013000Pubdate0209013.pdf?sequence=1>
- Gertler, P. y otros. La Evaluación de Impacto en la Práctica. Washington DC: The International Bank for Reconstruction and Development. 2011. p.1-47.
- Lineamientos para el Seguimiento, Evaluación y Gestión de la Evidencia del MIDIS. 2012.

SECCIÓN 1

CONCEPTOS BÁSICOS Y DISEÑO DE EVALUACIONES

MATERIALES INFORMATIVOS DE APOYO

UNIDAD 2:

- Logic Model – Kellogg Foundation. 2004. p. 1-26
- ZIGLA Consultores. Inversión social estratégica. Teoría de cambio. Ejemplo de aplicación. Sin fecha. Disponible en:
<http://www.mapa.ziglaconsultores.com/content/mapa-de-monitoreo-y-evaluacion-de-impacto-social>
- Retolaza, I. Teoría de Cambio. Un enfoque de pensamiento-acción para navegar en la complejidad de los procesos de cambio social. PNUD. Hivos. Guatemala. 2010. Disponible en:
<https://planificacionsocialunsj.files.wordpress.com/2011/09/teor3ada-del-cambio-pensamiento-accion-complejidad-cambio-social.pdf>
- La evaluación de impacto en la práctica – Banco Mundial. Primera parte: Capítulo 2. p. 22-30.
- La evaluación de impacto en la práctica – Banco Mundial. Segunda parte: Capítulos 3 y 4. 2011.
- Taylor & Bodgan. Introducción a los métodos cualitativos de investigación. Ed, Paidós. Capítulo 1: Introducción ir hacia la gente. 2000.
- Módulo 2: Propiedades y Construcción de Indicadores del Curso Indicadores para el desarrollo – INDES. p. 1-53.
- García. T. El cuestionario como instrumento de investigación / evaluación. Sin editor. Almendralejo. 2003. Disponible en:
http://cvonline.uaeh.edu.mx/Cursos/Maestria/MTE/Gen02/seminario_de_tesis/Unidad_4_anterior/Lect_El_Cuestionario.pdf
- FAO. Realización de encuestas nutricionales en pequeña escala. Manual de campo. Sin fecha. Disponible en:
<https://books.google.com.pe/books?id=Tb-kRdZrgIAC&pg=PA18&lpg=PA18&dq=planificar+encuestas&source=bl&ots=buYwh-SVrf&sig=qfaNfdo8bERDruwx8DTwi8HfO9A&hl=es&sa=X&ei=rGbRVMe0J5DCsATdjoLAAQ&ved=0CEMQ6AEwBzgK#v=onepage&q=planificar%20encuestas&f=false>
p. 93-108.

Reflexionando mi práctica sobre el diseño de evaluaciones

TE INVITAMOS A REVISAR Y ANALIZAR LO SIGUIENTE:

En 1625, Gustavus II Adolphus, el rey de Suecia, encargó la construcción de cuatro barcos de guerra para avanzar en sus metas imperiales. El más ambicioso de esos navíos, llamado Vasa, fue uno de los más grandes barcos de guerra de su tiempo. El 10 de agosto de 1628, el Vasa, resplandeciente en su carpintería pintada luminosamente y dorada, fue inaugurado en el puerto de Estocolmo con la aclamación de una multitud y una notable ceremonia. Pero la aclamación tuvo corta vida; alcanzado por una ráfaga de viento mientras todavía estaba en el puerto, el barco repentinamente se ladeó, se fue a pique y se sumergió.

Se ordenó inmediatamente una investigación, y se puso de manifiesto que el compartimento del lastre no había sido hecho suficientemente grande para equilibrar los dos pisos de artillería que el rey había especificado. Con solo 121 toneladas de lastre de piedra, el barco perdió estabilidad. Sin embargo, si los constructores simplemente hubieran agregado más lastre, el piso de artillería inferior hubiera sido llevado peligrosamente cerca del agua; el barco necesitaba flotación para contener tanto peso.

En términos abstractos, el diseño del Vasa –los modos por los cuales fueron planeados y construidos los diferentes componentes del barco– fue fatalmente defectuoso. El barco fue cuidadosamente construido, reuniendo todas las normas existentes, pero las características clave de sus diferentes partes –en particular el peso de los pisos de artillería y lastre y el tamaño de la bodega– no fueron compatibles y la interacción de estas hizo que el barco se dé vuelta. Los constructores de barcos de aquella época no tenían una teoría general de diseño de barcos; trabajaban principalmente con los modelos tradicionales, por prueba y error. Aparentemente el Vasa fue originalmente proyectado como un barco más pequeño y luego fue aumentado, ante la insistencia del rey... (Maxwell, Joseph A.; 1996).

Luego de la reflexión, RESPONDE:

- En esta experiencia, ¿por qué era importante un buen diseño?
- ¿Cuál es el riesgo de un diseño imperfecto?
- ¿Qué ilustra la tragedia de Vasa con respecto al "diseño"?

Si culminaste las actividades de reflexión, desarrolla los contenidos y actividades de las unidades. Si necesitas apoyo de los tutores, comunícate con ellos a través de la plataforma virtual.



UNIDAD DE APRENDIZAJE 1: CONCEPTOS BÁSICOS SOBRE EVALUACIÓN

Esta unidad presenta una breve descripción de los principales conceptos relacionados a la evaluación. Partiendo de una definición genérica se avanza hacia conceptos prácticos sobre lo que significa e implica realizar procesos de evaluación.

1.1 Bases conceptuales y metodológicas de la evaluación

1.1.1 ¿Qué es evaluación?

La evaluación es un procedimiento que intenta valorar una situación o una acción, es decir, comprende la medición para estimar el valor y la opción de reconocer, dar importancia o apreciar ciertas cualidades de ella (lo que no necesariamente se consigue midiendo). Existen múltiples definiciones de la evaluación. Algunos autores como Weiss (1998) usan como término la investigación evaluativa (evaluation research), que la define como el uso de la metodología propia de las ciencias sociales para analizar los efectos o las operaciones de un programa. Rossi y Freeman (1989) definen la investigación evaluativa como "la aplicación sistemática de los procedimientos de investigación social para valorar la conceptualización y el diseño, la ejecución y la utilidad de los programas de intervención social".

La Organización de Cooperación y Desarrollo Económico (OECD) (2010:20) define la evaluación como la "apreciación sistemática y objetiva de un proyecto, programa o política en curso o concluido, de su diseño, su puesta en práctica y sus resultados". La definición de la OECD usa el término "apreciar", cuyo significado es darle valor a algo o alguien y también estimar la magnitud de las cosas o sus cualidades. Agrega luego que tiene como objetivo "determinar la pertinencia y el logro de los objetivos, así como la eficiencia, la eficacia, el impacto y la sostenibilidad para el desarrollo. Una evaluación deberá proporcionar información creíble y útil que permita incorporar las enseñanzas aprendidas en el proceso de toma de decisiones de beneficiarios y donantes".

La evaluación también se refiere al proceso de determinar el valor o la significación de una actividad, política o programa.

1.1.2 ¿Cuáles son las funciones o propósitos de la evaluación?

Seguiremos aquí el planteamiento de EuropeAid,¹ que asigna tres funciones o propósitos a la evaluación:

► **Ayuda a la decisión**

Se puede conducir la evaluación en beneficio de las personas que van a decidir o negociar un cambio en la intervención evaluada. En este caso se la utiliza para corregir la puesta en marcha, para modificar la estrategia de intervención o para redefinir las orientaciones políticas.

El principal mecanismo de utilización de las evaluaciones por parte de los responsables de tomar las decisiones se denomina retroalimentación (*feedback*). Las recomendaciones son esenciales para este tipo de uso así como lo es el seguimiento de las mismas.

► **Ayuda al juicio e informa**

La evaluación puede estar destinada a una gama más amplia de partes interesadas con miras a ayudarlas a formar su propio juicio sobre los méritos. Además de los responsables de la toma de decisiones, se puede extender este uso a las instituciones políticas (Ej.: el Parlamento), a los países asociados, a la sociedad civil, a los medios y, en última instancia, a los ciudadanos con la perspectiva de informar a la opinión pública.

El principal mecanismo mediante el cual se utilizan las evaluaciones para formular juicios es la rendición de cuentas (*accountability*). Las conclusiones son esenciales para este tipo de uso, así como la transparencia de los criterios de juicio (también denominados "criterios de evaluación razonada").

► **Conocer y comprender**

Por último, la evaluación puede ayudar a los usuarios a aprender de la intervención, a comprender mejor lo que funciona y lo que no, y a acumular conocimientos. Contribuye indirectamente a la transferencia de los conocimientos adquiridos de ese modo en beneficio de las redes profesionales que no tienen necesariamente un vínculo directo con la intervención evaluada.

La identificación de buenas prácticas y de lecciones transferibles es esencial para este tipo de utilización.

Fuente: http://ec.europa.eu/europeaid/evaluation/methodology/methods/mth_use_es.htm

¹ EuropeAid. Cooperación y Desarrollo es una división de la Comisión Europea.

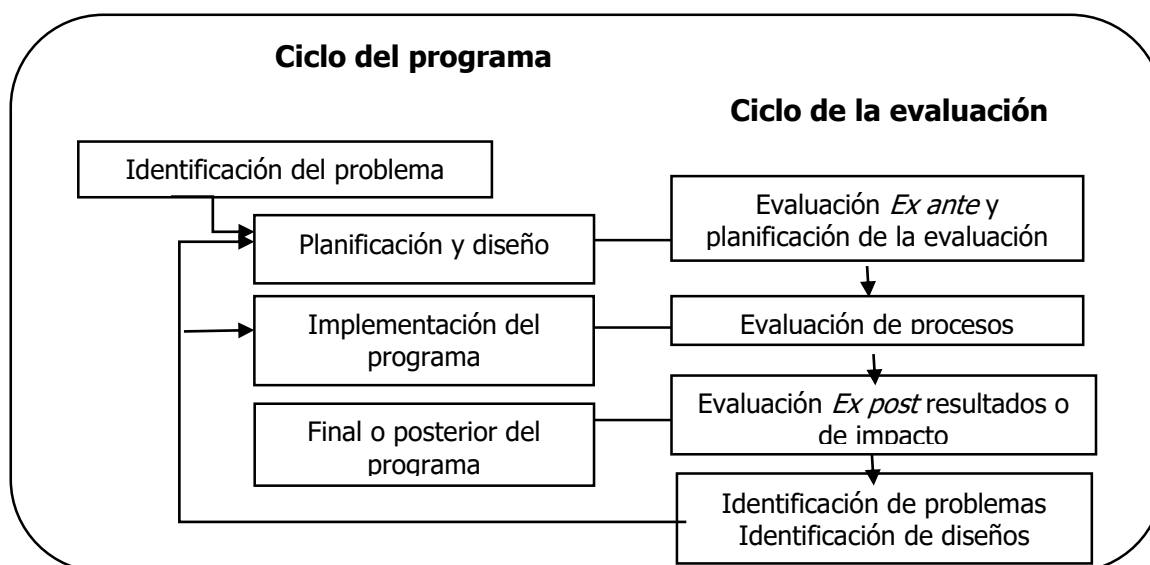
1.1.3 ¿En qué momento de la vida del plan, programa o proyecto se evalúa?

Considerando el ciclo del PPP, la evaluación se realiza desde el inicio, durante su vigencia e incluso posteriormente. La evaluación se inicia en la etapa del diseño o fase de preinversión; allí, además de la planificación de la evaluación, se realiza un tipo de evaluación, la *ex ante*, que nos ayuda a decidir cuáles proyectos implementar.

Durante la ejecución del PPP se realizan evaluaciones de proceso; estas son útiles para saber si los procesos operativos del programa son efectivos o se llevan de acuerdo a lo especificado. A la finalización del programa se desarrollan evaluaciones finales (a veces también de impactos más cercanos). Culminada la vida del PPP se pueden hacer evaluaciones de impacto sobre los efectos de más largo plazo.

Las evaluaciones *ex post* o las de impacto, si bien cierran un ciclo de evaluación, sirven para iniciar un nuevo ciclo de programas o proyectos porque la información que proviene de la evaluación sirve para i) identificar nuevos problemas que deben ser resueltos por un nuevo programa y ii) especificar nuevos diseños de programas o proyectos. A continuación, en la figura 1, se resumen las relaciones que hay entre ciclo de proyecto y ciclo de evaluación.

Figura 1. Relación entre el ciclo del PPP y el ciclo de la evaluación



Elaborado por Alejandro Bardales basado en Valen, V. & Baxter H. 2001.

Para ampliar el tema tratado, revisa la siguiente lectura:

Metodología de evaluación de la ayuda exterior de la comisión europea, Volumen 1, 2006. Páginas 11-22.

1.2 Principios éticos y enfoques de equidad, género e interculturalidad en la evaluación

1.2.1 ¿Bajo qué principios éticos se debe evaluar?

La evaluación se guía por un conjunto de principios universales como el respeto a las personas, la búsqueda del bien, el no hacer daño y la justicia. Se ha desarrollado un conjunto de principios particulares a la evaluación porque esta actividad se enfrenta a problemas éticos y requiere de un conjunto de normas que sirvan para dar pautas a los estudios de evaluación. Existen varios códigos de ética para la evaluación, por ejemplo el de la American Evaluation Association, EuropeAid, el Programa de las Naciones Unidas para el Desarrollo (PNUD), entre otros. En ellos se hace mención a los estándares de calidad, respeto a los participantes, respeto a la diversidad de intereses y valores. A continuación mostramos las normas éticas del PNUD para la evaluación:

Normas del PNUD para evaluaciones

La evaluación debe ser:

- **Independiente.** La administración no debe imponer restricciones en el alcance, contenido, comentarios y recomendaciones de los informes de evaluación. Los evaluadores deben estar libres de cualquier conflicto de intereses.
- **Intencionada.** El fundamento de una evaluación y de las decisiones que se deben basar en ella debería ser claro desde el principio.
- **Transparente.** Una consultación significativa con las partes interesadas es esencial para la credibilidad y utilidad de la evaluación.
- **Ética.** La evaluación no debería reflejar intereses personales o sectoriales. Los evaluadores deben tener integridad profesional, respetar los derechos de instituciones e individuos a proporcionar información de forma confidencial, y ser sensibles a las creencias y costumbres de los entornos culturales y sociales locales.
- **Imparcial.** Eliminar los sesgos y maximizar la objetividad son críticas para la credibilidad de la evaluación y su contribución al conocimiento.

Normas del PNUD para evaluaciones

- **De alta calidad.** Toda evaluación debería cumplir unos estándares mínimos de calidad definidos por la oficina de evaluación.
- **Oportuna.** Las evaluaciones deben estar diseñadas y completadas de forma oportuna, de manera que aseguren la utilidad de los hallazgos y recomendaciones.
- **Útil.** La evaluación es una disciplina de gestión que trata de proporcionar información para que sea utilizada en la adopción de decisiones basadas en las evidencias. Para aumentar la utilidad de las conclusiones y recomendaciones, las principales partes interesadas deben estar involucradas de diferentes formas en la realización de la evaluación.

Fuente: Normas para la Evaluación, PNUD.

1.2.2 ¿Existen evaluaciones que se guían por o sirven directamente a principios universales?

Sí, hay evaluaciones que no solo se guían por principios universales, sino que tienen como objetivo saber cómo se han materializado esos principios en los programas o en la sociedad en general. Estos son los casos de las evaluaciones con enfoque de equidad, de equidad de género, del respeto a la diversidad u otras enfocadas en poblaciones vulnerables como los niños, que tienen como marco de referencia la Declaración Universal de los Derechos Humanos.

a) El enfoque de equidad

El enfoque de equidad considera el trabajo con las poblaciones vulnerables, el respeto a la diversidad, incorpora a los grupos más discriminados de la sociedad y además acepta la necesidad de reducir la pobreza de forma integral en la aplicación del cumplimiento de los Objetivos de Desarrollo del Milenio (ODM).² En este enfoque vienen trabajando agencias de Naciones Unidas como el PNUD, UNICEF, UNEG, entre otros.

Para UNICEF “equidad significa que todos los niños y niñas tengan la oportunidad de sobrevivir, desarrollarse y alcanzar su máximo potencial, sin discriminación, parcialidad o favoritismo”. Este significado se extiende a la población en general y consiste en garantizar el respeto a los derechos fundamentales de todos los hombres y mujeres de cualquier edad,

² En los anexos se puede encontrar un resumen de los mismos y el avance alcanzado al 2013.

independientemente de su sexo, raza, creencias religiosas, ingresos, características físicas, ubicación geográfica o cualquier otra condición.

El enfoque de equidad no significa que todos sean (o deban ser) iguales, sino que los derechos, responsabilidades y oportunidades no dependerán del hecho de que una persona haya nacido hombre o mujer, su edad, lugar de residencia, cultura, etnia, raza, capacidades, color de piel. Implica que se tienen en cuenta los intereses, necesidades y prioridades de todos, reconociendo la diversidad entre los diferentes grupos de pobladores, con la única prioridad de atender a aquellos que se encuentran en situaciones de desventaja, pobreza o inequidad.

b) La equidad de género

La equidad de género se refiere a la igualdad de derechos, responsabilidades y oportunidades de mujeres y hombres, niñas y niños. Igualdad no significa que las mujeres y los hombres son o deban ser iguales, sino que los derechos, responsabilidades y oportunidades no dependerán del hecho que una persona haya nacido hombre o mujer. Implica que se tienen en cuenta los intereses, necesidades y prioridades tanto de hombres como mujeres, reconociendo la diversidad entre los diferentes grupos de mujeres y hombres. La igualdad entre mujeres y hombres, niñas y niños se considera un asunto de derechos humanos y una precondition e indicador del desarrollo sostenible centrado en los derechos humanos. Es también un componente esencial para la realización de todos los derechos humanos (Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género 13-14).

Para la Organización de las Naciones Unidas (ONU), una evaluación sensible a los derechos humanos (DD.HH.) y la igualdad de género aborda los principios de programación tal como lo requieren un enfoque basado en los derechos humanos y una estrategia de transversalización de género. Contribuye al proceso de cambio social y económico, que es esencial para gran parte de la programación de desarrollo, identificando y analizando las desigualdades, prácticas discriminatorias y relaciones desiguales de poder que son el centro de los problemas del desarrollo. La evaluación sensible a los DD.HH. y la igualdad de género puede ayudar a que las intervenciones sean más efectivas y los resultados más sostenibles. (Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género 20).

c) ¿Con qué principios se integra la igualdad de género en la evaluación?

Los principios que integran la igualdad de género en la evaluación han sido tomados del documento “Integración de los derechos humanos y la igualdad de género en la evaluación hacia una guía del United Nations Evaluation Group (UNEG)”, 2011, y son los siguientes:

Inclusión. Evaluar los DD.HH. y la igualdad de género requiere prestar atención a qué grupos beneficia y qué grupos contribuyen a la intervención evaluada. Los grupos deben estar desagregados en criterios relevantes: grupos desfavorecidos y favorecidos dependiendo de su género o situación (mujeres/hombres, clase, etnia, religión, edad, lugar de residencia, etc.), varios tipos de garantes de derechos y varios tipos de titulares de derechos, para determinar si los beneficios y contribuciones de la intervención evaluada fueron distribuidos equitativamente.

Participación. La evaluación de los DD.HH. y la igualdad de género deben ser participativas. Las partes interesadas de la intervención tienen derecho a ser consultadas y a tomar parte en las decisiones sobre qué será evaluado y cómo. Además, la evaluación analizará si las partes interesadas han tenido la oportunidad de participar en el diseño, implementación y supervisión de la intervención. Es importante medir la participación del grupo interesado en el proceso, así como el beneficio obtenido con los resultados.

Relaciones iguales de poder. Tanto los DD.HH. como la igualdad de género buscan, entre otras cosas, el equilibrio de poder en las relaciones entre grupos favorecidos y desfavorecidos y dentro de los mismos. La naturaleza de la relación entre la parte implementadora y la parte interesada de una intervención puede apoyar u obstaculizar el cambio. Cuando las evaluadoras o evaluadores miden el cambio en las relaciones de poder como resultado de una intervención, deben comprender perfectamente el contexto y llevar a cabo la evaluación de forma que apoye el empoderamiento de los grupos desfavorecidos, por ejemplo, el empoderamiento de las mujeres cuando estas son el género en desventaja en un determinado contexto. Además, los evaluadores y evaluadoras deben tener conciencia de su propia posición de poder, que puede influenciar las respuestas a las preguntas que se formulan durante la

interacción con las partes interesadas. Es necesario ser sensible a estas dinámicas.

Se deben usar métodos de evaluación mixtos. Independientemente de la dimensión de la intervención, la evaluación con enfoque de DD.HH. e igualdad de género debe usar métodos cuantitativos y cualitativos.

d) Interculturalidad

En las intervenciones de desarrollo nos encontramos con PPP en los que se involucran actores y comunidades con diversas identidades y tradiciones culturales. El Perú es un país que presenta una amplia diversidad cultural y lingüística, principalmente en las regiones alto andinas y amazónicas. El enfoque de interculturalidad plantea que en la evaluación se debe considerar cuidadosamente la situación en la que convergen diversas culturas y se dan procesos de integración cultural. La empatía y respeto a las identidades culturales es un principio fundamental, y el esclarecimiento de cómo el PPP que es evaluado afecta o promueve los derechos culturales, especialmente de las poblaciones vulnerables.

Para ampliar el tema tratado, realiza la siguiente lectura:

- Integración de los derechos humanos y la igualdad de género en la evaluación—Hacia una Guía del UNEG, 2011. p. 13-16.

1.3 Niveles de la evaluación: políticas, programas y proyectos

Cualquier acción organizada para "resolver" un problema social o situación que se considera indeseable (intervención social) puede ser evaluada. Las intervenciones sociales se clasifican en políticas, programas o proyectos.

Un **proyecto social** es la unidad mínima de asignación de recursos que a través de un conjunto integrado de procesos y actividades pretende transformar una parte de la realidad, disminuyendo o eliminando un déficit, o solucionando un problema.

Un proyecto social debe cumplir las siguientes condiciones:

- Definir el, o los problemas sociales, que se persigue resolver (especificar cuantitativa y cualitativamente el problema antes de iniciar el proyecto).

- Tener objetivos de impacto claramente definidos (proyectos con objetivos imprecisos no pueden ser evaluados).
- Identificar a la población objetivo a la que está destinada el proyecto (la que teniendo las necesidades no puede satisfacerlas autónomamente vía el mercado).
- Especificar la localización espacial de los beneficiarios.
- Establecer una fecha de comienzo y otra de finalización.

Los proyectos sociales producen y/o distribuyen bienes o servicios (~~productos~~) para satisfacer las necesidades de aquellos grupos que no poseen recursos para solventarlos, son autónomos, con una caracterización y localización espacio-temporal precisa y acotada.

Un **programa social** es un conjunto de proyectos que persiguen los mismos objetivos, que pueden diferenciarse por trabajar con poblaciones diferentes y/o utilizar distintas estrategias de intervención.

La **política social** es un conjunto de programas que pretende alcanzar los mismos fines. Da las orientaciones sobre qué problemas sociales priorizar y define las principales vías y/o límites para la intervención que la política plantea.

Los programas y proyectos sociales se enmarcan en una política, constituyen su traducción operacional, mediante la asignación de recursos para su implementación (ver cuadro 1).

En la literatura tradicional se ha asumido que los proyectos se definen por la existencia de inversión, esto es, porque se asignan recursos para la adquisición de bienes de capital (terreno, construcciones, equipamiento). Los programas, alternativamente, supondrían solo la utilización de recursos para gastos corrientes (por ejemplo, los sueldos y salarios que se requieren para la operación). Hoy existe acuerdo en que los proyectos no se definen por la presencia o ausencia de inversión, sino por ser una unidad organizada de gestión que busca solucionar un problema.

Cuadro 1. Ejemplo de niveles de evaluación

Política	Promoción de la Seguridad Alimentaria y Nutrición
Programas	Programa Articulado Nutricional
	Programa Juntos
	Programa Qali Warma
Proyecto	Proyecto de Nutrición Infantil del Gobierno Regional de Puno
	Proyecto Alianza por la Nutrición

Elaborado por Alejandro Bardales, 2014.

Respecto a las evaluaciones de programas y políticas, si bien la mayor parte de autores las considera similares, en este capítulo se presentan algunas diferencias entre ambos niveles. Siguiendo a Bustelo (39), el cuadro 2 presenta cinco dimensiones (o áreas) en las que estas se encuentran.

En cuatro de las dimensiones: objeto, analista, aplicación y plazos, la autora encuentra diferencias en su mayor parte relacionadas a la naturaleza misma de lo que se entiende por “un programa o una política”, es decir cuestiones como alcance, especificidad, control y otros elementos que los diferencian. Sin embargo, al referirse a la metodología coincide con los demás autores en que esta es igual para ambos niveles de intervención social.

Cuadro 2. Diferencias entre la evaluación de programas y la de políticas

Dimensiones	Evaluación de programas	Evaluación de políticas
Objeto de estudio	Planes, programas, proyectos Servicios, medidas	Políticas públicas
Agente evaluador	Consultores externos, Responsables internos	Analistas
Aplicación	Mejora del programa o servicio. Control de responsabilidades Contribución a la evaluación de una política	Iluminación para acciones futuras Ayuda a las decisiones sobre continuación o terminación
Metodología	Flexibilidad metodológica	Flexibilidad metodológica (relativo mayor apoyo en análisis secundario de datos)
Tiempos	Temporalidad ajustada	Puede tomarse más tiempo aunque debe reflejar el carácter dinámico de la política

Fuente: Diferencias entre la Evaluación de Programas y la Evaluación de Políticas. Bustelo.

1.4 Identificación de la "evaluabilidad" de un plan, programa o proyecto

El concepto de evaluabilidad de un proyecto/programa se define como la medida en que el valor generado o los resultados esperados de un proyecto o programa son verificables de manera fiable y creíble.

El **estudio de evaluabilidad** es un ejercicio que ayuda a identificar si una intervención puede ser evaluada, si se justifica hacer una evaluación y si esta es viable y tiene probabilidades de proporcionar información útil. Su propósito no solo es determinar si se puede realizar la evaluación o no, sino también preparar la intervención para que esta tenga todas las condiciones necesarias para ser evaluada (Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género 18).

En el análisis de evaluabilidad se examinan los siguientes componentes o aspectos del PPP: diagnóstico, lógica del proyecto, supuestos y riesgos, indicadores, datos de línea de base, y plan de monitoreo y evaluación.

Más adelante, en la preparación del diseño de evaluación se abordará con mayor detalle las características relevantes de estos componentes para determinar las condiciones de evaluabilidad del PPP.

1.5 Enfoques sobre evaluación

Existen diversos enfoques sobre la evaluación que la orientan o guían: i) en la elección sobre qué enfatizar o interesarse, ii) en definir la labor del evaluador, iii) en la elección de los métodos a usar iv) en cómo considerar a los intereses que hay en una evaluación.

De acuerdo a Virtanen & Uusikylä (2004), estos enfoques son tres:

- a) El enfoque libre de objetivos da importancia a la descripción de los procesos y busca responder a los intereses de los involucrados.
- b) El enfoque con objetivos predeterminados se orienta a saber si se alcanzaron los objetivos; intenta ubicar y demostrar las relaciones causales existentes entre lo actuado y los resultados obtenidos, intentando verificar el nivel de responsabilidad que tiene el programa en los logros.
- c) El enfoque realista viene a ser una combinación de los dos anteriores. Toma el análisis del contexto, los mecanismos de evaluación y los relaciona con los resultados obtenidos (CMO por las siglas en inglés de *context, mechanisms y outcomes*).

El cuadro 3 compara los tres enfoques de acuerdo a diversas dimensiones que amplían la definición y caracterización de cada enfoque de evaluación.

Cuadro 3. Enfoques principales en la evaluación y su caracterización respecto a dimensiones relevantes

DIMENSIONES	ENFOQUES		
	Enfoque libre de objetivos	Enfoque con objetivos predeterminados	Enfoque realista
Base filosófica	Fenomenología	Positivismo	Realismo
Autor ejemplo	Guba and Lincoln, 1989	Vedung, 1997 Rossi et al, 1999 Mohr, 1995	Pawson y Tilley, 1997
Base disciplinaria	Sociología Antropología	Teoría económica Teoría política	Política social Sociología
Enfoque de la metodología	Describir integralmente el programa y desde la perspectiva de los participantes	Juicio de valor sobre el programa Identificar vínculos causales	A fin de identificar y analizar las relaciones entre Mecanismos, Contexto y Resultado (CMR)
Metodología	Etnografía, estudios de caso; observación participante; triangulación	Diseños experimentales y cuasi experimentales diseños análisis costo-beneficio	Contexto + Mecanismo = Resultado
Interpretación de la causalidad	Reacios a trazar la conexión entre causas y efectos	Principal preocupación	Un asunto de la agenda de investigación, un énfasis especial en el marco CMR
Variables	Emergen en el proceso de evaluación	Predeterminadas como entrada-salida	Emergentes y predeterminadas
Grupos de comparación o control	No necesariamente	Sí	No necesariamente
Los participantes a la realización de evaluación	Activo	Ninguno	Activo
Papel del evaluador	Interactivo	Independiente del programa	Observación
Las presiones políticas (interna/externa)	Describir	Controlada en el diseño	Describir y explicar
Enfoque del informe de evaluación	Imagen integral del programa en proceso	Tomar una decisión de ir o no ir	Para comprender e informar las configuraciones CMO

Fuente: (Virtanen y Uusikylä 82).

1.6 Tipos de evaluaciones

Las evaluaciones se diferencian, dando lugar a diversos tipos, según la etapa en la que se encuentra el PPP y en razón de los resultados previstos, a lo que se puede agregar, en nuestro caso, el presupuesto por resultados.

1.6.1 Según la etapa del plan, programa o proyecto

Una de las principales formas de clasificar las evaluaciones prioriza el momento del programa o proyecto en que este se realiza. Considerando dicho parámetro las evaluaciones se clasifican en:

- a. *Ex ante*: las cuales se realizan antes de que la iniciativa se ponga en marcha, a fin de determinar principalmente la pertinencia y viabilidad de implementar la propuesta.
- b. De proceso: son aquellas evaluaciones que se realizan ya iniciada la ejecución del programa/proyecto, y tienen por finalidad analizar el avance en el logro de productos y/o el desempeño en la ejecución.
- c. *Ex post*: son las evaluaciones que se realizan habiendo finalizado la intervención, a fin de determinar, la eficacia en el cumplimiento de los productos y resultados del plan, programa o proyecto; evaluar la eficiencia en el uso de los recursos, estimar si se logró los impactos previstos, entre otros aspectos.³

El cuadro 4 presenta las características de estos tipos de evaluación de acuerdo con las preguntas clave ¿Qué? ¿Cuándo? ¿Por qué? ¿Cómo?

³Cabe indicar que si bien las evaluaciones de impacto se ejecutan usualmente al término del programa/proyecto, en caso la intervención tuviera un periodo largo de implementación se puede llevar a cabo una medición de impacto durante el desarrollo de la intervención en las denominadas evaluaciones intermedias.

Cuadro 4. Características de los tipos de evaluación según la etapa del PPP

Preguntas clave	Evaluación <i>ex ante</i>	Evaluación de procesos	Evaluación <i>ex post</i> ⁴	
			Evaluación de resultados	Evaluación de impacto
¿Qué?	Evalúa una propuesta de intervención que busca la solución de un problema específico.	Evalúa la dinámica de la intervención (actividades realizadas, productos obtenidos, problemas identificados, entre otros).	Evalúa en qué grado se logra el resultado o cambio deseado, es decir, en qué medida se cumplen con los objetivos de la intervención.	Determina relaciones de causa-efecto y estima magnitudes específicas de cambio. Es decir, permite medir el grado de cambio que se atribuye a la intervención.
¿Cuándo?	Se da antes de la intervención del programa o proyecto.	Durante la intervención del programa o proyecto para responder preguntas de gestión y rendición de cuentas.	Se aplica en un momento específico de la intervención, usualmente durante su implementación.	Se aplica al término de la intervención. Preferentemente después de transcurrido un tiempo prudencial.
¿Por qué?	Permite elevar la probabilidad del buen uso de los recursos y sincerar los impactos esperados con los recursos disponibles.	Permite identificar los aspectos positivos en la ejecución de la intervención para intensificarlos y los negativos para mejorarlos.	Permite determinar si se están logrando los objetivos de la intervención.	Permite comprobar la efectividad de un programa o proyecto.
¿Cómo?	Puede basarse en variados tipos de análisis: costo-beneficio, costo-impacto, costo-eficiencia y de diseño (basado en la pertinencia y coherencia de la intervención).	Se describe e interpreta lo que ocurrió. Ello se puede realizar a partir de la lectura de documentos, las entrevistas a personas, los registros del administrador, etc.	Se basa en un amplio rango de métodos: cualitativos y/o cuantitativos. Compara las características del diseño del programa con el actual desempeño del programa. No requiere un grupo de comparación.	Responde a preguntas de atribución de los resultados observados ¿El programa es efectivo comparado con la ausencia del mismo? Se compara lo que ocurrió respecto a lo que habría ocurrido en la ausencia del programa (contrafactual).

Elaborado por Alejandro Bardales, 2014.

⁴ En algunos textos sobre evaluación de programas/proyectos se encuentra que la evaluación *ex post* se define como aquella que se ejecuta luego de iniciada la implementación o que se realiza al finalizar la intervención. No existe consenso en la literatura al respecto.

Evaluación formativa y evaluación sumativa son otros dos tipos de evaluación que también tienen relación con la etapa del PPP. Cuando la evaluación se lleva a cabo en un punto intermedio del ciclo del PPP, los resultados de la evaluación aportan al mejoramiento del diseño y ejecución, el interés primario es hacer las correcciones y ajustes a tiempo para asegurar el éxito del PPP. Cuando la evaluación se lleva al término y post término, el interés primario es determinar el mérito o valor del PPP, se la considera una evaluación sumativa. Se podría decir que evaluaciones de desempeño y resultados, cuando se realizan a mitad de término del proyecto tienden a tener un carácter formativo, y las que se realizan ex post, de resultados e impacto, un carácter sumativo.

1.6.2 Según los resultados del plan, programa o proyecto

Por otro lado, una segunda clasificación se enfoca en los resultados y distingue entre evaluación de impacto y evaluación de desempeño (ver cuadro 5).

- a. De desempeño: medida en que una intervención para el desarrollo o una entidad que se ocupa de fomentar el desarrollo actúa conforme a criterios/normas/directrices específicos u obtiene resultados de conformidad con metas o planes establecidos (OECD, 2010).
- b. De impacto: este tipo de evaluación se enfoca en las relaciones de causa-efecto a fin de determinar la atribución de los cambios observados a la intervención, esto mediante una comparación de lo que ocurrió con respecto a lo que habría ocurrido en la ausencia del programa/proyecto.

Si bien las evaluaciones de planes, programas o proyectos según el aspecto de la intervención que analizan pueden diferenciarse claramente en los tres tipos señalados líneas arriba, en la práctica, dado que estas evaluaciones analizan aspectos complementarios de una intervención, es posible desarrollar diseños mixtos de la evaluación donde, por ejemplo, se considere objetivos de evaluación tanto de desempeño como de diseño o como una evaluación final que combine un análisis de las dimensiones sobre el desempeño que tuvo el programa con una evaluación de impacto y resultados.

Cuadro 5. Diferencia entre la evaluación de desempeño y la evaluación de impacto – USAID

	Evaluación de desempeño	Evaluación de impacto
¿Qué?	Evaluar cómo se consiguen los resultados.	Determinar la causa, el efecto y la magnitud del cambio específico.
¿Cuándo?	Se planean los tiempos específicos para dar respuesta a las cuestiones de gestión y rendición de cuentas.	Se usa el <i>ex post</i> para probar las nuevas hipótesis de desarrollo, demostrar conceptos, informar para la toma de decisión.
¿Por qué?	Se utiliza para ver e informar lo que el programa ha hecho, cómo se aplica, si se están produciendo los resultados esperados y otros muy pertinentes de la gestión y las preguntas relacionadas con el diseño.	Se utiliza para la elaboración de políticas de USAID y aprender acerca de lo que funciona y lo que no en la comunidad del desarrollo. También se puede abordar el cómo y el por qué cuando se combina con la evaluación de desempeño.
¿Cómo?	Puede estar basado en una amplia gama de métodos cualitativos, cuantitativos o mixtos.	Generalmente se requiere una situación hipotética (grupo control) combinada con una fuerte línea de base y post encuesta del programa.
¿Cómo	Comparar las características de diseño del programa para el rendimiento y hacer juicios sobre el valor del programa.	Comparar los impactos reales con las proyecciones de la etapa de diseño.

Fuente: USAID. R-1: Monitoring and Evaluation.

Para ampliar el tema tratado, revisa las siguientes lecturas:

- Independent Evaluation Group. 2013. World Bank Group Impact Evaluations, Relevance and Efectiveness. The World Bank Group. Disponible en:
<https://openknowledge.worldbank.org/bitstream/handle/10986/13100/757230PUB0EPI00013000Pubdate0209013.pdf?sequence=>
- Gertler, P. y otros. La Evaluación de Impacto en la Práctica. Washington DC. The International Bank for Reconstruction and Development, 2011. p. 1-47.

1.6.3 Tipos de evaluación del Presupuesto por Resultados (PpR) considerados en el Ministerio de Economía y Finanzas (MEF) del Perú

- a. Evaluación de Desempeño y Ejecución Presupuestal (EDEP): se enfoca en el análisis de la lógica del diseño, así como en la implementación y desempeño (eficiencia, eficacia, calidad) en la entrega de bienes.
- b. Evaluación de Impacto: se centra en la medición de los resultados enteramente atribuibles a los productos entregados mediante la estimación de escenarios contrafactuales.

1.6.4 Tipos de evaluación considerados en el Ministerio de Desarrollo e Inclusión Social (MIDIS) del Perú

El MIDIS define los siguientes tipos de evaluación a fin de garantizar que las evaluaciones de políticas, planes, programas y proyectos sociales en materia de desarrollo e inclusión social estén orientadas a resultados.

- a. Evaluación de diseño: analiza sistemáticamente el diseño de las intervenciones públicas sujetas a evaluación con el fin de determinar si las intervenciones son pertinentes para la consecución de los objetivos planteados.
- b. Evaluación de procesos: analiza mediante trabajo de campo si el programa lleva a cabo sus procesos operativos de manera eficaz y eficiente y de calidad.
- c. Evaluaciones de resultados: analiza sistemáticamente el diseño y desempeño de los programas para mejorar su gestión y medir el logro de sus objetivos.
- d. Evaluación de impacto: identifica, mediante metodologías rigurosas, si existen cambios en los indicadores de resultados que sean atribuibles a la ejecución de las intervenciones públicas.
- e. Revisiones sistemáticas y metanálisis de evaluaciones de impacto: analiza y sintetiza sistemáticamente la literatura de evaluaciones de impacto aplicando los filtros correspondientes a fin de garantizar su rigurosidad.

- f. Evaluaciones *ex ante*: analiza la pertinencia, la factibilidad y la sostenibilidad potencial de una intervención para el desarrollo antes de tomar la decisión de otorgar financiamiento.

1.6.5 Sobreposición entre los tipos según etapa y resultados

En la práctica de la evaluación los tipos de evaluación mencionados no son mutuamente excluyentes. Una evaluación puede ubicarse en la intersección de los diversos tipos tal como se muestra en el Cuadro 6, en donde también se ha ubicado en las cuadrículas las referencias realizadas a los tipos de evaluación de los presupuestos por resultados del MEF, y los tipos de evaluación del MIDIS.

Cuadro 6. Coincidencias de los tipos de evaluación considerando ambos criterios de clasificación

		TIPOS SEGÚN ETAPA DEL PPP			
		EX ANTE	DESEMPEÑO	EX POST	
				RESULTADO	IMPACTO
		MIDIS	MIDIS MEF-EDEP		
TIPOS SEGÚN RESULTADOS	RESULTADO		X	X MIDIS	
	IMPACTO			X	X MIDIS MEF-IMP

Para profundizar el tema tratado, revisa la siguiente lectura:

- Lineamientos para el Seguimiento, Evaluación, y Gestión de la Evidencia del MIDIS, 2012.

1.7 Validez y confiabilidad del diseño de evaluación

La validez y la confiabilidad de las evaluaciones radican en los siguientes aspectos.

Pertinencia de los indicadores de resultados. Los indicadores deben tener estrecha relación con la cadena de resultados del proyecto, deben estar claramente definidos, las unidades de medida deben estar claramente establecidas, deben ser diferentes de los indicadores de productos y los instrumentos diseñados para su obtención deben haber sido adecuadamente validados.

Solidez de la metodología de recolección de datos. El universo de análisis debe estar claramente establecido, el diseño de la muestra debe estar claramente establecido y tener pertinencia desde el punto de vista de precisión y eficiencia en costos, el plan de recolección de datos debe estar bien detallado y los instrumentos deben haber sido validados y/o probados en campo previamente.

Calidad de datos. Las fuentes de los datos a emplear o empleados, los métodos de verificación de la calidad de los datos cuando son obtenidos de fuentes primarias (encuestas), los criterios de depuración de datos, métodos empleados para imputar los datos faltantes, así como tasas de no respuesta de las encuestas de donde se obtienen los datos deben estar establecidos claramente.

El método empleado para medir las atribuciones del proyecto en el impacto. Se debe tener claramente establecido que el método aplicado sea pertinente de acuerdo a la naturaleza de los datos y que se cumplan los supuestos que asumen los modelos a emplear.

La robustez del análisis. Los resultados deben ser obtenidos con métodos estadísticos de estimación adecuados para el tipo de modelo empleado y los tipos de datos disponibles, deben tener claramente establecidos los márgenes de error de los resultados.

1.8 El proceso de diseño de una evaluación

Cada tipo de evaluación analiza diferentes aspectos de un plan/programa/proyecto, por lo cual, en primer lugar, es necesario definir los objetivos de la evaluación que se va a realizar; es decir, se debe precisar si el objetivo es hacer una evaluación de desempeño, de resultados a la mitad del periodo de ejecución, de impacto, etc., según los diferentes tipos revisados anteriormente; estos objetivos tienen que estar claramente establecidos para poder diseñar la evaluación y, de manera más específica, para definir las preguntas de evaluación.

En el diseño de cada evaluación se requiere definir una metodología apropiada que combine métodos e instrumentos cuantitativos y/o cualitativos específicos. Sin embargo, de manera general se pueden identificar los pasos que se presentan en el cuadro 7⁵.

Cuadro 7. Pasos a seguir en el diseño de la evaluación

1. Diagnóstico del contexto (1)
2. Descripción del plan, programa o proyecto. (3)
3. Identificación de objetivos / cadenas de cambio. (4)
4. Identificación de los interesados. (2)
5. Evaluabilidad del PPP
6. Definición de las preguntas de evaluación. (5)
7. Definición de los métodos: cuantitativos, cualitativos o mixtos. (6)
8. Identificación de las dimensiones, variables e indicadores a medir/verificar y las fuentes de información. (7)
9. Determinación de la muestra. (10) [Población Unidad]
[Tamaño efecto (8)]
10. Diseño de los instrumentos de recojo de información (9)
11. Preparación del plan de recolección y análisis. (11)

En las siguientes unidades se discutirá cada uno de estos pasos; sin embargo, es

⁵Holden, D., Zimmerman M. (eds.). A Practical Guide to Program Evaluation Planning. Theory and Case Examples. California. Sage Publications, Inc. 2009. Chapter II. 2010.

importante indicar que las evaluaciones de impacto requieren el uso de métodos cuantitativos sofisticados y un diseño diferenciado de acuerdo al tipo de método que se haya determinado utilizar; para profundizar sobre la evaluación de impacto, se requiere la revisión de la lectura recomendada.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 1

Ahora procede a desarrollar la actividad aplicativa de la unidad



Actividad 1.

Participa en el FORO: Funciones de la Evaluación

1. Ingresa a la sección de foros de la plataforma de aprendizaje virtual (PAV) y participa con opiniones, comentarios y aportes sobre el tema.
2. Intercambia puntos de vista y experiencias personales e institucionales en torno a la evaluación.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

1. ¿En cuál de los siguientes tipos de evaluación se halla incluido el monitoreo?
 - a) La evaluación *ex ante*
 - b) La evaluación de procesos
 - c) La evaluación *ex post*

2. Estimar los beneficios sociales del programa en un estudio de preinversión es una típica...
 - a) Evaluación *ex ante*
 - b) Evaluación de procesos
 - c) Monitoreo
 - d) Evaluación de impacto

3. ¿Qué enfoque de evaluación está relacionado con la siguiente afirmación: "El principal desafío para llevar a cabo evaluaciones eficaces de impacto es identificar la relación causal entre el proyecto, el programa o la política y los resultados de interés" (La evaluación de impacto en la práctica, p. 4, 2011).
 - a) Enfoque libre de objetivos
 - b) Enfoque realista
 - c) Enfoque con objetivos predeterminados



UNIDAD DE APRENDIZAJE 2: DISEÑO DE EVALUACIONES**2.1 Diagnóstico del contexto****2.1.1 Análisis del entorno**

El análisis del contexto donde se da la intervención del programa ayuda a entender el desempeño del PPP y a identificar las oportunidades y riesgos para su implementación. En este análisis se identifican los elementos que están fuera de las fronteras del programa y que tienen el potencial para afectarlo en todo o en parte, **conjunto de elementos que constituyen el entorno del programa.**

Para Bamberger (2007, p. 25) "El análisis contextual evalúa la influencia de los factores económicos, políticos, institucionales y medioambientales en la ejecución y el efecto directo de los proyectos. Además, examina la influencia de las características socioculturales preexistentes de las poblaciones objetivo en la respuesta de distintos grupos al proyecto. Si bien la mayor parte del análisis contextual es cualitativo (por ejemplo, entrevistas con informantes clave, revisión de la documentación del proyecto y **observación participante**, este puede incluir también un análisis cuantitativo de los datos recogidos en encuestas de hogares".

Este enfoque supone, razonablemente, que las características del entorno influyen en la estructura, estrategias, procesos y capacidades de los programas, los cuales a su vez tienen un efecto sobre el desempeño y el impacto que el programa pueda tener. Las variables del entorno se estudian en los modelos cuantitativos como variables moderadoras sobre el desempeño o impacto del programa. Los factores exógenos actúan como fuerzas institucionales, presiones de interesados, regulaciones, necesidades.

2.1.2 ¿Cómo hacer un análisis del entorno?

Primero, hay que identificar las variables de entorno que puedan influir en el programa y luego seleccionar aquellas que son relevantes. La relevancia de una variable se reconoce porque i) podría determinar directamente el éxito o fracaso del programa ii) tiene múltiples impactos sobre otras variables iii) genera un alto nivel de incertidumbre (Licha, 2000).

En el análisis del entorno se recoge información sobre los siguientes aspectos:

- Las necesidades o problemas que el programa intenta cambiar
- El político e institucional
- El económico
- El sociocultural

Cada programa tiene sus variables de entorno relevantes, por lo cual no es necesario hacer esfuerzos sistemáticos (recoger y clasificar muchas variables de todos los entornos posibles) para estudiar el entorno del programa.

1º. Análisis de las necesidades que el programa intenta o intentará satisfacer. Para ello se analizan diagnósticos, estudios de línea de base o informes de otras instituciones.

2º. Análisis del contexto político e institucional que puede afectar el programa.

3º. Análisis de las variables del entorno económico que pueden afectar el programa, tales como:

- El gasto público
- El crecimiento económico

4º. Análisis de las variables del entorno sociocultural que son relevantes para su desenvolvimiento: creencias, valores, actitudes y formas de vida de las personas. Por ejemplo, en el caso de salud:

- Creencias y conocimientos sobre cómo se transmiten las enfermedades
- Percepciones y creencias sobre la higiene
- Conocimientos y prácticas sobre higiene

A continuación, a manera de ilustración, se presenta el caso de un programa de agua y saneamiento rural, respecto al análisis de necesidades y análisis político-institucional (Cuadros 8 y 9).

Cuadro 8. Ejemplo del análisis de necesidades para el Programa de Agua y Saneamiento Rural para el Perú

El problema de agua y saneamiento en zonas rurales
1 de cada 100 hogares rurales tiene acceso a agua segura o potable
16 de cada 10 hogares cuenta con servicios de saneamiento
90% de la población rural no se lava las manos
98% manipula el agua de manera inadecuada
54% mantiene sus viviendas y letrinas sucias
12,8% de los niños menores de 5 años padece de enfermedades diarreicas
31,9% de niños sufre de desnutrición crónica

Fuente: <http://pnsr.vivienda.gob.pe/index.php?accion=problematica-rural>

Esta información puede complementarse con fuentes cualitativas, como puede verse en el siguiente ejemplo:

El tema de agua en todos los lugares investigados es altamente relevante e incluso coincidente con las necesidades y prioridades establecidas por las comunidades rurales o zonas urbanas marginales de Lima. En todos los ámbitos evaluados la intervención del agua servirá para atender a una demanda identificada por la población. "El agua es lo de más preocupación porque en años secos o sequías sufrimos escasez de agua, lo que pasa es que tenemos un pozo pero ya no sale agua. Lo que hacemos entonces es ir a los ríos, donde hacemos cola; esa es la preocupación más grande..."

Cuadro 9. Ejemplo del análisis político institucional para un Programa de Agua y Saneamiento Rural para el Perú

Institución	Responsabilidad
Ministerio de Vivienda	Financiamiento y supervisión de obras en el medio rural.
	Órgano rector del sector.
Ministerio de Salud Dirección General de Salud Ambiental (DIGESA)	Responsable por la calidad del recurso hídrico: <ul style="list-style-type: none"> ▪ Formula políticas y dicta normas de calidad sanitaria y protección ambiental. ▪ Vigila la calidad de las aguas. ▪ Apoya en la formulación de proyectos de saneamiento en el medio rural. ▪ Autoriza vertimientos. ▪ Aprueba proyectos de plantas de tratamiento de agua y aguas residuales.
Ministerio de Economía y Finanzas	Negocia préstamos, regula la actividad financiera de las empresas del Estado.
Superintendencia Nacional de Servicios de Saneamiento (SUNASS)	<ul style="list-style-type: none"> ▪ Regula tarifas y fiscaliza la prestación de los servicios de saneamiento. ▪ Resolver en última instancia administrativa, las controversias por los servicios de las Entidades Prestadoras de Salud (EPS), entre EPS y usuarios. ▪ Emite normas complementarias y promueve el desarrollo de las EPS (Registro EPS, información técnica, comercial, tarifaria).
Ministerio de Agricultura	Otorga los derechos de uso del recurso hídrico.
Municipalidades, centros poblados rurales, Juntas Administradoras de Servicio y Saneamiento (JASS), Asociaciones de usuarios	Responsables de la prestación de los servicios, otorgan derecho de gestión de servicios, apoyan en construcción de proyectos de saneamiento, administración.

A esta información podemos agregarle el estado de la legislación vigente y las principales políticas sobre agua y saneamiento.

Las leyes relevantes sobre agua y saneamiento son las siguientes:

- Ley N° 26284, Ley General de la Superintendencia Nacional de Servicios de Saneamiento.
- D.S: N° 24-94-PRES. Reglamento de la Ley General de la Superintendencia Nacional de Servicios de Saneamiento.
- Ley N° 26338. General de Servicios de Saneamiento (julio 94).
- D.S: N° 09-95-PRES. Reglamento de la Ley General de Servicios de Saneamiento (1995).
- Ley General de Servicios de Saneamiento (Decreto Ley N° 26338).
- Texto Único Ordenado del Reglamento de la Ley General de Servicios de Saneamiento (Decreto Supremo N° 023-2005-VIVIENDA).
- Centro poblado rural: No sobrepasa 2 mil habitantes. Pequeña ciudad: entre 2 mil y 15 mil habitantes. N° 908 – SUNASS – Ley de fomento y desarrollo del sector saneamiento (2000).
- Ley de Residuos Sólidos.
- Ley Orgánica del Poder Ejecutivo (Ley N° 29158).
- Ley de Organización y Funciones del Ministerio de Vivienda, Construcción y Saneamiento (MVCS) (Ley N° 27792).
- Reglamento de Organización y Funciones del MVCS (Decreto Supremo N° 002-2002-VIVIENDA).
- Creación del Programa Nacional de Saneamiento Rural en el MVCS (Decreto Supremo N° 002-2012-VIVIENDA).
- Manual de Operaciones del Programa Nacional de Saneamiento Rural (Resolución Ministerial N° 096-2012-VIVIENDA).

Programa Nacional de Saneamiento Rural (PNSR): programa del MVCS orientado a posibilitar el acceso de la población del ámbito rural, al agua y saneamiento de calidad y sostenible. Lo integran los programas del Sistema Nacional de Inversión Pública (SNIP), el Programa Nacional de Agua y Saneamiento Rural (PRONASAR), el Programa de Mejoramiento y Ampliación de Servicios de Agua y Saneamiento en Perú (PROCOES) y Amazonía Rural.

Las Políticas sobre agua y saneamiento están contenidas en la [Resolución Ministerial N° 031-2013](#) y en el Plan de Mediano Plazo 2012-2016 del PNSR aprobado por el Ministerio de Vivienda, Construcción y Saneamiento. Este Plan incluye los Lineamientos de Política en Agua y Saneamiento Rural:

- Gestión del servicio por parte de los operadores.
- Marco normativo.
- Fortalecimiento institucional.
- Articulación intergubernamental, intersectorial e interinstitucional.
- Sistema de gestión.
- Comunicación para el cambio social y de comportamiento.
- Cobertura y calidad del servicio.
- Sostenibilidad en la provisión de servicios de calidad.
- Fortalecimiento institucional.

Fuente: <http://pnsr.vivienda.gob.pe/index.php?accion=problematica-rural>

2.2 Descripción del plan, programa o proyecto

La descripción del PPP es el paso en el cual el evaluador revisa minuciosamente la información documental y otras publicaciones relacionadas y establece claramente los resultados esperados, esto con la finalidad de conocer las características intrínsecas de la intervención y tener un punto de referencia para llevar a cabo la comparación entre “lo que es” y lo que “debería ser” dicho plan, programa o proyecto. Este proceso provee la información necesaria para evaluar la forma en que la intervención atiende a sus beneficiarios (Holden, D.; Zimmerman M.; 2010).

2.2.1 ¿Cómo realizarlo?

Es de vital importancia conocer las etapas de desarrollo del PPP, así como el contexto en el que se encuentra funcionando. Para ello se debe:

- Recopilar y revisar minuciosamente toda la información básica del diseño y planificación del PPP, informes de monitoreo y/o de evaluaciones anteriores, planes de trabajo, informes de misiones, así como toda información que brinde información relevante sobre el desarrollo de las actividades de la intervención.
- Se puede entrevistar al personal clave y a las partes interesadas para comprender la historia del PPP desde su implementación y evolución, a fin de luego desarrollar modelos lógicos o marcos conceptuales y crear herramientas que puedan identificar temas de interés para la evaluación.

2.2.2 Herramientas

El proceso de descripción del programa ha sido tratado por muchos investigadores que sugieren el uso de herramientas tales como modelos lógicos o marcos conceptuales a fin de aclarar el propósito para el cual está diseñado el programa y qué productos y actividades son relevantes para los objetivos de la evaluación.

En términos generales, dichas herramientas pueden ayudar a los evaluadores a considerar si:

- Los objetivos del PPP están bien definidos.
- Las relaciones causales entre las actividades y objetivos del plan, programa o proyecto son realistas y verificables.

- Los usos previstos de la información de evaluación están bien definidos.

A continuación se presenta la información básica que se debe recabar en el proceso de descripción del plan, programa o proyecto:

Información básica

1. Componentes, Procesos, Actividades.
 - a. Protocolos, procedimientos y manuales.
 - b. Cualquier información que permita responder a las preguntas: ¿qué sucede con los beneficiarios? ¿Qué es lo que hace el programa?
 - c. Observación de las actividades del programa.
2. Planes, informes de evaluaciones anteriores, memorándums internos concernientes a la intervención.
3. Características sociodemográficas de los beneficiarios.
4. Uso del servicio por parte de los beneficiarios (tipo y calidad de los servicios recibidos).
5. Información sobre la satisfacción de los beneficiarios, reporte de las percepciones de los beneficiarios acerca de las fortalezas y debilidades del plan, programa o proyecto.
6. Información financiera, costos del programa y gastos en general.

2.3 Explicitación de la cadena de cambios: lógica de la intervención

El siguiente paso en el diseño de la evaluación es identificar claramente los objetivos, productos y actividades del PPP que serán revisados en la evaluación. Para ello, es necesario que el evaluador reconozca la lógica en la cual se plantearon las actividades que llevaría a cabo el programa a fin de lograr sus objetivos.

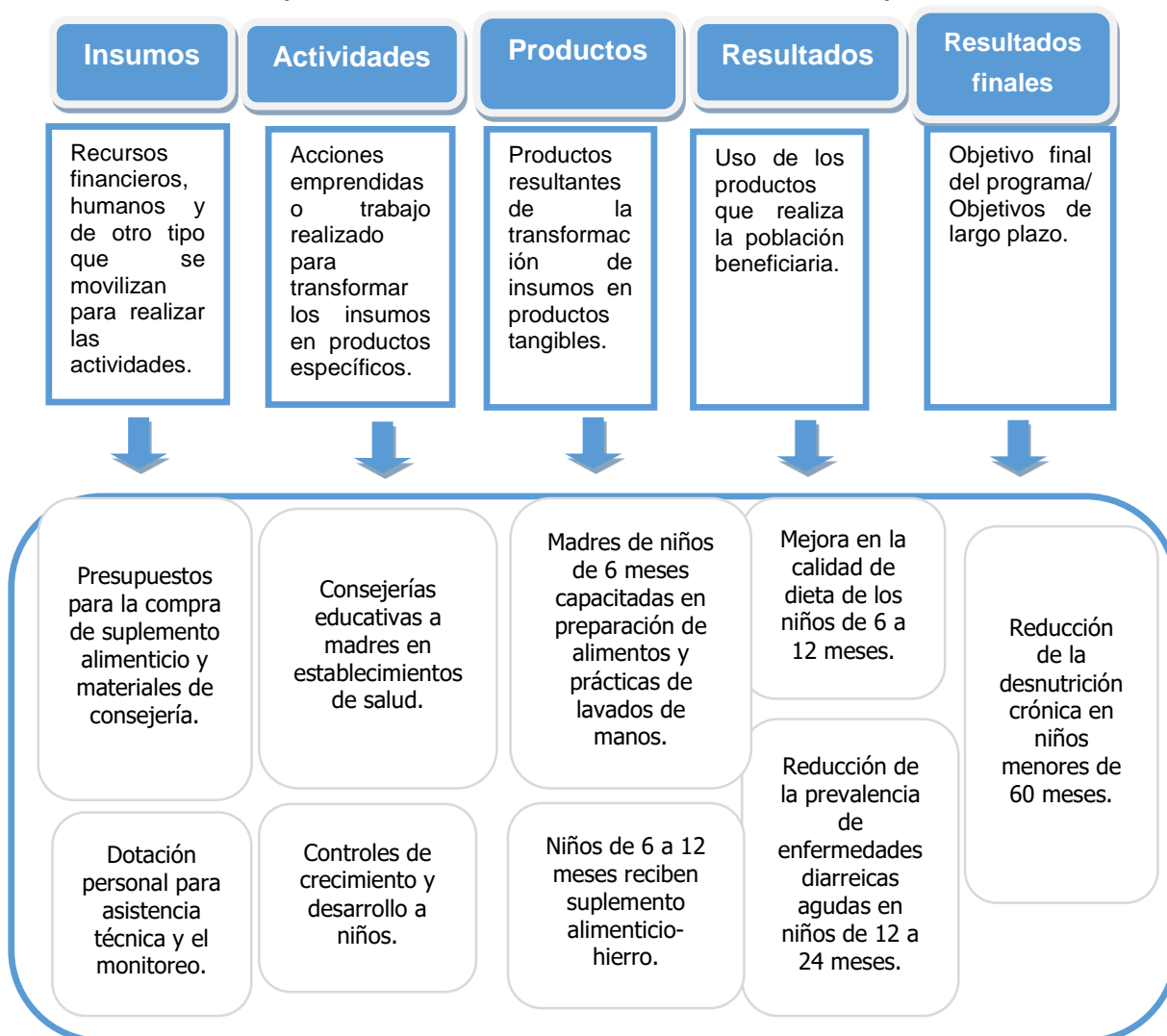
La cadena de cambios prevista en el PPP da una definición lógica y plausible de cómo la secuencia de insumos, actividades y productos relacionados constituye la vía para lograr los impactos.

Una cadena básica de resultados contiene los siguientes elementos:

- **Insumos:** los recursos de que dispone el proyecto, que incluyen el personal y el presupuesto.
- **Actividades:** las acciones emprendidas y el trabajo realizado para transformar los insumos en productos.
- **Productos:** los bienes tangibles y los servicios que producen las actividades del programa (están directamente bajo el control del organismo ejecutor).
- **Resultados:** los resultados que se espera alcanzar una vez que la población se beneficie de los productos del proyecto (los resultados se dan normalmente a corto o mediano plazo).
- **Resultados finales:** Los objetivos finales del proyecto pueden estar influidos por múltiples factores y se alcanzan normalmente después de un largo periodo. También se les denomina impactos.

En la Figura N° 2, que se ofrece en la siguiente página, se ilustra esta cadena básica.

Figura 2. Cadena de resultados
(Con referencia a un PPP sobre salud infantil)



Fuente: Liz Girón, 2014.

Las relaciones de causalidad señaladas son conceptuadas como la teoría de cambio del PPP. Para la evaluación es necesario explicitarla y analizarla.

2.3.1 ¿Qué es una Teoría de Cambio?

Una Teoría de Cambio es una descripción de cómo se supone que una intervención conseguirá los resultados deseados. Describe la lógica causal de cómo y por qué un PPP o una política lograrán los resultados deseados o previstos. Así, la Teoría del Cambio es fundamental para cualquier evaluación, ya que al

describir en secuencia lógica la estrategia empleada para alcanzar los objetivos que debe lograr la intervención, permite identificar y formular concretamente las preguntas y precisar los objetivos de la evaluación (Gertler P., et. al; 2011).

Las teorías del cambio describen una secuencia de eventos que generan resultados:

- Examinan las condiciones y las suposiciones necesarias para que se produzca el cambio.
- Explicitan la lógica causal detrás del programa y
- Trazan el mapa de las intervenciones del programa a lo largo de las vías lógicas causales.

Las teorías del cambio pueden ayudar a determinar qué insumos y actividades se requieren, qué productos se generan y cuáles son los resultados finales derivados de los cambios de comportamiento de los beneficiarios.

Es importante que la Teoría del Cambio se formule al inicio de la intervención con la participación de los beneficiarios a fin de que queden claramente definidos los objetivos. Además, al diseñar el programa se debe revisar la bibliografía especializada en busca de documentos que describan experiencias y programas similares, y se deben comprobar los contextos y las suposiciones de la Teoría del Cambio.

2.3.2 De la Teoría de Cambio a la cadena de cambio o de resultados

"La cadena de cambio es una forma de esquematizar una Teoría de Cambio"

Las Teorías de Cambio se pueden esquematizar de varias maneras. Por ejemplo, mediante modelos teóricos, modelos lógicos, marcos lógicos, modelos de resultados y cadenas de resultados. Todos ellos incluyen los elementos básicos de una Teoría del Cambio, es decir, la cadena causal, que debe ser capaz de discernir entre los logros del programa y las condiciones o influencias ajenas al programa.

Los resultados no dependen solo del control directo del proyecto sino también de cambios de comportamiento de los beneficiarios del programa. En otras palabras, dependen de las interacciones entre el lado de la oferta (implementación) y el

lado de la demanda (beneficiarios). Estos son los aspectos que se someten a la evaluación de impacto para medir la efectividad.

Conforme los resultados se alejan de los buscados con el programa, más factores externos influyen sobre su calidad y magnitud. En una evaluación que busque conocer efectos causales del programa, se necesita ejercer control sobre estos factores externos pues se convertirán en posibles efectos causales alternativos al programa.

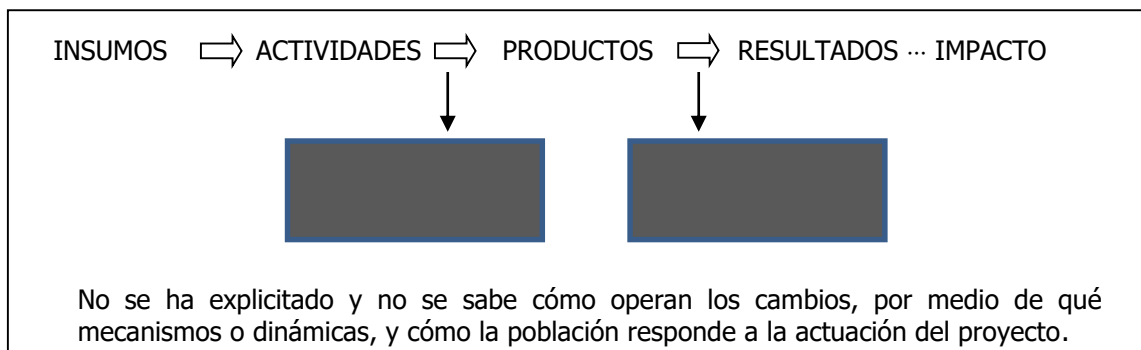
Las cadenas de cambio facilitan el debate en torno a la evaluación al poner de manifiesto qué información es necesario recoger y qué cambios de los resultados hay que tener en cuenta cuando se evalúa el proyecto.

En la etapa del diseño de la evaluación corresponde identificar claramente los elementos de la cadena de resultados del PPP, los cuales a su vez deberán estar correctamente definidos a fin que se pueda desarrollar la verificación y/o medición en el proceso de evaluación.

En la práctica actual lo usual es que en el diseño y planeamiento de las intervenciones se aplique la metodología del Marco Lógico, con una cadena de cambios en la que no se explicitan las condiciones y mecanismos para que la cadena de causalidad opere efectivamente; es decir, lo que debe ocurrir para que con la actividad se obtenga el producto previsto, e igualmente para que del producto obtenido se sigan los resultados esperados.

Cuando no se explicitan estas condiciones y mecanismos se habla de una "caja negra" (Figura 3); asumimos que hay un proceso dinámico que tiene que ocurrir pero no sabemos en qué consiste este proceso y sí se está dando en la realidad en el proceso de ejecución del PPP.

Figura 3. La caja negra



El caso de un proyecto sobre desnutrición en Bangladesh, permite ilustrar la aplicación de la teoría de cambio⁶. La argumentación planteada en la formulación del proyecto es que la desnutrición se debe a la ignorancia más que a la pobreza, lo cual es sustentado en los datos que muestran que aún en los niños del quintil más rico hay desnutrición. En la población existe la creencia que durante el embarazo las mujeres deben disminuir su alimentación, lo que afecta el desarrollo del feto y hace que los niños nazcan con un bajo peso.

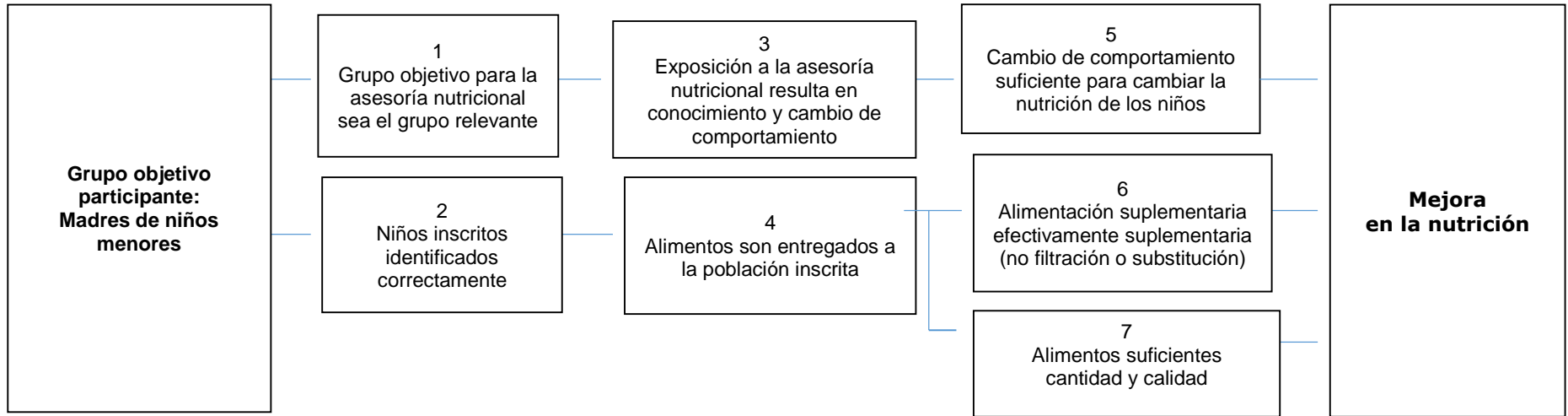
El grupo objetivo estuvo conformado por madres con hijos con desnutrición, y mujeres embarazadas. La estrategia principal de intervención fue la asesoría y capacitación, y de manera complementaria por el suplemento de alimentos. La selección de madres-hogares se realizó en función a la medición de peso de los niños y su ubicación en las curvas de desarrollo. La medición y selección estuvo a cargo de promotoras comunitarias entrenadas por el programa, y un monitoreo semanal.

El proyecto no tuvo los efectos deseados, lo cual se debió a la falta de atención a nexos débiles y ausentes en la cadena causal.

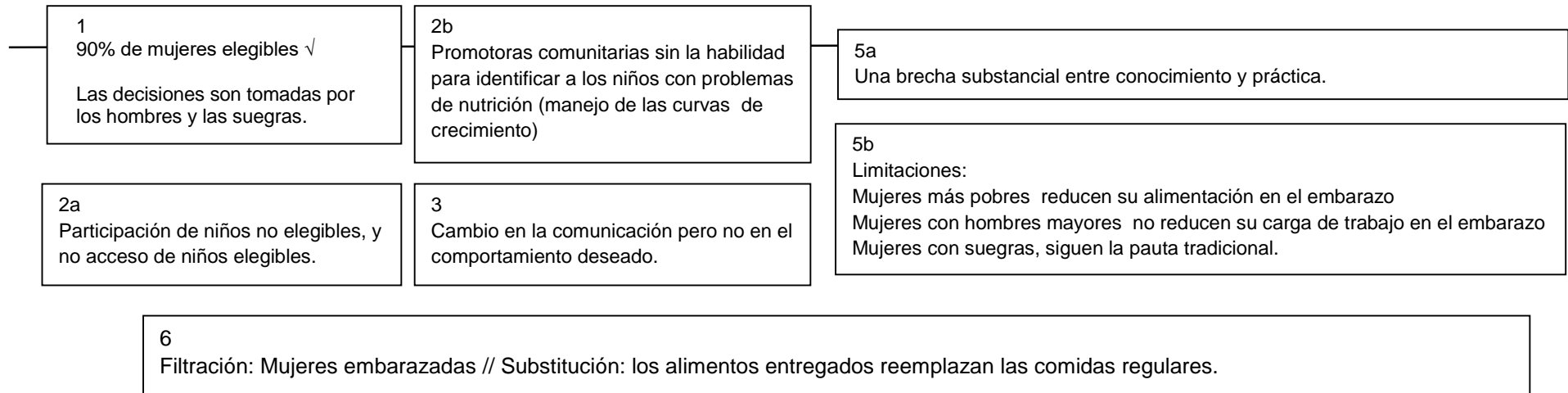
Los aprendizajes resultantes de la evaluación estuvieron referidos, principalmente, a los nexos de la cadena de causalidad: Incluir a suegras y esposos, afinar las poblaciones objetivo, y mejorar la selección y entrenamiento de las promotoras comunitarias. En la Figura 4, incluida en la siguiente página, se muestra la cadena causal del proyecto y los principales puntos de observación resultantes del análisis de la teoría de cambio.

⁶ Theory-based Impact Evaluation: Principles and Practice. Howard White (2011) Executive Director, International Initiative on Impact Evaluation, 3ie

Figura 4. Proyecto sobre desnutrición infantil: cadena de causalidad, condiciones de efectividad



Observaciones a la cadena causal, anotadas siguiendo la numeración correspondiente



Para ampliar el tema tratado, realiza las siguientes lecturas:

- Logic Model – Kellogg Foundation. 2004. p. 1-26
- ZIGLA Consultores. Inversión social estratégica. Teoría de cambio. Ejemplo de aplicación. Sin fecha. Disponible en:
<http://www.mapa.ziglaconsultores.com/content/teoria-del-cambio-ejemplo-de-aplicacion>
- Retolaza, I. Teoría de Cambio. Un enfoque de pensamiento-acción para navegar en la complejidad de los procesos de cambio social. PNUD. Hivos. Guatemala. 2010. Disponible en:
<https://planificacionsocialunsj.files.wordpress.com/2011/09/teorc3ada-del-cambio-pensamiento-accion-complejidad-cambio-social.pdf>

2.4 Identificación de los interesados

La identificación o análisis de interesados es una técnica que se utiliza para evaluar la importancia de las personas clave, grupos o instituciones en la evaluación. Permite establecer de manera sistemática los intereses y necesidades de los interesados y genera información que es crítica para la planificación e implementación de su evaluación.

El análisis de los interesados ayuda a anticipar la influencia que los diferentes grupos de interés pueden tener en la movilización de apoyo o resistencia a la evaluación, así como la utilización de los resultados de evaluación y aplicación de las recomendaciones. También ayuda a proporcionar una base y una estrategia para su participación en la evaluación, que a menudo conduce a exámenes más útiles y rentables.

2.4.1 ¿Quiénes son los interesados?

Los interesados son todas aquellas personas y organizaciones que son afectados por el PPP; entre ellos se encuentran por ejemplo: la población de beneficiarios organizada, las organizaciones de base de la comunidad, expertos interesados en la temática, entidades del gobierno (nacional, regional y local), donantes, entre otros.

Las partes interesadas en la evaluación varían significativamente entre los diferentes programas/proyectos. Asimismo, las partes interesadas y el grado en el que estas están involucradas varía dependiendo del método de evaluación. Por ejemplo, una evaluación basada en la comunidad de un nuevo programa escolar podrá invitar a las partes interesadas incluso a ser parte del propio proceso de planificación de la evaluación (Holden, D., Zimmerman M.; 2010).

2.4.2 ¿Cómo realizarlo?

La realización de un análisis de los interesados en la evaluación es un proceso escalonado. En tanto, de manera general se puede plantear los siguientes cuatro pasos y algunas pautas:

a) Identificar e invitar a las partes interesadas

Primero, se debe identificar a los principales usuarios de la evaluación: las instituciones que financian el PPP, los patrocinadores de la evaluación, los tomadores de decisiones o gerentes del PPP, los gestores del programa y el personal en campo y/o los beneficiarios del programa. Se debe consultar a

los usuarios qué es lo que realmente necesitan conocer sobre el programa y cómo se imaginan utilizando los resultados de la evaluación. Este paso ayuda a identificar la información que se necesita incluir en la evaluación y cómo conseguirla.

Segundo, identificar los clientes y otros interesados clave. Este grupo es probablemente mayor que el grupo que identificó en el paso anterior. Un grupo más amplio de interesados son aquellos que pueden estar interesados en y/o ser afectados por la evaluación, tanto positiva como negativamente. Se debe asegurar la inclusión de los grupos marginados de ser el caso. Este grupo de actores tiene intereses importantes, pero tendrá menos voz en la evaluación.

Pautas:

- Se debe asegurar que cada uno de estos grupos pueda plantear de manera oportuna temas y preguntas que se puedan abordar en la evaluación.
- En escenarios en que el PPP en estudio tenga un alto componente político, es fundamental involucrar a los interesados de fuera del programa en las primeras etapas de planificación para garantizar la utilidad de los resultados de la evaluación. En esta situación es importante entrevistar a las partes interesadas, como los líderes comunitarios y funcionarios electos para obtener su opinión sobre los aspectos importantes de un programa en particular.

b) Evaluar los intereses de las partes

Se debe analizar los intereses, el nivel de influencia, la importancia y el potencial impacto de las partes interesadas sobre la evaluación. Para desarrollar este paso se puede utilizar la siguiente matriz que permitirá determinar en qué medida cada individuo o grupo podrían desempeñar un papel en la planificación, implementación o uso de la evaluación. Al hacerlo, se deben considerar las relaciones entre los actores, puntos de acuerdo y conflicto entre ellos y las implicancias de corto y largo plazo de estas sobre la evaluación.

Pautas:

- Considerando que incluir la participación de los interesados puede requerir una gran cantidad de tiempo y recursos, dos estrategias para asegurarse de que el tiempo es bien invertido pueden ser:
 - (iii) el definir de forma proactiva los roles y las expectativas de las partes interesadas antes de participar en la planificación de la evaluación y,
 - (ii) el establecer un proceso de participación permanente en evaluación.

c) Desarrollar estrategias para involucrar adecuadamente a los interesados

Las estrategias utilizadas deben ser adecuadas para el contexto y las necesidades de la evaluación. Incluso si los grupos o individuos específicos no están involucrados de manera explícita, sigue siendo importante que sus intereses y necesidades se hayan tenido en cuenta.

Existen cuatro principios específicos para determinar el grado de participación de las partes interesadas:

- i. Comprender, respetar y tener en cuenta las diferencias entre las partes interesadas, tales como intereses, cultura, religión, discapacidad, edad, orientación sexual, origen étnico, etc.
- ii. Incluir las perspectivas e intereses de todas las partes interesadas.
- iii. Permitir a los interesados el acceso a la información de la evaluación.
- iv. Mantener un equilibrio entre el patrocinador de la evaluación y las necesidades e intereses de las partes interesadas.

Pautas:

- Una herramienta eficaz cuando se invita a participar a los interesados en el proceso de evaluación es proporcionarles una descripción escrita de sus roles y captar sus expectativas. En general, las personas aprecian conocer el nivel de compromiso y el tiempo que se les solicita, cuáles serían las expectativas de su participación y cómo su participación puede beneficiar al programa y a ellos mismos.
- Se debe considerar ofrecer actualizaciones o información periódica a las partes interesadas durante el proceso de la evaluación. Esto ayuda a

mejorar la aceptación y el uso eventual de los resultados de la evaluación.

d) Asegurarse de que el informe de evaluación indique la naturaleza y el alcance de la participación de las partes interesadas

Para cada parte interesada debe estar expresado su rol, sus aportes, sus posibles reacciones ante hallazgos, su interés y también su identificación como opuesto a la evaluación. El cuadro 10 permite sistematizar toda la información.

Cuadro 10. Matriz de análisis de interesados

(1) Partes interesadas	(2) Posición	(3) Interés en la evaluación	(4) Posible reacción ante hallazgos
Identificar el nombre de todos los interesados (afectado) por la evaluación. En la práctica son los afectados o interesados por el PPP. Pueden ser el personal del proyecto, los beneficiarios, los financiadores, los hacedores de políticas, funcionarios de instituciones reguladoras, expertos, miembros de la academia, etc	Definir cuál es el rol que tienen en el diseño de la evaluación: normativo, tomador de decisiones o financiador o beneficiario, etc	Identificar el interés que pueden tener en la evaluación.	Identificar cuáles pueden ser las respuestas de cada interesado ante los hallazgos de la evaluación

Ejemplo: El MIMP financia un nuevo modelo de atención a las mujeres que sufren violencia y decide evaluar su efectividad

Parte interesada	Posición	Interés en la evaluación	Posible reacción ante hallazgos
Ministerio de la Mujer y Poblaciones Vulnerables (MIMP)	Rol normativo, financia el proyecto sobre un nuevo modelo de atención a las mujeres que sufren violencia, supervisará todo el proceso de evaluación	Necesita saber si el nuevo modelo de atención a las mujeres que sufren violencia es efectivo.	Si la evaluación es positiva promoverá que el nuevo modelo de atención se generalice en todo el país.

2.5 Estudio y verificación de la evaluabilidad del PPP

Mediante este paso se busca determinar si la intervención reúne las condiciones para ser evaluada, es decir, que es viable hacer la evaluación, y si estamos en su inicio asegurar que tales condiciones se cumplan para su evaluación, que se pueda establecer, de manera fiable y creíble, en qué medida se han logrado los resultados previstos, se ha generado valor. (Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género 18)

El cuadro 11 muestra una Lista de Verificación de Evaluabilidad para Proyectos o Programas, la cual se divide en siete áreas clave. Para cada una de ellas se incluye un conjunto de características que ofrecen mayor especificación y claridad para el análisis. La evaluabilidad de un proyecto o programa dependerá de la presencia o ausencia de las características para cada área especificada en este instrumento. La calificación de evaluabilidad para cada área clave ha sido definida en escalas de cinco puntos (dos áreas) y diez puntos (cinco áreas), de pobre a excelente, con una calificación total máxima de 60 puntos.

Más allá de la calificación numérica resultante, es importante que se tenga una idea clara de las implicaciones o restricciones que la usencia o condiciones no favorables, encontradas en las áreas analizadas, puedan tener para la evaluabilidad; y, de lo que se pueda hacer para superar dichas limitaciones.

La intención es que esta lista de verificación sea utilizada por equipos de diseño de proyectos durante la preparación de un proyecto o programa, como un recordatorio de los aspectos clave a ser considerados para asegurar que la intervención sea evaluable. El instrumento también puede ser utilizado para analizar la evaluabilidad de proyectos o programas una vez completado el diseño; para ello será útil revisar también el Informe de Proyecto (IP), el Marco Lógico, el calendario de ejecución del proyecto, si existe, y el contrato de préstamo, si fuera el caso.

Cuadro 11. Lista de verificación de evaluabilidad para proyectos o programas

Criterios	Escala de calificación
<p>A. Diagnóstico</p> <p>El problema o necesidad abordado por el proyecto o programa ha sido claramente identificado por medio de consultas a los involucrados (prestatarios, entidades ejecutoras, beneficiarios, otras partes interesadas).</p> <ul style="list-style-type: none"> Las causas del problema han sido identificadas. Los beneficiarios potenciales están claramente identificados. <p>Los resultados esperados al final de la ejecución del proyecto o programa están claramente vinculados a los problemas y necesidades identificados en el diagnóstico.</p>	<p>0 No hay diagnóstico.</p> <p>1-4 <i>Regular</i>: diagnóstico poco claro.</p> <p>5-7 <i>Bueno</i>: diagnóstico relativamente claro.</p> <p>8-10 <i>Excelente</i>: los objetivos del proyecto o programa están claramente relacionados con el diagnóstico.</p>
<p>B. Lógica del proyecto</p> <ul style="list-style-type: none"> Los objetivos del proyecto o programa (Fin, Propósito, Componentes/Productos) están claramente expresados. Cada Componente/Producto es necesario (esencial) para lograr el Propósito. Tanto la Lógica Vertical (Componentes/Productos a Propósito) como la Lógica Horizontal (Supuestos) reflejan las condiciones necesarias y suficientes para que la intervención tenga éxito. 	<p>0 Falta de lógica o no está clara.</p> <p>1-2 <i>Regular</i>: la lógica es poco clara y no parece que todos los Componentes sean esenciales para lograr el Propósito, o los Supuestos no reflejan las condiciones "suficientes" para lograr el nivel siguiente de objetivos.</p> <p>3 <i>Buena</i>: lógica del proyecto relativamente clara.</p> <p>4-5 <i>Excelente</i>: la Lógica Vertical y Horizontal del proyecto reflejan las condiciones necesarias y suficientes para producir los Componentes (productos) y lograr el Propósito.</p>
<p>C. Supuestos y riesgos</p> <ul style="list-style-type: none"> Los Supuestos para la ejecución del proyecto o programa y el logro de sus objetivos han sido identificados, reflejan los riesgos mencionados en el documento de proyecto y pueden ser monitoreados 	<p>0 No están definidos los Supuestos y riesgos o no están claros.</p> <p>1-4 <i>Regular</i>: los Supuestos y riesgos han sido definidos parcialmente y colocados en su nivel apropiado.</p> <p>5-7 <i>Bueno</i>: los Supuestos y riesgos están claramente definidos y vinculados al análisis de involucrados.</p>

Criterios	Escala de calificación
<ul style="list-style-type: none"> Los Supuestos han sido colocados en el nivel apropiado, es decir, a nivel de Actividades para producir los Componentes (Productos) y a nivel de Componentes para lograr el Propósito. Los Supuestos están vinculados al análisis de involucrados y reflejan sus resultados. Las acciones de seguimiento para monitorear los Supuestos han sido identificadas, con objeto de aumentar su probabilidad de ocurrencia. 	<p>8-10 <i>Excelente</i>: los Supuestos y riesgos están claramente definidos y pueden ser monitoreados; se han identificado acciones de seguimiento para monitorear los Supuestos.</p>
<p>D. Indicadores de Componentes (productos)</p> <ul style="list-style-type: none"> Los indicadores de los Componentes (Productos) identifican medidas cuantitativas o cualitativas de los bienes y servicios que se espera serán producidos por medio de la ejecución del proyecto. Los indicadores de los Componentes (Productos) han sido acordados con los involucrados principales. Los indicadores de los Componentes (Productos) reflejan y especifican claramente las metas durante (es decir, metas intermedias – trimestrales, semestrales, anuales) y al final de la ejecución del proyecto. Los indicadores de los Componentes (Productos) son específicos, medibles, realizables, relevantes, enmarcados en el tiempo e independientes. 	<p>0 Falta de definición de todos los indicadores.</p> <p>1-3 <i>Regular</i>: la mayoría de los indicadores no están bien especificados.</p> <p>4-7 <i>Buena</i>: la mayoría de los indicadores están bien especificados, incluyen metas intermedias y fueron acordados con los involucrados.</p> <p>8-10 <i>Excelente</i>: al menos un indicador para cada Componente (Producto) cumple todos los criterios especificados.</p>
<p>E. Indicadores de Propósito (efecto directo)</p> <ul style="list-style-type: none"> Los indicadores del Propósito identifican medidas cuantitativas o cualitativas de los efectos directos esperados (beneficios para los clientes o beneficiarios) al final de la ejecución del proyecto. Los indicadores del Propósito especifican claramente las metas esperadas durante y al final de la ejecución del proyecto. Los indicadores del Propósito son específicos, medibles, realizables, relevantes, enmarcados en el tiempo e independientes. Al menos un indicador a nivel de Propósito es un indicador “de avanzada”, es decir que se refiere a un efecto directo “próximo” o intermedio. 	<p>0 Falta de definición de todos los indicadores.</p> <p>1-4 <i>Regular</i>: la mayoría de los indicadores no están bien especificados.</p> <p>5-7 <i>Buena</i>: la mayoría de los indicadores están bien especificados, incluyen metas cualitativas y fueron acordados con los involucrados; al menos un indicador a nivel de Propósito cumple todos los criterios especificados.</p>

Criterios	Escala de calificación
	8-10 <i>Excelente</i> : hay al menos un indicador “de avanzada” a nivel de Propósito.
<p>F. Datos de línea de base para objetivos</p> <ul style="list-style-type: none"> • Los datos de línea de base para los indicadores del Propósito se presentan en el documento de proyecto, junto con las series históricas correspondientes, cuando es apropiado. • Los datos de línea de base para los indicadores de los Componentes (Productos) se presentan en el documento de proyecto, junto con las series históricas correspondientes, cuando es apropiado. 	<p>0 No hay especificación de datos de línea de base o es pobre</p> <p>1-3 <i>Regular</i>: no se proporcionan datos de línea de base; el proyecto podría incluir provisión para desarrollar datos de línea de base como actividad inicial.</p> <p>4-5 <i>Excelente</i>: los datos de línea de base (a nivel de Componentes/Productos y Propósito) están especificados en el documento de proyecto junto con las series históricas correspondientes cuando es apropiado y serán revisados al comienzo de la ejecución del proyecto.</p>
<p>G. Plan de Monitoreo y Evaluación</p> <ul style="list-style-type: none"> • Los prestatarios y el banco han definido un sistema de recolección de información para generar información sobre indicadores. • Hay un Plan de Monitoreo y Evaluación que tiene financiamiento adecuado del proyecto o programa (los recursos han sido identificados y asignados para asegurar que la información identificada será recolectada y analizada). • Se han hecho arreglos pertinentes para utilizar la información para el monitoreo del proyecto o programa. • Las fuentes de información están especificadas para todos los indicadores. • El Plan de Monitoreo y Evaluación incluye la divulgación de información a los involucrados. • Se espera que los involucrados, incluyendo los beneficiarios, participen en el Monitoreo y Evaluación. 	<p>0 Falta de, o pobre, definición de un sistema de Monitoreo y Evaluación: los recursos asignados para M&E son inadecuados.</p> <p>1-4 <i>Regular</i>: el sistema de M&E está solo parcialmente definido; posiblemente tenga financiamiento adecuado.</p> <p>5-7 <i>Buena</i>: el sistema de M&E está bien definido y financiado íntegramente.</p> <p>8-10 <i>Excelente</i>: el sistema de M&E está bien definido, incluyendo la participación de los involucrados y la divulgación oportuna de información a los beneficiarios y otros involucrados.</p>

Fuente: Materiales del curso a distancia Monitoreo y Evaluación de Proyectos Octubre 2001. Banco Interamericano de Desarrollo (BID)

2.6 Definición de las preguntas de evaluación e hipótesis

Luego de haber identificado la lógica causal de la intervención a ser evaluada, y haber realizado el análisis de los interesados, el siguiente paso que abordamos es definir las preguntas de evaluación, las cuales, a su vez, nos permitirán definir claramente los objetivos de evaluación.

Como se vio en la Unidad 1, debemos recordar que cada evaluación responde a diferentes preguntas que se derivan de: a) Las necesidades de información que tiene el programa o sus interesados y/o b) los tipos o enfoques de evaluación.

Responder a las preguntas requiere identificar y elegir los métodos a utilizarse y conocer cuánto costaría responderlas. De manera general, las evaluaciones pueden responder a tres tipos de preguntas, las cuales a su vez son complementarias (Bertrand, J., Manuel S.; 2000):

- Preguntas descriptivas. ***¿Cómo está funcionando el programa/proyecto?*** La evaluación intenta determinar qué está ocurriendo y describe los procesos, condiciones, relaciones organizativas y opiniones de las partes interesadas – Evaluación de desempeño.
- Preguntas normativas. ***¿Se han logrado los objetivos planteados por el programa/proyecto?*** La evaluación compara lo que está ocurriendo con lo que debería ocurrir; evalúa si se logran o no los objetivos, por qué y cómo – Evaluación de resultados.
- Preguntas sobre causa y efecto. ***¿Se logró o no el cambio deseado? (Si se hubiera logrado) ¿En qué medida se puede atribuir el cambio al programa/proyecto?*** La evaluación examina los resultados finales e intenta determinar la diferencia que ejerce la intervención sobre estos – Evaluación de impacto.

Cuando se trata de hacer una evaluación causal o conocer posibles explicaciones acerca de un estado del programa es posible formular hipótesis. En la evaluación de impacto la hipótesis es que el programa es el causante del cambio del indicador de impacto. Todo el diseño de la evaluación busca verificar si esa hipótesis tiene alguna evidencia para no rechazarla.

Otras hipótesis que se pueden formular en evaluación pueden referirse a identificar los factores que afectan algún aspecto del desempeño del programa. Por ejemplo en un programa de fortalecimiento de capacidades de promotores comunitarios de

planificación familiar se puede formular la hipótesis de que su desempeño está afectado tanto por el conocimiento que tienen acerca de métodos anticonceptivos como por el grado de apoyo que reciben de la comunidad o por el material educativo que disponen o por el grado de supervisión que reciben del programa.

IMPORTANTE

Se debe resaltar que cada evaluación es en cierto modo un caso único por su especificidad y, por lo tanto, las diversas clasificaciones que existen acerca de los tipos de evaluación, dimensiones de evaluación y preguntas de evaluación que se pueden hacer respecto a ellos son simplificaciones que un evaluador debe tomar de manera referencial.

2.6.1 Preguntas guía según dimensiones de evaluación

En evaluación existen cinco dimensiones o criterios que se analizan para dar cuenta de aspectos clave sobre el desempeño y logros del programa, estos son: relevancia, eficiencia, efectividad, impacto y sostenibilidad.

Toda evaluación de un plan, programa o proyecto requiere definir cuáles de estas cinco dimensiones serán analizadas durante la evaluación. Así, luego de haber identificado las dimensiones a evaluarse, el siguiente paso es identificar las preguntas de cada dimensión que son relevantes para responderse con la evaluación. Para revisar a detalle ver el cuadro 12.

Cuadro 12. Preguntas guía según dimensiones de evaluación

DIMENSIONES	DEFINICIÓN
1. Relevancia del programa	Qué tan bien el programa responde a las causas relevantes del problema identificado.
2. Eficiencia del programa	Es la medida en la cual el programa ha usado los menores recursos posibles para lograr la calidad requerida de los productos/resultados.
3. Efectividad del programa	Una medición de efectividad analiza si es que el programa ha logrado lo que se propuso lograr en sus planes.
4. Impacto de programa	El análisis sistemático de los cambios en la vida de las personas producidos por el programa.
5. Sostenibilidad	La medida en la cual las metas del programa continuarán siendo cumplidas después de que el programa termine.

Fuente: Plan Internacional. Lineamientos sobre la evaluación de programa de Plan Internacional.

a) Relevancia

Cuando se realiza una evaluación de un PPP la primera dimensión sobre la cual se necesita trabajar es la relevancia del programa:

¿Qué tan bien el programa responde a las causas relevantes del problema identificado?

Para analizar esta pregunta se necesita seguir dos pasos:

- i. Analizar qué ha sucedido durante el periodo del ciclo de vida de los programas y de esta manera desarrollar un contexto histórico específicamente concerniente a aquellos problemas para los cuales el programa fue originalmente diseñado a resolver, y ver si estos problemas son o no todavía dominantes al momento de haber completado el programa.
- ii. Luego, se debe explorar el grado de ajuste entre el programa y este contexto histórico –en otras palabras, el evaluador necesitará analizar hasta qué punto los problemas están siendo abordados por el programa y cómo el programa se ha ajustado para calzar con el contexto cambiante.

Preguntas asociadas:

¿El modelo causal del programa reúne efectivamente las principales o más relevantes causas del problema?

¿La intervención que supone el programa se hace a nivel de las causas del problema señaladas en el diagnóstico?

¿El Fin del programa da cuenta de una solución o mitigación del problema o necesidad detectada en el diagnóstico?

¿Las actividades del programa son adecuadas al perfil de los beneficiarios y la localidad en que habitan?

b) Eficiencia

La eficiencia consiste en lograr las metas, en términos de cantidad y calidad, con el menor uso de recursos posible. El análisis de la eficiencia presta atención entonces a los recursos gastados para alcanzar las metas.

Los evaluadores tienen una definición específica de eficiencia como la medición de:

¿Hasta qué punto el programa ha utilizado los menores recursos posibles para lograr la calidad requerida de los productos/resultados?

Para analizar la eficiencia de un PPP, el punto de partida viene de las comparaciones, tanto cotejando lo que realmente sucedió con lo que estaba planificado como haciendo paralelos entre productos similares producidos (a menudo empezando con una comparación de costos). Sin embargo, para hacer dichas comparaciones es importante trabajar con información comparable.

La comparación le permitirá decir si un PPP es eficiente con relación a otro similar. A través de dichas comparaciones, se pueden identificar razones para las variaciones en la eficiencia y se pueden reportar aprendizajes que se pongan en práctica en futuros proyectos.

Preguntas asociadas:

¿El costo estimado del programa por beneficiario o producto son razonables o existen formas alternativas de menor costo para lograr los mismos productos?

¿Los productos propuestos al costo estimado se pueden lograr en los plazos programados?

¿En qué medida el modelo de gestión favorece el uso eficiente de los recursos?

c) Efectividad

Es la relación entre lo logrado con lo planificado.

Una evaluación de la efectividad analiza si *el programa ha logrado lo que propuso en sus planes.*

La medición de la efectividad del programa necesita pensamiento sistemático porque un PPP comprende un amplio rango de metas específicas en los diferentes niveles de implementación, las que variarán desde el logro de una contribución específica a la meta del programa, hasta el de la producción de un número de productos. Para medir si un programa ha sido “efectivo” es necesario dividir el plan original y evaluar si cada una de las metas planificadas ha sido lograda o no.

Preguntas asociadas:

¿En qué medida el modelo de gestión permite alcanzar adecuadamente los objetivos del programa?

¿En qué medida las actividades programadas garantizan el logro de los productos en el plazo estimado? (evidencia).

¿En qué medida los recursos humanos requeridos que implementará el programa dan garantías de cumplimiento de los productos?

d) Impacto

Una medición del impacto del programa puede ser definida como un análisis sistemático de los cambios en la vida de las personas beneficiarias (niños, familias y comunidades) en el largo plazo que pueden ser atribuidos a la intervención del PPP.

Considerando esta definición básica, un análisis del impacto del programa necesita observar tres áreas clave:

- Los cambios positivos en la vida de las personas.
- Los cambios negativos en la vida de las personas.
- El grado en el cual es posible concluir que estos cambios en la vida de las personas pueden ser atribuibles al programa.

La última pregunta, conocida como **“la pregunta de atribución”**, es lo que hace a la evaluación del impacto tan difícil. Esto se debe a que la evaluación de un cambio debido a un programa necesita tomar en cuenta que hay muchos otros factores que producen cambios en la vida de las personas, y cuyos efectos deben ser aislados.

IMPORTANTE

Esta dimensión a su vez se constituye en un tipo de evaluación en sí misma, por lo tanto, en la etapa del diseño de evaluación lo que se debe definir es si será necesario realizar la evaluación de impacto del programa considerando que este tipo de evaluaciones son de alto costo y requieren un alto nivel de experticia técnica.

e) Sostenibilidad

La sostenibilidad puede ser definida como la medida en la cual los objetivos del programa continuarán cumpliéndose después de que termine el programa. Este tema debe ser manejado como parte de la evaluación de programa y responde a la pregunta:

¿Hasta qué punto continuará el impacto del programa después de que este haya finalizado?

La sostenibilidad del programa en este contexto no debe ser confundida con "sostenibilidad ambiental", que se enfoca solamente en el impacto ambiental de los programas y es un tipo especial de evaluación de sostenibilidad.

Preguntas asociadas:

¿Las características del bien o servicio que el PPP pretende otorgar junto con las condiciones externas (supuestos) permiten prever una duración prolongada del Propósito y Fin del programa?

¿El nivel de participación de los beneficiarios le da sostenibilidad al programa?

¿La institucionalidad pública en que se inserta el programa favorece la mantención de beneficios en aquellas personas o familias que egresan del programa?

2.6.2 Énfasis de las preguntas según tipos de evaluación

Como se vio en el ítem anterior, el evaluador debe identificar las dimensiones y preguntas que interesa que sean analizadas. Sin embargo, en esta parte debemos indicar que cada tipo de evaluación puede tener énfasis en el análisis de ciertas dimensiones y, en otros casos, considerar preguntas específicas.

A continuación se presentan las principales preguntas que se responden con cada uno de los tipos de evaluación y se indica qué dimensiones son frecuentemente utilizadas en cada uno de estos. De otro lado, como se podrá identificar a través de esta sección, las preguntas de los diferentes tipos de evaluaciones son complementarias entre sí, por lo cual es posible diseñar una evaluación que responda a más de un tipo de preguntas en los denominados diseños mixtos de evaluación.

a) Evaluación *ex ante*

Dado que la evaluación *ex ante* se realiza sobre una propuesta de PPP antes de que haya sido implementado, las siguientes preguntas permiten indagar sobre aspectos básicos que deben estar adecuadamente sustentados en cualquier propuesta de PPP.

¿Están claramente definidos los objetivos del programa/proyecto? ¿Es posible medir el cumplimiento de los objetivos? ¿Las metas son posibles de cumplirse?

¿Se presenta en el diagnóstico un modelo explicativo que identifique causas y efectos del problema y los indicadores que respaldan dicha relación?

¿Qué población objetivo se atenderá?

¿Qué bienes y/o servicios serán provistos? (Entregables claramente definidos).

¿Se identifican las áreas geográficas en las que se realizará la intervención?
¿Se entrega criterios que justifiquen la priorización que se ha hecho de algunos territorios en relación a otros?

¿Cuál es el mejor sistema de entrega de los servicios? (Análisis de alternativas).

¿Cómo será organizado el programa/proyecto?

¿Qué recursos son necesarios y apropiados para el programa/proyecto?

¿La propuesta programática es coherente con políticas específicas y generales que orienta la acción pública?

Dimensiones:

En la evaluación *ex ante* es usual analizar las dimensiones de relevancia, eficiencia y sostenibilidad en la medida que la propuesta de PPP brinde información que permita determinar si las acciones planteadas llevarán al resultado esperado al menor costo posible.

Finalmente, en una evaluación *ex ante* se debe incluir un análisis sobre la evaluabilidad del PPP en los siguientes términos:

Preguntas asociadas:

¿La claridad con que han sido expuestos los objetivos facilitará la evaluación futura del programa?

¿La operacionalización de los productos y actividades facilitará el monitoreo del programa? (calidad de los indicadores y validez de las metas) ¿La operacionalización de los objetivos facilitará la evaluación *ex post* del programa?

¿El sistema de evaluación *ex post* propuesto es adecuado a los objetivos del programa y perfil de los beneficiarios?

b) Evaluación de proceso

De manera general, este tipo de evaluación se enfoca en la implementación del PPP y responde a las siguientes preguntas:

¿Cuánto hemos hecho? (durante un periodo de tiempo específico: un mes, un trimestre, un año)

Ejemplo:

- Número de eventos realizados (charlas, talleres, grupos de apoyo, etc.).
- Número de participantes en el evento o actividad.
- Número de productos o materiales diseñados / producidos / distribuidos.

¿Cuán bien lo hemos hecho?

Ejemplo:

- Grado de satisfacción con el servicio o actividad, medido por múltiples aspectos (indicadores ilustrativos):
 - La información fue entendible.
 - La información fue suficiente.
 - Pudo participar (hacer preguntas, etc.).
 - El trato por parte del personal fue satisfactorio.

¿Cómo podríamos mejorar?

Enfocándose siempre en la implementación de las actividades de un programa/proyecto.

c) Evaluación de resultados

La evaluación de resultados se enfoca principalmente en verificar si los productos planificados fueron obtenidos en los tiempos y costos previstos.

- ¿Han sido alcanzados los productos y resultados trazados? (Contrastar lo planificado versus ejecutado).
- ¿Fueron entregados los servicios a la población objetivo?
- ¿Está satisfecha la población con el servicio? Puede evaluarse esto a través de una encuesta de percepciones.

Dimensiones:

Una evaluación de resultados por lo general se enfoca en el análisis de las dimensiones de eficacia y eficiencia.

d. Evaluación de impactos

A diferencia de otros tipos de evaluación, que pueden responder a muchas preguntas, las evaluaciones de impacto se preocupan por conocer cuál es el impacto de un PPP sobre un resultado de interés, única dimensión analizada. La pregunta fundamental en este tipo de evaluación es:

¿Cuál es el impacto (o efecto causal) de un programa sobre un resultado de interés?

Por ejemplo: ¿cuál es el impacto sobre el acceso a la atención a la salud de la contratación de la atención primaria a proveedores privados? ¿La mejora de las carreteras y el acceso a mercados laborales aumenta el ingreso de los hogares? En caso afirmativo, ¿cuánto?

A partir de la formulación de la pregunta central de interés pueden ensayarse algunas hipótesis. En este caso, como ejemplo, formulamos las hipótesis a partir de la evaluación de un programa que busca elevar las tasas de finalización de la secundaria aplicando un nuevo programa de matemáticas en alumnos de secundaria; para ello se llevan a cabo dos actividades: la capacitación de maestros en el nuevo programa y la entrega de textos a los

alumnos. A partir dicha Teoría del Cambio, se pueden plantear las siguientes hipótesis:

- Si se capacita a los profesores y se distribuyen los libros de texto, los resultados de los exámenes de matemáticas mejorarán en promedio en 5 puntos.
- El desempeño en los estudios de matemáticas en secundaria influye significativamente en las tasas de finalización y el desempeño en el mercado laboral (Gertler P., et. al; 2011).

Para ampliar el tema tratado, revisa la siguiente lectura:

- La evaluación de impacto en la práctica – Banco Mundial. Primera parte: Capítulo 2. 2011. p. 22-30.

2.7 Elección de métodos: cuantitativos, cualitativos y mixtos.

Las evaluaciones de PPP usualmente se valen de los métodos de investigación utilizados por las ciencias sociales para analizar un fenómeno. Dichos tipos de métodos, por su naturaleza y objetivo, se dividen entre cualitativos y cuantitativos (Hernández, R., et. al; 2008). En los siguientes ítems se discutirá a detalle la naturaleza, tipos y principales usos de dichos métodos, así como los criterios que deben considerarse cuando debemos realizar un diseño mixto que combine tanto métodos cuantitativos como cualitativos. Sin embargo, las evaluaciones requieren también del uso de métodos propios de la temática o campo en el que se evalúa. Por ejemplo en una evaluación nutricional se requiere usar indicadores como los del estado nutricional, cuya medición se debe hacer con las técnicas y procedimientos propios de las ciencias nutricionales.

Asimismo, cabe indicar que debido a la diversidad de tipos de PPP, preguntas de evaluación, disponibilidad de datos, restricciones de tiempo y circunstancias del país, cada estudio de evaluación será diferente y requerirá cierta combinación de las metodologías adecuadas, tanto cuantitativas como cualitativas. Por lo tanto, el evaluador debe examinar con cuidado las opciones metodológicas al diseñar el estudio, con el objetivo de producir los resultados más sólidos posibles.

2.7.1 Las metodologías cuantitativas

Para describir diferentes aspectos de la intervención en términos numéricos y cuantificables (dimensionan el fenómeno), se utilizan las metodologías cuantitativas.

Por ejemplo: el número de regiones intervenidas por el PPP, el porcentaje de la población beneficiaria, la tasa de mortalidad infantil, la variación del flujo de la inversión, el ingreso promedio de los habitantes de una determinada región, entre otros. En este caso el evaluador define lo que considera importante a cuantificar y canaliza sus esfuerzos en obtener información robusta y suficiente para el cálculo de los indicadores cuantitativos que se hayan identificado a partir de la cadena de resultados e información cuantitativa adicional que se requiera para la evaluación.

Las metodologías cuantitativas que se usan frecuentemente en la evaluación de los diversos tipos de intervenciones son: (i) recopilación de información estadística, (ii) métodos experimentales y (iii) métodos cuasi experimentales; los dos últimos se aplican fundamentalmente para las evaluaciones de impacto. En la metodología cuantitativa usualmente recopilan datos usando diversas técnicas específicas como encuestas con entrevistas estructuradas, lo cual se estudia a detalle en las siguientes secciones.

2.7.2 Recopilación de información estadística:

En todas las evaluaciones se utiliza regularmente la recopilación de la información estadística existente que ha generado el programa durante su implementación. La información estadística de una determinada intervención puede provenir de:

- Hojas de registros del programa/proyecto.
- Archivos sobre actividades realizadas, servicios prestados y productos entregados.
- Base de datos de beneficiarios de la intervención.

Una de las grandes ventajas de la información estadística es que gran parte de esta se recolecta con cierta regularidad, como parte del manejo de la intervención, y no representa un costo ni esfuerzo adicional, sobre todo si los sistemas de información están diseñados tomando en cuenta los indicadores utilizados para la evaluación. Sin embargo, dentro de las desventajas puede decirse que:

- ⇒ Solo se puede obtener información sobre las variables existentes en el sistema.
- ⇒ En ciertas ocasiones la calidad de los datos no es fiable por el mismo proceso de captura de la información.

Es usual que la información de los registros del proyecto sea usada para las actividades de monitoreo del programa, por ejemplo en los proyectos de nutrición es usual que se evalúe la talla y el peso de los niños de manera periódica. Si las mediciones y registros se han ejecutado con prolijidad, esa información será muy útil para evaluar si el programa ha conseguido resultados. Se podría comparar si los niños que recibieron más tiempo el programa han conseguido mejores logros nutricionales que aquellos que estuvieron expuestos menos tiempo al programa. Asunto clave para el análisis de esta información de rutina es considerar i) los posibles sesgos que se pueda incurrir al usarla, por ejemplo, que haya autoselección para participar en el programa y ii) si los resultados del análisis son representativos de toda la población o una población en particular, por ejemplo, los niños que acuden al puesto de salud.

2.7.3 Las metodologías cualitativas

Por medio de las metodologías cualitativas se trata de captar la realidad de la intervención desde la perspectiva de las personas entrevistadas (beneficiarios, prestadores de servicio, directivos de la intervención, etc.) a fin de conocer a profundidad la dinámica, comportamientos y percepciones en torno a la intervención. Estas metodologías no tratan de cuantificar las características del proyecto, sin embargo, en ciertas ocasiones existe cuantificación de fenómenos observados a través de técnicas cualitativas.

Aunque existen ciertos parámetros al momento de la discusión u observación, las personas que participan en la evaluación tienen mucha más libertad de expresar sus ideas o describir una situación según su propio punto de vista. Las metodologías cualitativas más comunes en la evaluación de intervenciones son: i) grupos focales, ii) entrevistas a profundidad y iii) la observación.

El cuadro 13 sintetiza y compara las diversas características de ambos métodos.

Cuadro 13. Características de los métodos cuantitativos y cualitativos

	Métodos cuantitativos	Métodos cualitativos
Objetivo	Cuantifica y estima magnitudes.	Ofrece descripciones detalladas. Responde a "cómo" y "por qué".
Tipo de preguntas	Predominan preguntas cerradas (¿Usted trabaja?: Si, No).	Predominan preguntas abiertas (¿Cómo se siente con la intervención del programa?).
Información que proporciona	Proporciona información numérica y cifras que favorecen interpretaciones similares por distintos evaluadores.	Proporciona datos sobre percepciones, creencias, valores, etc., que pueden ser interpretados en forma diferente por distintos evaluadores.
Tamaño de muestra	Requieren muestras grandes basadas en algún método estadístico (métodos aleatorios).	Requiere muestras más limitadas, generalmente con selección dirigida a personas que cumplen ciertas características. Generalmente el tamaño está a discreción del investigador.
Especialistas	Generalmente se requiere personal con experiencia en métodos estadísticos sofisticados.	Requiere especialistas que analicen datos cualitativos.
Resultados	Los resultados pueden generalizarse al total de la población objetivo, si la muestra es representativa.	Los resultados no pueden generalizarse, son indicativos solo de las percepciones, creencias, valores, etc.

IMPORTANTE

En la literatura sobre evaluación, los métodos experimentales y los cuasi experimentales usualmente son aplicados para la estimación de los impactos de un programa/proyecto, sin embargo, cabe indicar que estos métodos pueden ser utilizados a otros niveles de la cadena de resultados a fin de estimar los denominados "efectos causales", es decir, los cambios observados que pueden ser atribuibles a la intervención y no a otros factores.

2.7.4 Uso de los Métodos cuantitativos⁷ en la Evaluación de impacto

Los métodos cuantitativos frecuentemente utilizados en el proceso de evaluación de impacto; pueden clasificarse en: (i) métodos experimentales (aleatorios) y ii) métodos cuasi experimentales (no aleatorios).

¿Cuál es la pregunta clave?

El objetivo de aplicar métodos experimentales o cuasiexperimentales en evaluación es estimar los efectos causales de la intervención, es decir, los cambios generados en algunos resultados del programa/proyecto debido únicamente a la intervención del PPP.

La mayoría de las preguntas de políticas conllevan relaciones de causa y efecto, por ejemplo: ¿la capacitación de los profesores mejora las calificaciones de los estudiantes en los exámenes? ¿Los programas de transferencias monetarias condicionadas mejoran la salud de los niños? ¿Los programas de formación profesional incrementan los ingresos de quienes los han cursado? Aunque las preguntas sobre causa y efecto son habituales, determinar que una relación es causal no es tan sencillo debido a que las intervenciones se desenvuelven en contextos donde una serie de factores concomitantes inciden sobre el mismo resultado.

La respuesta

La respuesta a la pregunta básica de inferencia causal *¿Cuál es el impacto o efecto causal de un programa P sobre un resultado de interés Y?* se obtiene mediante la siguiente fórmula básica:

$$\alpha = (Y | P = 1) - (Y | P = 0)$$

Según esta fórmula, el impacto causal (α) de un programa (P) sobre un resultado (Y) es la diferencia entre el resultado (Y) con el programa (es decir, cuando $P = 1$) y el mismo resultado (Y) sin el programa (es decir, cuando $P = 0$).

Por ejemplo, si P representa un programa de formación profesional e Y representa el ingreso, el impacto causal del programa de formación profesional (α) es la diferencia entre el ingreso de la persona (Y) cuando participa en el programa de formación

⁷ Gertler P., et .al (2011); Baker, July (2000).

profesional (es decir, cuando $P = 1$) y el ingreso de la misma persona (Y) en ese mismo momento si no hubiera participado en el programa (es decir, cuando $P = 0$).

Dicho de otro modo, sería necesario medir el ingreso en un momento para la misma unidad de observación (en este caso una persona), pero en dos realidades diferentes. Si esto fuera posible, se observaría el nivel de ingresos de una persona en un mismo momento, tanto con el programa como sin él, de manera que la única explicación posible de cualquier diferencia en su ingreso sería el programa y se eliminaría cualquier factor externo que pudiera explicar también la diferencia en los resultados. En este caso se podría confiar en que la relación entre el programa de formación profesional y el ingreso es causal. Sin embargo, esto presenta un problema contrafactual.

¿Qué es el contrafactual? y ¿cuál es el problema contrafactual?

Como se explicó anteriormente, el impacto (α) de un programa es la diferencia entre los resultados (Y) de la misma persona cuando ha participado y cuando no ha participado en el programa. Sin embargo, es imposible medir a la misma persona en dos situaciones diferentes en el mismo momento. Una persona habrá participado en un programa o no lo habrá hecho. No se la puede observar simultáneamente en dos estados diferentes. *Este es el "problema contrafactual": ¿cómo se mide lo que habría ocurrido si hubiera prevalecido la otra circunstancia?*

Aunque se puede observar y medir el resultado (Y) para los participantes en el programa ($Y | P = 1$), no se cuenta con datos para establecer cuál habría sido el resultado en ausencia del programa ($Y | P = 0$). En la fórmula básica, la expresión ($Y | P = 0$) representa el contrafactual, que puede entenderse como lo que habría pasado si un participante no hubiera participado en el programa. En otras palabras, *el contrafactual es lo que habría sido el resultado (Y) en ausencia del programa (P).*

¿Cómo estimar el contrafactual?

El primer paso consiste en construir el escenario contrafactual. En esencia se trata de seleccionar un grupo que no está siendo afectado por el programa, al que se denomina grupo de control, y contra el cual se compara el grupo intervenido (grupo de tratamiento o intervención). Para esto, desde antes de la implementación del PPP, se establece un grupo de control que debe estar formado por una población de similares características, en promedio, a la que va a ser intervenida. La comparación de las situaciones existentes en los dos grupos al momento de la evaluación con

respecto a sus situaciones en el momento previo al inicio del PPP dará la medida del efecto debido al PPP.

Para determinar dicho escenario contrafactual se pueden utilizar diversos métodos que entran en dos categorías generales que se mencionaron líneas arriba:

- Métodos experimentales (aleatorios).
- Métodos cuasi experimentales (no aleatorios).

A continuación se explica en qué consisten los principales métodos, así como sus fortalezas y debilidades. Como lo descubrirá el lector, ninguna técnica es perfecta y, por lo tanto, el evaluador debe tomar decisiones frente a las ventajas y desventajas de cada uno de los métodos elegidos.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Banco Mundial. La evaluación de impacto en la práctica. Segunda parte: Capítulos 3 y 4. 2011.

a. Métodos experimentales (aleatorios)

Los métodos experimentales de evaluación consisten en construir el escenario contrafactual de manera aleatoria. Es decir, se debe seleccionar al azar al grupo que recibió el proyecto y el grupo que no lo recibió.

La aleatorización permite que los grupos de control y de tratamiento sean semejantes en promedio en todas sus características, por tanto, la única diferencia entre ambos grupos será su participación en el programa (tratamiento o intervención). La asignación aleatoria puede hacerse a nivel de individuos, hogares, comunidades o regiones.

i. Uso

Este método es usado cuando se trata de probar la eficacia de las intervenciones antes de su implementación masiva. En la práctica, en la implementación de los PPP mayormente se definen los criterios de inclusión, las características que deben cumplir los que van a recibir la intervención del PPP y no se recurre a métodos aleatorios; sin embargo, en algunos proyectos que tienden a la cobertura total y no tienen los recursos suficientes para cubrir a toda la

población objetivo, se inicia la intervención con una selección aleatoria de los participantes.

¿Por qué la asignación aleatoria genera una estimación excelente del contrafactual?

El proceso de asignar aleatoriamente unidades a los grupos de tratamiento y de comparación producirá dos grupos estadísticamente iguales en cuanto a sus promedios, siempre que el número al que se aplique la asignación aleatoria sea lo suficientemente grande. A su vez, estos promedios también pueden generalizarse para la población de la que se extraen los participantes.

Por lo tanto, si después de poner en marcha el programa se observan diferencias en los resultados entre los grupos de tratamiento y de comparación, se sabrá que solo pueden deberse a la introducción del programa, ya que la selección de los dos grupos fue idéntica y ambos están expuestos a los mismos factores externos. En otras palabras, el grupo de comparación contiene todos los factores que podrían explicar también el resultado de interés.

¿Cuándo se puede usar la asignación aleatoria?

Existen normalmente dos situaciones en las que la asignación aleatoria es viable para la evaluación en el caso del grupo de intervención:

⇒ *Cuando la población elegible es superior al tamaño de la muestra determinado.* Entonces se hace una selección aleatoria.

⇒ *Cuando la población elegible es menor o igual al tamaño de la muestra determinado.* Entonces se toman todas las unidades elegibles.

En el caso del grupo de control, necesariamente se debe tomar una muestra aleatoria. Para asegurar esto, se debe escoger una población suficientemente grande como para tomar una muestra.

ii. Ventajas y limitaciones

En el diseño de la evaluación se debe tomar en cuenta:

- En programas de amplio alcance, es posible que no hayan grupos sin tratamiento, como en el caso de un cambio de política. Entre los ejemplos se incluye los préstamos de ajuste de la banca de desarrollo o programas aplicados a nivel nacional.

- Durante el experimento los individuos de los grupos de control podrían cambiar ciertas características que los identifican, lo que podría invalidar o contaminar los resultados. Por ejemplo, si las personas se trasladan hacia un área de proyecto o salen de él, podrían ingresar o salir del grupo de tratamiento o de control. Alternativamente, las personas a las que se les niega el beneficio del programa pueden buscarlo a través de otras fuentes, o bien aquellas a las que se les ofrece un programa podrían no aceptar la intervención. Pueden haber también casos de unidades que formaron parte del grupo de intervención al inicio y luego salieron, sea por decisión propia, o del PPP.
- Podría resultar difícil garantizar que la asignación sea realmente aleatoria. Un ejemplo de esto podrían constituirlo los administradores de préstamos que excluyen solicitantes de alto riesgo para lograr mejores resultados. Y por último, los diseños experimentales pueden ser costosos y prolongados en ciertas situaciones, especialmente en la recopilación de nuevos datos.

iii. **¿Cómo determinar aleatoriamente el tratamiento?**

Habiendo conocido los conceptos básicos sobre asignación aleatoria, a continuación se presentan los principales pasos que se deben seguir para obtener una asignación aleatoria exitosa:

Paso 1: Definir las unidades elegibles

Según el programa, esta puede ser una persona, un centro de salud, una escuela o incluso todo un pueblo o municipalidad. El conjunto de unidades elegibles está formado por aquellas acerca de las cuales interesa conocer el impacto de un programa. Por ejemplo, si se está implementando un programa de formación para los maestros de escuelas primarias en áreas rurales, los profesores de centros de secundaria o los maestros de escuelas de primaria de áreas urbanas no formarán parte del conjunto de unidades elegibles. A la relación de unidades elegibles se le denomina marco de muestreo o también marco muestral.

Paso 2: Determinar la muestra de evaluación

Una vez que se haya determinado la población de unidades elegibles, se debe comparar el tamaño del grupo con el número de observaciones requeridas para la evaluación. Si la población elegible es pequeña, puede ser necesario incluir

todas las unidades en la evaluación. Si hay más unidades elegibles que las requeridas para la evaluación, el siguiente paso consiste en seleccionar una muestra de la población. Nótese que este segundo paso se realiza principalmente para limitar los costos.

Si se observa que los datos de los sistemas de monitoreo existentes pueden emplearse para la evaluación y que estos sistemas abarcan al conjunto de unidades elegibles, no se necesitará extraer una muestra de evaluación diferente.

En nuestro ejemplo del programa de formación de docente, el conjunto de las unidades elegibles incluye a decenas de miles de profesores de todas las escuelas públicas del país, y se necesita recolectar información detallada acerca de los conocimientos pedagógicos de estos. Aunque entrevistar a todos ellos puede no ser viable, podría bastar con seleccionar una muestra de 1000 profesores distribuidos en 100 escuelas. En la medida en que la muestra de escuelas y profesores sea representativa de la población total de profesores de escuelas públicas, cualquier resultado de la evaluación puede generalizarse.

Paso 3: Asignación aleatoria

El tercer paso consistirá en formar los grupos de tratamiento y de comparación a partir de las unidades de la muestra de evaluación. Esto requiere definir un procedimiento aleatorio de selección de los participantes. La selección debe basarse en la aplicación de métodos científicos de muestreo.

Por ejemplo, si se necesita asignar 40 de las 100 unidades de la muestra de evaluación al grupo de tratamiento, puede decidirse asignar las unidades con las cifras aleatorias más altas a este grupo y dejar el resto para el de comparación. Si requerimos asignar un gran número de unidades se puede utilizar una hoja de cálculo o un software especializado en estadística para asignar un número aleatorio a cada unidad de observación de la muestra de evaluación, y se usa la regla elegida para formar ambos grupos.

b. Métodos cuasiexperimentales (no aleatorios)

En los diseños de evaluaciones cuasiexperimentales no hay asignación aleatoria de los grupos que reciben el beneficio del programa/proyecto. Se busca que algún grupo que no es parte de los beneficiarios conforme el contrafactual. Estos métodos generan grupos de comparación que se asemejan al grupo de tratamiento, al menos

en las características observadas, usando métodos econométricos. Por lo tanto, se deben aplicar controles estadísticos para abordar las diferencias entre los grupos de tratamiento y de comparación y emplear técnicas de pareo sofisticadas para crear un grupo de comparación que sea lo más similar posible al grupo de tratamiento. En algunos casos también se selecciona un grupo de comparación antes del tratamiento, aunque la selección no es aleatoria.

Una situación común que obliga al uso de métodos cuasiexperimentales se da cuando los PPP no cuentan con una línea de base aplicada a un grupo de control.

En las evaluaciones no experimentales existen (más que en los diseños experimentales) riesgos o hipótesis alternativas a aquella que interesa *"es el proyecto la causa de cambios estadísticamente significativos entre el grupo de tratamiento y el de control"*. Las hipótesis alternativas se describen en términos de aquellos otros factores distintos al tratamiento que podrían influir en el cambio observado. El reto es cómo aislar el efecto del PPP de los efectos de los otros factores.

Por ejemplo: en un programa de desarrollo de infraestructura vial se espera un cambio en el ingreso de la población debido a que las comunidades beneficiarias tienen mayor acceso al mercado y menores costos de transporte de sus productos debido a la nueva carretera. Sin embargo, los beneficiarios pueden haber mejorado en su productividad. En este caso, la explicación del cambio en el ingreso debido al cambio en la productividad es una hipótesis alternativa.

Al respecto, debemos mencionar que en el uso de métodos experimentales de asignación aleatoria se debe diseñar una muestra de población con diferentes niveles de productividad, de forma que este factor esté controlado y no sea fuente de los cambios observados. Sin embargo, en los métodos cuasiexperimentales, estas hipótesis alternativas deben ser controladas con diferentes técnicas.

De otro lado, también se debe indicar que existen algunos tipos de hipótesis alternativas que amenazan la validez de los resultados al aplicar cualquiera de los dos métodos indicados (experimentales y cuasiexperimentales). Entre ellos se encuentran:

- ⇒ Difusión de resultados o contaminación de los efectos entre el grupo que recibe el proyecto y el que no. Por ejemplo, si las personas del grupo de control al enterarse del tratamiento buscan y logran participar de este.

⇒ Efectos de igualdad compensatoria. Hacen referencia al caso en que el grupo control, al no participar en el programa, se organiza e impulsa y obtiene por sí mismo el tratamiento.

En el caso de los métodos cuasiexperimentales, una de las principales hipótesis alternativas que nos puede generar problemas al estimar los efectos causales es el denominado sesgo de selección.

i. ¿Qué es el sesgo de selección?

El sesgo de selección en el lenguaje de la evaluación de proyectos es el conjunto de factores no observables que definen la participación de los potenciales beneficiarios en un PPP.

Por ejemplo, en un programa de formación profesional para jóvenes desempleados. Dos años después de la puesta en marcha del programa, una evaluación intenta estimar su impacto sobre el ingreso comparando la media de ingresos de un grupo de jóvenes que decidió inscribirse con la de otro grupo que eligió no hacerlo.

Supóngase que, según los resultados, los jóvenes inscritos ganan el doble que los que no se inscribieron, que son los que constituyen el contrafactual. ¿Cómo se deben interpretar estos resultados? Habría que considerar que quienes deciden participar en el programa pueden tener muchos medios para mejorar su ingreso y pueden esperar un gran beneficio del programa. Por el contrario, los que deciden no inscribirse pueden ser jóvenes desanimados que no esperan beneficiarse de este tipo de programa. Es posible que el desempeño de estos grupos en el mercado laboral sea diferente y tengan diferentes ingresos incluso sin el programa de formación profesional. Si en la evaluación se observara una diferencia en los ingresos de ambos grupos, no se podría determinar si esta se debe al programa de formación o a diferencias en la motivación u otros factores no observados previas entre ambos grupos. Cuando las personas menos motivadas deciden no inscribirse en un programa se genera un sesgo de selección. El problema consiste entonces en la buena determinación del contrafactual y en el control del efecto de las variables que motivan el sesgo de selección para participar en el programa.

ii. Ventajas y limitaciones

La ventaja principal de los diseños cuasiexperimentales es que se pueden basar en fuentes de datos existentes y, por lo tanto, a menudo son más rápidos y menos costosos de implementar. Además, se pueden realizar una vez que el programa se ha implementado, a condición de que existan suficientes datos.

Las desventajas principales de los métodos cuasiexperimentales son:

- (a) Con frecuencia se reduce la confiabilidad de los resultados, puesto que los métodos son menos sólidos estadísticamente.
- (b) Los métodos pueden ser estadísticamente complejos.
- (c) Conllevan el problema de sesgo de selección.

Al generar un grupo de comparación en lugar de asignarlo aleatoriamente, hay muchos factores que pueden afectar la confiabilidad de los resultados. La complejidad estadística requiere conocimientos especializados considerables en el diseño de la evaluación y el análisis e interpretación de los resultados.

iii. Principales métodos cuasiexperimentales.

Entre los métodos más conocidos de evaluación cuasiexperimental se encuentran: métodos de doble diferencia, los métodos de pareamiento y los métodos de variables instrumentales. A continuación pasamos a explicarlos.

Doble diferencia (diferencias en diferencias - DD)

Este método consiste en comparar los cambios a lo largo del tiempo en la variable de interés entre el grupo de tratamiento y los del grupo de control. Dentro de los métodos cuasiexperimentales, el método de diferencias en diferencias es uno de los más fuertes estadísticamente y frecuentemente utilizado.

Este método se denomina doble diferencia porque la comparación de los cambios obtenidos se realizará en tres pasos:

- a) Se determina la diferencia en la variable de interés entre el antes y el después de la intervención en la población intervenida por el proyecto. Los cambios observados podrán explicarse como consecuencia del proyecto y del contexto. Llamemos R a esta diferencia.

- b) Se determina la diferencia en la misma variable de interés entre el antes y el después en el grupo de control. Los cambios se podrán explicar como consecuencia del contexto. Denominemos S a la diferencia.
- c) Se determina la diferencia en los dos resultados, es decir R-S., Esta diferencia da la medida del cambio que se podrá atribuir a la intervención.

Aplicación del método. Considérese por ejemplo el caso de un programa de construcción de carreteras que no se puede asignar aleatoriamente. Uno de los objetivos del programa es mejorar el acceso a los mercados laborales y uno de los indicadores de resultado es el empleo.

En este caso, la simple observación del cambio antes-después en las tasas de empleo de las áreas afectadas por el programa no nos mostrará el efecto causal del programa, porque es probable que muchos otros factores influyan en el empleo. Al mismo tiempo, la comparación de las áreas que recibieron y no recibieron el programa de carreteras será problemática si existen razones no observables por las que unas lo recibieron y otras no.

Sin embargo, ¿qué pasaría si se combinaran los dos métodos y se compara los cambios antes-después de los resultados de un grupo inscrito en el programa con los cambios antes-después de un grupo no inscrito en el programa? La diferencia de los resultados antes-después del grupo inscrito (la primera diferencia) considera factores constantes en el tiempo para dicho grupo, ya que se compara al grupo consigo mismo. Sin embargo, todavía nos quedan los factores externos que varían con el tiempo. Una manera de observar estos factores variables en el tiempo es medir el cambio antes-después de los resultados de un grupo que no se inscribió en el programa, pero estuvo expuesto a las mismas condiciones ambientales (la segunda diferencia).

Si se "limpia" la primera diferencia de otros factores variables en el tiempo que afectan al resultado de interés sustrayéndole la segunda diferencia, se habrá eliminado la principal causa de sesgo que nos preocupaba en el caso de las comparaciones simples de antes-después. Por lo tanto, el método de diferencias en diferencias incluiría: comparaciones antes-después y comparaciones inscritos-no inscritos en el programa para generar una mejor estimación del contrafactual. En nuestro caso sobre carreteras, el método de diferencias en diferencias podría comparar el cambio en el empleo antes y después de la implementación del programa en el caso de las personas que

viven en áreas afectadas por el programa de construcción de carreteras con los cambios en el empleo en las áreas en las que no se implementó el programa.

Si bien este método nos permite tener en cuenta las características constantes en el tiempo entre los grupos de tratamiento y de comparación, no nos ayuda a eliminar el sesgo producido por aquellas características que varían a lo largo del tiempo entre los grupos de tratamiento y de comparación.

En el ejemplo anterior del programa de carreteras, si las zonas de tratamiento se benefician también de la construcción de un puerto marítimo, no se podrá tener en cuenta el efecto de la construcción del puerto marítimo mediante un método de diferencias en diferencias. Para que este método genere una estimación válida del contrafactual, se debe suponer la inexistencia de estas diferencias que aparecen a lo largo del tiempo entre los grupos de tratamiento y de comparación.

Cuando se usa el método de diferencias en diferencias, se debe suponer que, en ausencia del programa, el resultado del grupo de tratamiento habría evolucionado paralelamente al grupo de control; a esto se le denomina ***el supuesto de tendencia común***.

A pesar de que no se puede demostrar, el supuesto fundamental de igualdad de tendencias puede ser comprobado comparando el comportamiento en los grupos de control y de intervención en un periodo anterior a la implementación del PPP. Si los resultados evolucionaron en paralelo antes del inicio del programa, se tendrá más confianza en que los resultados habrían seguido progresando simultáneamente en el periodo posterior a la intervención en ausencia de la intervención.

Pareamiento

El pareamiento es un método que utiliza técnicas estadísticas avanzadas para construir un grupo de comparación artificial: se usa cuando previo a la intervención del PPP se invita a los potenciales beneficiarios a participar en el proyecto; unos aceptan y otros no, obviamente. En este caso forman parte del grupo de intervención aquellos que decidieron participar en el proyecto.

La idea para la evaluación consiste en organizar pares de individuos participantes-no participantes que tengan las características más similares posibles entre ellos.

El método consiste en calcular un puntaje de propensión a participar del PPP, que puede ser expresado en una probabilidad de participar, de todos los potenciales beneficiarios, tomando en cuenta varias variables. Luego se forman parejas beneficiario-no beneficiario que tengan el mismo puntaje de propensión, o que estén muy próximas entre sí. Estas "unidades próximas" se convierten en el grupo de control y se convierten, por tanto, en el contrafactual.

Los de pareamiento son métodos alternativos que nos permiten estimar el efecto causal de una intervención sobre la cual solamente tenemos una medición (no hay un antes y después como en el caso del método de doble diferencia), no hay asignación aleatoria ni se conoce la regla de asignación a la intervención.

Los métodos de pareamiento se pueden aplicar a casi todas las reglas de asignación de un programa, siempre que se cuente con un grupo que no haya participado en este. Estos se basan normalmente en características observadas para construir un grupo de comparación y, por lo tanto, suponen que no hay diferencias no observadas entre los grupos de tratamiento y de comparación que estén asociadas también con los resultados de interés.

Ejemplo de aplicación

Consideremos un caso en el que se debe evaluar el impacto de un programa y se cuenta con una serie de datos tanto de hogares inscritos en el programa como de no inscritos, por ejemplo una encuesta demográfica y de salud. El programa no tiene ninguna regla de asignación clara (como una asignación aleatoria o un índice de elegibilidad) que explique por qué algunos hogares están inscritos y otros no. En dicho contexto, los métodos de pareamiento permitirán identificar, mediante las características disponibles en la base de datos, a los hogares no inscritos más parecidos a los hogares inscritos (o de tratamiento), esto por medio de las características disponibles en su base de datos. Estos hogares no inscritos "pareados" se convierten así en el grupo de control para estimar el contrafactual.

Ventajas y limitaciones. Aunque la técnica del pareamiento requiere una cantidad significativa de datos, se trata de un método relativamente versátil que se ha empleado para evaluar programas de desarrollo en numerosas situaciones. Este método presenta otras desventajas: el pareamiento solo se puede realizar en función de las características observadas que se hayan definido en la fuente de donde se obtienen los datos, no se pueden incorporar características no observadas al cálculo de la puntuación de la propensión. Por lo tanto, para que el procedimiento de pareamiento identifique a un grupo de comparación válido, es necesario asegurarse de que no existen diferencias sistemáticas entre las características no observadas de las unidades de tratamiento y las unidades de comparación pareadas que pudieran influir en el resultado. Es decir, deben controlarse los posibles sesgos de selección. Además, para el pareamiento se debe emplear variables disponibles en fuentes de datos cercanas al inicio de la intervención.

Métodos de variables instrumentales

Son métodos que se usan para controlar el efecto de factores o aspectos confusores en las estimaciones de los impactos de los PPP, y que por tanto pueden producir errores. Estas variables generalmente no forman parte de aquellas observables. Para controlar estos errores se incorporan variables relacionadas a ellas a las cuales se les conoce como variables instrumentales.

Una variable instrumental debe contar con dos características para poder considerarse válida en el caso de PPP:

- ⇒ Debe estar correlacionada con la participación en el programa.
- ⇒ No puede estar correlacionada con los resultados (más allá de su participación en el programa) o con variables no observables.

Cuando existe focalización de un PPP y se han conformado grupos de control, es posible que dicha conformación sea “incompleta” porque no podríamos saber si estamos incluyendo todas las variables pertinentes; algunas no las podríamos observar, de modo que el cambio en los indicadores de resultado podría deberse justamente a esas variables. Así, algunas variables omitidas podrían ser las que expliquen la participación en un proyecto y por lo tanto influyen en el cambio de los indicadores de impacto. Por lo tanto, debemos entonces considerarlas en las estimaciones del efecto causal.

Una solución a este problema es utilizar “una o más variables que influyan en la participación, pero no en los resultados dada la participación. Esto identifica la variación exógena en los resultados atribuibles al programa, reconociendo que su establecimiento no es aleatorio sino intencional. Las “variables instrumentales” se usan primero para predecir la participación en el programa y luego se observa cómo varía el indicador de resultados con los valores proyectados”.⁸

Ejemplo de uso. La evaluación de un programa de educación realizado en la sierra de Perú se basa en la comparación de los años promedio de estudios alcanzado por los niños de doce años, que ya deben haber concluido la primaria pues la edad normativa es de los seis a los once años de edad. La encuesta de evaluación se toma en el mes de agosto.

La simple comparación de los años promedio alcanzados obtenido en la encuesta puede contener un error, pues el solo uso de la variable edad puede generar sesgos debido a que el sistema educativo indica que la matrícula al primer grado solo procede para los niños que tengan seis años cumplidos hasta el 31 de marzo de ese año; los nacidos en fecha posterior tendrán que esperar hasta el siguiente año escolar.

En los datos de la encuesta tomada en el mes de agosto, en los niños que tienen doce años se encontrará hasta un máximo de seis años de estudio en los que tienen fecha de nacimiento anterior al 31 de marzo, mientras que los que tienen fecha de nacimiento posterior pueden tener solo hasta un máximo de cinco años de estudio. Este hecho por si solo ya está ocasionando una diferencia de un año en la educación alcanzada, diferencia que es ajena a la intervención del PPP. Si la encuesta tiene el dato de la fecha de nacimiento se puede construir una variable instrumental con dos valores: 1 para aquellos que tienen fecha de nacimiento anterior al 31 de marzo, y 0 para el segundo para los que la tienen luego de tal fecha, y esta variable se incorpora en el modelo.

⁸ “Evaluación del impacto de los proyectos de desarrollo en la pobreza” Baker Judy. Banco Mundial p. 7.

Limitaciones

La principal limitación se da en la identificación de una buena variable relacionada con la participación en el PPP. Otro aspecto es que no debe ser empleado cuando el efecto de la variable instrumental es muy grande.

En el análisis se emplean métodos estadísticos avanzados como el enfoque de los mínimos cuadrados en dos fases, lo que significa que se debe hacer una doble regresión.

Método de regresión discontinua

El método de regresión discontinua se usa en la evaluación de PPP que se enfocan a individuos que son elegibles según su posición con respecto a un índice continuo o escala en un momento próximo al inicio de la intervención (línea de base), como por ejemplo a las personas cuyo ingreso per cápita están por debajo de la línea de pobreza, o cuya edad está por encima de los sesenta y cinco años. El método requiere disponer de un índice continuo y un punto de corte específico.

El impacto se mide comparando dos grupos cuyos índices se encuentren próximos al punto de corte. Por ejemplo, si la intervención se da a los que están por debajo del punto de corte, el grupo de control o contrafactual estará constituido por los que están por encima del punto de corte. Se puede considerar que todos los que están próximos al punto de corte tienen las mismas características y que la única diferencia en los resultados se debería a la intervención del PPP.

Este método requiere de muestras suficientemente grandes de tal manera que las vecindades construidas en torno a los puntos de corte para formar los grupos cuenten con suficientes casos para hacer inferencias; además necesita del empleo de métodos estadísticos avanzados.

2.7.5 Métodos cualitativos

Los métodos cualitativos en evaluación tratan de dar respuesta a los porqués de los resultados observados, añadiendo más profundidad a los hallazgos de la evaluación de un PPP. Dentro de los principales métodos cualitativos se encuentran: (i) la observación, (ii) las entrevistas, (iii) los grupos focales y (iv) la revisión de documentos, registros, materiales y artefactos.

a. La observación

Este método se basa en el hecho de captar con todos los sentidos los hechos o acontecimientos en el transcurso de una situación determinada. Por ejemplo, si uno se encuentra en un salón de clases, es posible observar el color de las paredes, el número de ventanas, las características de los alumnos o el simple olor del recinto. Es necesario utilizar todos los sentidos, puesto que todo puede ser relevante: clima, espacio físico, colores, aromas, iluminación, entre otros.

Los propósitos de la observación cualitativa son: a) explorar ambientes, contextos y los diferentes espacios de interacción social, b) descripción de comunidades, contextos o ambientes, c) comprender los procesos, vinculaciones entre personas o situaciones, d) identificar problemas y e) generar hipótesis para estudios posteriores.

Ejemplo: percepción de una capacitación a ciudadanos para la lucha contra la inseguridad ciudadana

En un escenario donde se llevan a cabo capacitaciones a vecinos de una determinada área urbana, un observador puede tener a la mano una lista de cotejo que contenga los siguientes aspectos:

- Los vecinos son capacitados en la reacción ante un robo o hecho violento.
- El capacitador brinda los materiales educativos a los vecinos.
- El capacitador absuelve las preguntas hechas por los capacitados.
- Se brindan ejemplos y casos prácticos durante la capacitación.
- El capacitador brinda información clave de contactos para la denuncia de hechos delictivos.

Teniendo estos puntos en cuenta, el observador puede crearse un criterio de lo que puede estar pasando en este tipo de actividades comparándolas con un modelo de actividad (cómo debería ser) para emitir las recomendaciones del caso.

Existen por lo menos tres formas de observación: directa (con lista de cotejo), por medio de cliente simulado y observación etnográfica.

i. Observación directa

Se utiliza frecuentemente para evaluar el desempeño de una actividad considerada importante para la intervención, por ejemplo, la competencia técnica del médico, educador o de asistencia técnica. En otras palabras, el

observador asiste a la actividad, observa si la persona que realiza la actividad la hace bien y ofrece sugerencias para su mejoramiento. Sin embargo, este tipo de observación puede ser muy subjetiva si los parámetros incluidos en la lista de cotejo no son claramente observables, es decir, que estos parámetros no puedan ser comparados con lo que debería ser.

ii. Observación sistemática por medio de cliente simulado(a)

En esta metodología un miembro del equipo de evaluación se hace pasar como un beneficiario de la intervención para observar el nivel de servicio que ofrece el personal del PPP cuando no saben que están bajo observación. Una ventaja de esta técnica es que el evaluador puede observar lo que le pasa al beneficiario bajo "circunstancias normales".

iii. Observación etnográfica

La observación etnográfica consiste en observar el comportamiento y las relaciones socioculturales que rodean un fenómeno o evento en una población determinada y durante un tiempo determinado. Esta técnica se usa para describir y explicar las secuencias de comportamientos y formas de interacción que se presentan en el sistema sociocultural en una población en relación a un cierto fenómeno o evento. En la observación etnográfica, el etnógrafo vive y trabaja por un tiempo en contacto con el grupo o comunidad, observando la interacción mediante su participación en la comunidad o grupo que está siendo evaluado. Debido al largo plazo que requiere este tipo de metodología, no es práctica para una evaluación.

b. Las entrevistas a profundidad

La entrevista a profundidad es una técnica de recojo de información íntima, flexible y abierta que se define como una reunión para intercambiar información entre el entrevistador y el (los) entrevistado (s). En la entrevista se logra una comunicación y la construcción conjunta de significados respecto a un tema.

Las entrevistas se dividen en semiestructuradas y no estructuradas o abiertas. Las entrevistas semiestructuradas se basan en una guía de asuntos o preguntas y el entrevistador tiene la libertad de introducir preguntas adicionales para precisar conceptos u obtener mayor información sobre los temas deseados. Las entrevistas abiertas se fundamentan en una guía general de contenido y el entrevistado posee toda la flexibilidad para manejarla.

En cuanto al tipo de preguntas que pueden contener las entrevistas, estas se pueden dividir en dos grupos:

Preguntas de opinión. Como su nombre lo dice, son preguntas que piden al entrevistado la apreciación de un determinado fenómeno (que haya sucedido, suceda o podría suceder); la expresión de sentimientos, donde se solicita expresar el sentir o una descripción profunda de determinados eventos que atañen al entrevistado; la manifestación de conocimientos, con el propósito de evaluar si el entrevistado cuenta con cierto conjunto de información; preguntas sensitivas, las que se encuentran relacionadas a los sentidos (que oyó, vio, escuchó, etc.); preguntas de antecedentes, donde se indaga sobre eventos previos al fenómeno de interés; y preguntas de simulación, donde se intenta obtener la opinión sobre eventos hipotéticos.

c. Grupos focales

Los grupos focales pueden ser considerados como una especie de entrevistas grupales. Consisten en reuniones de grupos pequeños o medianos (de tres a diez personas) en las cuales los participantes conversan en torno a uno o varios temas en un ambiente relajado e informal bajo la conducción de un especialista en dinámicas grupales. El tamaño de los grupos depende del tema: de tres a cinco personas cuando se expresan emociones profundas o temas complejos y de seis a diez participantes si las cuestiones a tratar versan sobre asuntos más cotidianos, aunque las sesiones no deben excederse de un número manejable de individuos. Asimismo, se deben construir tantos grupos como sean necesarios para diferenciar la información proporcionada.

El cuadro 14 presenta dos ejemplos de grupos focales:

Cuadro 14. Ejemplos de grupos focales

Naturaleza del estudio	Grupos que podrían integrar el estudio
Comprender las razones por las cuales mujeres que de manera constante son agredidas físicamente por sus esposos, mantienen la relación marital a pesar del abuso.	Tres o cuatro grupos pequeños (cinco participantes por grupo). Los grupos podrían integrarse por el grado de agresión física o el tiempo de abuso o tomando en cuenta ambos elementos. Son recomendables de cuatro a cinco sesiones.

Naturaleza del estudio	Grupos que podrían integrar el estudio
Analizar los problemas en la atención a pacientes de un hospital.	Un grupo formado de médicos, otro de enfermeras, uno de residentes, uno de personal auxiliar, dos grupos mixtos (médicos, enfermeras, residentes, auxiliares) y dos de pacientes. Seis o siete participantes por grupo. Más de una sesión por grupo, de ser necesario.

Una vez identificado el tipo de grupo focal a llevarse a cabo, es importante considerar los siguientes pasos para la realización del mismo:

- Se determina un número de grupos y sesiones a realizar.
- Se definen los perfiles de personas que participarán en la sesiones.
- Se detecta a las personas del tipo elegido.
- Se invita a estas personas a la sesión.
- Se organizan las sesiones. Cada sesión debe efectuarse en un lugar confortable, silencioso y aislado.
- Se lleva a cabo la reunión.
- Se elabora el reporte de la sesión.

Es recomendable grabar las sesiones para poder tener registrada toda la información aportada, así como para realizar análisis más profundos donde los tonos de voz, los énfasis, etc., sean importantes para interpretar la información.

d. Revisión de documentos, registros y materiales

Los documentos, registros y materiales son una fuente valiosa de información cualitativa en una evaluación que pueden ayudar a entender la intervención materia de estudio. Estos elementos sirven al investigador para conocer los antecedentes de un ambiente, experiencia, vivencias o situaciones y su funcionamiento cotidiano. Existen diversas fuentes de documentación, aunque las más comunes son las individuales y las grupales.

i. Individuales

- Documentos escritos personales. Son de tres tipos: a) los documentos o registros preparados por razones oficiales tales como certificados de nacimiento o de matrimonio, licencias de manejo, cédulas profesionales, escrituras de propiedades, entre otros; b) documentos preparados por razones personales como cartas, diarios, manuscritos y notas; y c) documentos preparados por razones profesionales (reportes, libros, artículos periodísticos, etc.), cuya difusión es pública.
- Materiales audiovisuales. Conformados por imágenes (fotografías, retratos, pinturas, etc.). Asimismo, cintas de audio-video generadas con un propósito específico.
- Artefactos individuales. Artículos creados con ciertos fines por una persona: ropa, herramientas, mobiliario, juguetes, etc.
- Archivos personales. Colecciones o registros privados de un individuo.

ii. Grupales

- Documentos grupales. Ejemplos de ellos son las actas institucionales, ponencias, declaraciones o manifiestos.
- Materiales audiovisuales grupales. Imágenes, cintas de audio-video, páginas web, entre otras, producidas por un grupo con objetivos oficiales, profesionales u otros.
- Artefactos o construcciones grupales. Elementos como edificaciones o esculturas colectivas.
- Documentos y materiales organizacionales. Memos, reportes, planes, evaluaciones, cartas, mensajes en los medios de comunicación, avisos, entre otros.
- Registros en archivos públicos.
- Huellas digitales o de otro tipo, rastros, vestigios, medidas de erosión o desgaste y de acumulación, entre otros.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Taylor & Bodgan. Introducción a los métodos cualitativos de investigación. Ed, Paidós. Capítulo 1: Introducción ir hacia la gente. 2000.

2.7.6 Métodos mixtos

El enfoque mixto, al cual denominamos métodos mixtos, es un proceso en el que se recolecta, analiza y vincula datos cuantitativos y cualitativos en una misma evaluación u estudio para responder a las preguntas e hipótesis planteadas. Es importante señalar que el enfoque mixto va más allá de la simple recolección de datos de diferentes modos sobre la misma intervención; también tiene implicancias en todas las etapas de una evaluación, tanto en el diseño, la recolección y análisis de los datos como en el reporte del estudio, ya que finalmente no se presentan los hallazgos de información cuantitativa y cualitativa por separado o según el instrumento que se empleó; por el contrario, trabajar bajo este enfoque implica analizar de manera conjunta toda la evidencia recopilada a fin de emitir juicio sobre los avances, logros y dinámicas de la gestión de un programa/proyecto.

Una de las principales ventajas del enfoque mixto es que este ofrece una perspectiva más precisa, integral y completa sobre un PPP en evaluación. Asimismo, ayuda a presentar los resultados de manera más amigable y permite un mayor grado de aceptación.

Sin embargo, la implementación de un enfoque mixto implica una serie de retos. Por ejemplo, el equipo evaluador debe poseer un fuerte entrenamiento tanto en los métodos cualitativos y cuantitativos,

y los costos de recopilación de información primaria podrían ser sustancialmente mayores al utilizarse ambos métodos. Por tanto, en la etapa de diseño hay que evaluar a detalle si se requerirá utilizar únicamente métodos cuantitativos o cualitativos o si la evaluación implementará ambos métodos; ello dependerá en buena medida del tipo de evaluación que se esté implementando (revisar preguntas de investigación) así como de la naturaleza de las variables del PPP, tal como se verá en el siguiente paso.

2.8 Identificación de variables e indicadores a medir/verificar y las fuentes de información

Hasta este punto ya se tienen definidos el tipo de evaluación a implementarse, las preguntas de evaluación, la Teoría de Cambio del PPP, así como los niveles que se van a evaluar (resultados y/o productos) y se han formulado las hipótesis y métodos a utilizar. Ahora, el siguiente paso consiste en identificar las variables e indicadores de cada una de las actividades, productos y/o resultados que se espera evaluar, los cuales nos ayudarán a generarnos una idea de situación y avance de estos aspectos

de la iniciativa. Para todo ello, los conceptos que debemos manejar a detalle son aquellos relacionados con los indicadores.

2.8.1 Las variables

Una variable es una propiedad o característica cuyos valores o categorías pueden variar y que son susceptibles de medirse u observarse. Ejemplos de variables son el género, la religión, la resistencia al material, la productividad de algún tipo de semilla, el ingreso familiar, etc.

El concepto de variable se aplica a personas u otros seres vivos, objetos, hechos y fenómenos.

2.8.2 ¿Qué es un indicador?

El indicador puede definirse como una medición que sirve para describir un comportamiento, concepto o fenómeno. Tal medición puede que no abarque todos los aspectos del fenómeno de interés, pero (tal como permite intuir su nombre) permite alguna "indicación" de ello. Un indicador puede ser en algunos casos una variable, y en otros puede ser un constructo que combina varias variables. Por ejemplo, imagine que se quiere medir el "bienestar de la población estudiantil que recibe desayunos escolares". De primera impresión, el bienestar es un concepto muy complejo y abarca muchas dimensiones, sin embargo es posible determinar algunas características relevantes como:

- Nutrición del estudiante.
- Rendimiento académico
- Hábitos alimenticios.
- Estado anímico del estudiante.

Cada uno de estos aspectos puede constituirse en un indicador del bienestar del estudiante. Es evidente que el fenómeno en cuestión es mucho más grande que los cuatro aspectos arriba señalados, pero a pesar de ello estos indicadores permiten tener la posibilidad de cuantificar el fenómeno de interés y conocer sus características intrínsecas.

Un indicador puede expresarse por medio de un número absoluto, un porcentaje, una tasa (prevalencia o incidencia) o por una expresión (un sí o un

no). Sin embargo, el indicador no solo abarca aspectos cuantitativos, sino también aspectos cualitativos que permiten una mayor profundidad y detalle.

Ejemplo de indicadores aplicados a la evaluación de procesos

La evaluación de proceso puede tener aspectos cuantitativos y cualitativos. Por un lado, hay indicadores cuantitativos que miden número, porcentaje, promedio u otras estadísticas, y por otro hay aspectos cualitativos que se desea monitorear y evaluar.

Por ejemplo, para evaluar el grado de satisfacción de los beneficiarios de una intervención que mejore la productividad en las micro y pequeñas empresas, es posible aplicar un cuestionario estructurado para medir y determinar de manera puntual qué porcentaje opinó que las capacitaciones en productividad fueron útiles para sus propias empresas. Asimismo, es posible estimar cuántos trabajadores en promedio tienen las empresas que participan en la intervención. Los dos ejemplos anteriores corresponden a indicadores de tipo cuantitativo, los cuales permiten cuantificar y medir magnitudes.

De manera alternativa, es posible reunir a todos los participantes de la capacitación y conocer su percepción acerca de la intervención, sus actitudes y principales reacciones al conocimiento que les es transmitido. De la misma forma, es posible entrevistar a algunos empleados de las micro y pequeñas empresas para poder identificar algunas limitaciones y problemas que puedan tener en su funcionamiento cotidiano. Estos últimos ejemplos corresponden a los aspectos cualitativos de los indicadores, dado que indagan de manera detallada y un tanto más subjetiva los aspectos de interés de la intervención. Ambos tipos de indicadores tienen ventajas y desventajas, pero ambos trabajan de manera complementaria y sirven para obtener un conocimiento más profundo sobre la dinámica de la intervención.

Los indicadores son un elemento clave dentro de cualquier evaluación dado que es a través de la información recopilada sobre ellos que el (los) evaluador (res) podrá(n) emitir juicios sobre los avances y logros de un PPP. Así, cabe recordar que el diseño de evaluación debe llevarse a cabo conjuntamente con el diseño del PPP, a fin de garantizar que los indicadores y en general los objetivos planteados se encuentren claramente definidos para luego puedan ser medidos y analizados durante el proceso de evaluación. Es importante indicar que para determinar los indicadores más pertinentes en una evaluación es necesario preguntarse qué se desea aprender de esa evaluación. Así, al momento de seleccionar los indicadores es necesario considerar:

- Para cada resultado, objetivo o producto es necesario tener al menos un indicador para dar el debido seguimiento a su desempeño. Al planificar la evaluación, se deben establecer los principales indicadores de resultados y

efectos, en lo posible como parte de un enfoque de Marco Lógico, ello para garantizar que la evaluación pueda examinar los resultados durante un periodo que responda a las necesidades de las autoridades responsables.

- Si un determinado indicador mide un fenómeno que no presenta cambios en el corto plazo (un indicador de resultado final, por ejemplo), no es eficiente medirlo en el corto plazo, solo en intervalos de tiempo de mayor duración, y es necesario que los responsables de la evaluación puedan planificar la medición de indicadores en diversos periodos de tiempo, lo que permitirá captar antes los efectos más inmediatos y seguir evaluando las medidas de los resultados finales.
- Si el indicador propuesto no es útil para mejorar el programa o los resultados del mismo, es mejor evitar medirlo.

También se pueden considerar indicadores que capten alguna información sobre características de la población beneficiaria no relacionadas estrictamente con la evaluación o los productos y actividades de la intervención, pero que pueden ser de interés para el análisis, como su nivel de pobreza o su opinión acerca del programa/proyecto. Además, el evaluador también podría desear incluir medidas de costos para realizar algún análisis de eficacia en función de los costos u otras evaluaciones complementarias que no tengan relación estricta con la evaluación.

Una vez identificados los indicadores que medirán las dimensiones de interés de la intervención, es necesario identificar su escala de medición y las fuentes de donde se recabará la información (por ejemplo: encuestas, registros administrativos, entre otros). Cabe mencionar que para cada indicador es necesario recolectar los datos de alguna fuente; en algunos casos los datos ya están disponibles, solo es cuestión de analizarlos. En otros casos esto implica diseñar y conducir un estudio primario para obtenerlos; es por ello que luego de identificar las fuentes de información hay que determinar los instrumentos de recolección apropiados que se utilizarán en la evaluación.

Es importante que los evaluadores y el personal técnico del programa participen en la elaboración de la matriz de evaluación (ver cuadro 15), la cual contiene el detalle de todos los aspectos antes mencionados sobre los indicadores: variables, enfoque, fuentes y métodos.

Cuadro 15. Ejemplo de matriz de evaluación

Objetivos de evaluación	Indicadores	VARIABLES	escala de medición	Fuentes	Método de recolección
Pregunta de evaluación 1:	(1)	(1)			
		(2)			
		(..)			
	(2)	(1)			
		(2)			
		(..)			
Pregunta de evaluación 2	(1)				
	(..)				

Fuente: elaborado por Alejandro Bardales, 2014.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Módulo 2: Propiedades y Construcción de Indicadores del Curso Indicadores para el desarrollo – INDES. Págs. 1-53.

2.9 Determinación de la muestra

La determinación de la muestra es un paso clave para que la información recopilada para la evaluación sea representativa de la población total en estudio (personas, hogares, instituciones, distritos, regiones, o cualquier otra unidad de análisis). A continuación se presentan los conceptos clave para entender en qué consiste el muestreo cuantitativo y cualitativo.

2.9.1 Muestreo cuantitativo

Las mediciones cuantitativas requieren de muestras probabilísticas, es decir aquellas en donde se conoce la probabilidad de seleccionar algún elemento del universo. El diseño de la muestra debe responder a dos cuestiones, una es cuántos elementos del universo deben seleccionarse, otra es cómo deben ser seleccionados. Se trata de definir el tamaño de la muestra y el procedimiento del muestreo.

El tamaño de muestra depende de la variabilidad de los datos, del nivel de confianza y la precisión de los resultados esperados.

Hay que tener en cuenta que los procedimientos de muestreo probabilístico pueden ser:

- Aleatorio simple. En el cual cada elemento de la población tiene igual probabilidad de ser elegido.
- Con probabilidades proporcionales al tamaño de las unidades muestrales. Por ejemplo, los hospitales con mayor cantidad de pacientes atendidos tienen una mayor probabilidad de selección que los pequeños.
- Sistemáticos. Cuando la muestra se elige con intervalos de muestreo.
- Por conglomerados. Cuando se selecciona un grupo de elementos de la población en lugar de elementos individuales, por ejemplo, alumnos de un aula completa, todas las personas que buscan atención en un servicio en el turno de la mañana, etc.
- Estratificados. Si la población se divide en subpoblaciones o estratos.
- Por etapas, cuando la selección se realiza de manera escalonada, por ejemplo, selección de centros educativos de una ciudad en una primera fase, y luego selección de alumnos dentro de tales centros educativos.

El muestreo aleatorio simple se utiliza cuando cada una de las unidades de la población (de tamaño N) tiene la misma probabilidad de ser elegida para integrar la muestra (de tamaño n). Esta probabilidad es igual a $f = n / N$, a la que se le denomina fracción de muestreo.

Para seleccionar una muestra aleatoria simple se debe tener en cuenta las siguientes recomendaciones:

1. Enumerar los elementos de la población del 1 hasta N (N es el tamaño de la población).
2. Utilizar algún procedimiento para seleccionar los n elementos de la población que conforman la muestra, por ejemplo utilizando la función ALEATORIO.ENTRE de la hoja de cálculo Excel® o una tabla de números aleatorios.

Tamaño de muestra

El tamaño de muestra depende de los parámetros que se desean estimar. Es decir, podemos estar interesados en conocer una media aritmética poblacional (μ), una proporción poblacional (P), diferencias de medias o una diferencia de proporciones. Para cada uno de los casos mencionados anteriormente existe una determinada fórmula para determinar el tamaño de la muestra. A continuación se presenta la

fórmula para calcular el tamaño de muestra que permita estimar la media aritmética poblacional con muestreo aleatorio simple.

$$n = \frac{Z_{\alpha/2}^2 \sigma^2 N}{(N - 1)E^2 + Z_{\alpha/2}^2 \sigma^2}$$

Donde:

- $Z_{\alpha/2}^2$ es el coeficiente cuyo valor depende del nivel de confianza que se preestablece. Es decir, para un nivel de confianza del 90% le corresponde $Z_{\alpha/2}^2 = 1,61$ para 95% le corresponde $Z_{\alpha/2}^2 = 1,96$ y para 99% le corresponde $Z_{\alpha/2}^2 = 2,57$. Generalmente se emplea 95% con $Z = 1,96$.
- σ^2 es la varianza de la característica que se está estudiando. Nos indica que a mayor variabilidad en la población le corresponde mayor tamaño en la muestra y a menor variabilidad en la población le corresponde menor tamaño de muestra.
- E es el error máximo permisible que estamos dispuestos a cometer para estimar μ para un nivel de confianza establecido. A E también se le conoce como margen de error, error absoluto o precisión de la estimación.
- N es el tamaño de la población.

Para calcular el tamaño de muestra para estimar una media se necesita conocer la varianza de la población objetivo, para lo cual se tiene en cuenta las siguientes recomendaciones:

- Recurrir a estudios similares al que se quiere estudiar y de ahí obtener el valor que corresponde a la varianza.
- Realizar un estudio piloto y estimar el valor que le corresponde a σ^2
- Existe otra posibilidad: si de acuerdo con la experiencia en el tema se puede tener una idea de la variación de los datos desde un mínimo (Y_{\min}) hasta un máximo (Y_{\max}), la diferencia entre estos dos valores constituye el rango; si además se asume que la variable en estudio tiene una distribución normal, entonces el valor de la desviación estándar puede ser estimado con la siguiente fórmula. : $\sigma \approx (V_{\max} - V_{\min})/6$

Ejemplo con Muestreo Aleatorio Simple

De una población de 20 000 universitarios se desea obtener una muestra para conocer la estatura promedio. La estimación muestral deberá tener un error máximo de 1 cm. respecto del verdadero promedio, con un nivel de confianza del 95%. Esto es que en 95 de cada 100 intervalos el error no será mayor de 1. Un estudio preliminar nos indica que la desviación estándar será de 5 cm.

Si se desea hallar el tamaño de muestra, solo basta con reemplazar los datos en la fórmula mostrada arriba de la siguiente manera:

$$n = \frac{(1.96^2)5^2 20000}{(20000 - 1)1^2 + (1.96^2)5^2}$$

De donde se obtiene que la muestra bajo estudio asciende a 96 personas aproximadamente.

Cuando se va a estimar el impacto de un PPP se tiene que incorporar la diferencia esperada en la variable de interés con respecto a la del grupo de control al momento de la evaluación, así como el efecto del diseño cuando se emplean combinaciones de métodos de muestreo (conocido como muestreo complejo) y las tasas de no respuesta esperadas. Luego, el diseño de la muestra debe especificar claramente la forma como van a ser seleccionados y localizados los informantes, lo cual debe ser incluido en los manuales con instrucciones precisas para su ejecución en el campo. Asimismo, el diseño debe incluir las variables que deben considerarse en los cuestionarios o en los reportes de campo para la determinación de las probabilidades de selección, lo que va ser necesario para el momento del análisis.

Para garantizar una buena muestra se requiere contar con la participación de especialistas en muestreo para que empleen las fórmulas adecuadas para el tipo de diseño de la muestra y para la estimación de los errores muestrales (desviaciones estándar e intervalos de confianza de los resultados y los efectos de diseño cuando se trata de muestras complejas) de las estimaciones.

En el caso de las evaluaciones de impacto, para la determinación del tamaño de la muestra, es necesario considerar, además los conceptos y procedimiento de potencia y tamaño del efecto, que son abordados a continuación.

Potencia y tamaño del efecto en la determinación del tamaño da muestra

Un asunto clave en los diseños de causalidad es estimar la probabilidad de encontrar un efecto en un experimento, digamos por ejemplo detectar si hay una diferencia significativa entre el grupo de control y el intervenido (nuestro programa). Detectar

una diferencia significa que nuestro programa tuvo mejor desempeño que el grupo de control. ¿Qué podemos hacer para detectar que realmente existe esa diferencia y no equivocarnos cuando trabajamos con datos provenientes de muestras?

Previamente tenemos que recordar que en estadística se asume que la hipótesis nula (H_0) es aquella según la cual los promedios o cualquier indicador de interés no son estadísticamente diferentes entre los grupos, mientras que la hipótesis alternativa (H_1) es aquella según la cual sí hay diferencias estadísticas significativas entre los promedios de los grupos comparados.

Los tipos de error

De otro lado, las equivocaciones o errores son de dos tipos. Supongamos que trabajamos con un programa que busca aliviar la pobreza de zonas rurales:

- a) Error Tipo 1: consiste en rechazar una hipótesis que es verdadera. Por ejemplo, podemos pensar que en realidad no hay diferencias entre el grupo de control y el programa en cuanto a su nivel de pobreza (hipótesis nula es verdadera). Sin embargo podemos concluir que hay un efecto cuando los datos presentan diferencias puntuales.

En nuestro ejemplo, al medir la pobreza en una muestra de los hogares de nuestro programa esta llega al 34% mientras que en el control alcanza al 40%. ¿Los hogares que se han beneficiado de nuestro programa son menos pobres que los del control? Como trabajamos con muestras, esas diferencias podrían no ser significativas, sin embargo, podríamos estar afirmando que tienen un diferente nivel de pobreza; entonces estaríamos cometiendo el error Tipo I.

Niveles de significancia típicos: 0.05, 0.01, 0.1.

- b) Error Tipo II: consiste en no rechazar una hipótesis falsa. Por ejemplo, se puede estar aceptando que un programa no tiene efecto cuando en realidad sí lo tiene, es decir hay diferencias en los niveles de pobreza entre la población de control y la intervenida.

Esto ocurre cuando en realidad hay diferencias en los parámetros poblacionales de ambos grupos (intervención y control), pero el evaluador desestima la diferencia mostrada por los datos provenientes de una muestra. Desde el punto de vista estadístico, esto puede ocurrir cuando hay mucha variabilidad en la muestra, o cuando el test empleado en el análisis es poco

sensitivo; en ambos casos ello está relacionado con el número de observaciones sobre las que se basa el estudio (tamaño de la muestra).

Obsérvese que de una u otra manera siempre nos podemos equivocar. Sin embargo, las consecuencias de cometer el error tipo II, es decir, si aceptamos que un programa de alivio a la pobreza no tiene impacto (acepto H_0 nula), cuando en realidad sí lo tiene, podría ser algo grave, puesto que nuestro error podría servir para que se tome la decisión de cerrar un programa de alivio a la pobreza que en realidad ha sido efectivo. ¿Qué implicaría para los pobres cerrar un programa por el hecho de suponer que no ha tenido impacto?

Tratándose de un programa que tiene como objetivo aliviar la pobreza nos interesaría que nuestro diseño controle la probabilidad de equivocarnos aceptando que no hay impacto (no hay diferencias entre los grupos de control e intervenido) cuando realmente este sí existió. Interesa, como dijimos al principio, saber si existen diferencias reales entre el grupo de control y el del programa. En otras palabras:

IMPORTANTE:

Se necesita controlar la probabilidad de cometer el error tipo II.

La potencia o poder de una prueba

Es el grado de probabilidad de rechazar estadísticamente la H_0 cuando esta es falsa. En nuestra investigación imaginaria, entonces, la potencia es el grado de probabilidad que nuestros evaluadores tienen de efectivamente detectar estadísticamente diferencias entre los promedios de los grupos de intervención y el de control.

Pr (no rechazar H_0/H_a es verdadera) = β

Pr (rechazar H_0/H_a es verdadera) = $1 - \beta$ = potencia

Un poder de 80% nos dice que en un 80% de los experimentos, para un tamaño de muestra dado, si H_0 es en realidad falso (ej: el efecto del programa no es cero), vamos a poder rechazar H_0 .

Una mayor potencia es mejor porque va a ser más probable que pueda reportar un efecto si es que existe.

De qué depende el poder estadístico

El poder ($1 - \beta$) de un test estadístico es una función de los siguientes parámetros:

- El nivel de significancia: a un α mayor, mayor es el poder (el error tipo I y el tipo II son inversos).
- La precisión de la muestra. A un mayor nivel de precisión con el resto de parámetros constante se incrementa la probabilidad de rechazar la falsa hipótesis nula.
- El tamaño de efecto que representa la magnitud o fuerza de las relaciones entre las variables de la población: a mayor grado del efecto mayor la probabilidad que podría ser detectada y rechazada una falsa hipótesis nula.

Niveles de potencia típicos: 0,80 y 0,90.

Tamaño de efecto

Es la medida estadística que cuantifica o estima la diferencia entre dos grupos. La d de Cohen es una medida muy difundida; en ella el tamaño del efecto se calcula restando la media obtenida por el grupo experimental menos la media del grupo de control y dividiendo el resultado entre la desviación típica de la población a la que pertenecen ambos grupos (no suele ser habitual conocer la desviación típica de la población, de modo que se suele utilizar la del grupo de control o se combinan la del grupo experimental y la del grupo de control).

Se puede considerar que la d de Cohen representa el número de desviaciones estándar que separan a dos grupos. Por ejemplo, una d de 0,5 indica que la diferencia entre el grupo experimental y el grupo de control en la variable evaluada es de media desviación estándar. Cohen dio algunas referencias para interpretar la magnitud de los tamaños del efecto:

- $d = 0,20$: tamaño del efecto pequeño.
- $d = 0,50$: tamaño del efecto mediano.
- $d = 0,80$: tamaño del efecto grande.

Aplicaciones del poder estadístico en la evaluación causal

a. En la planificación o diseño de una evaluación

Se debe estimar *a priori* del poder estadístico. El poder mide la probabilidad *a priori* de detectar efectos significativos, entonces se puede usar como una herramienta de planificación para que un estudio tenga una probabilidad razonable de detectar un hallazgo significativo.

Al inicio del estudio el evaluador puede usar la función de potencia estadística para determinar el tamaño de muestra necesaria para encontrar un efecto si este existe. Asignando un alto nivel al poder estadístico (0,80) se ofrece una probabilidad no trivial.

Se debe entonces calcular un tamaño de muestra que incorpore el poder estadístico. Usando algún software podemos por ejemplo estimar los tamaños de muestra de una evaluación en la cual ***deseamos o asumimos los siguientes parámetros:***

- Deseamos que el poder estadístico sea de 0,8 o más.
- Asumimos un nivel de significación de 0,05.
- Asumimos que habrá un efecto de tamaño grande entre el grupo de control y el intervenido $d=0,8$.

Con estos parámetros el tamaño de muestra será de 21 para cada grupo. Si por ejemplo asumimos que el efecto de tamaño será mediano ($d=0,5$), sin cambiar los valores de los otros parámetros (nivel de significación=0,05 y poder estadístico mayor o igual a 0,8), entonces el tamaño de muestra será de 51 para cada grupo. A más pequeño el efecto de tamaño, más grande tendrá que ser la muestra para mantener un poder de 0,8.

b. Para saber si las conclusiones de una evaluación se basaron en un adecuado poder estadístico

Se puede usar el análisis de poder para llevar a cabo evaluaciones *a posteriori* del poder de las pruebas estadísticas usadas por estudios de evaluación analizando el tipo de prueba, el tamaño de muestra y el nivel de significación. Ello conduce a saber si la hipótesis nula de una evaluación no fue rechazada

porque tuvo un bajo poder, es decir, si el estudio no fue suficientemente sensible para detectar una relación que sí estuvo presente en la población.

¿Cómo incrementar el poder estadístico de una evaluación?

➤ **Establecer un apropiado diseño de investigación. Esto implica:**

Elegir las variables que son absolutamente necesarias para la investigación.

También es posible usar estrategias que emplean bloques, estratificaciones, pareo o control de factores extraños. El poder se incrementa solo si las variables elegidas correlacionan altamente con la variable dependiente.

Reducir los errores en la medición que puedan producirse por instrumentos inapropiados.

Se puede reducir el error por un diseño de mediciones repetidas (que hacen posible la covariación entre ellas).

➤ **Asignación adecuada de los sujetos al tratamiento y el control**

Cuando se trata de asignar individuos a un grupo de tratamiento, y a un control se debería asignar igual número de sujetos a ambos grupos; si los grupos no son iguales se pierde poder estadístico.

➤ **Relajando alfa**

El nivel de alfa (nivel de significación) debería determinarse por la relativa importancia de los errores Tipo I y Tipo II. Para esto los investigadores deberían conocer los costos de ambos tipos de error.

➤ **Por número de colas**

Se puede incrementar el poder del estudio usando pruebas de una cola en lugar de dos colas.

➤ **Efectos cuantificables grandes**

Desarrollar procedimientos e instrumentos de investigación que permitan efectos cuantificables suficientemente grandes como para ser detectados por experimentos de tamaño razonable, de forma similar a lo ocurrido en la astronomía con el desarrollo de telescopios más sensitivos para detectar estrellas de pequeña magnitud.

➤ Uso de enfoques alternativos

Usar enfoques de investigación alternativos, que no dependan de la estadística inferencial. Por ejemplo los procedimientos de la estadística bayesiana, que no fuerza una dependencia de un sí o un no al rechazar una hipótesis nula.

IMPORTANTE:

Para saber más sobre los tipos de muestreo puedes consultar el siguiente enlace:
<http://www.estadistica.mat.uson.mx/Material/elmuestreo.pdf>

2.9.2 Muestreo cualitativo

En los estudios cualitativos casi siempre se emplean muestras pequeñas no aleatorias y se aplican criterios distintos para seleccionar a los participantes.

Una de las limitaciones frecuentemente planteadas con relación al método cualitativo es que la representatividad de los resultados se pone en duda pues no pueden ser generalizados para la población objetivo; sin embargo debemos tener en cuenta que el interés de la investigación cualitativa en ocasiones es reflejar realidades múltiples dentro de la población objetivo. La literatura señala los siguientes tipos de muestreo como los comúnmente usados para los estudios cualitativos:

a. Casos extremos

Una estrategia para la identificación de las observaciones está representada por la selección de los casos extremos, los cuales poseen gran información dado su carácter de poco común o de representar "casos especiales" de una determinada realidad. Este tipo de muestreo ofrece la posibilidad de analizar los modelos de buenas prácticas o de fracasos en algunas intervenciones, así como de indagar acerca de los factores que han causado tales resultados.

Este tipo de estrategia tiene un papel importante en la mejora de intervenciones de características similares, sin embargo se debe tener cautela al momento de establecer los criterios de selección de las observaciones. Si el identificar casos extremos se torna de alta complejidad, es posible usar aquellos casos que varían grandemente del promedio de los casos disponibles y representan casos inusuales. Tales observaciones poseen características similares a los casos extremos dado que tienen las mismas características de ser poco comunes.

b. Casos intensivos

Esta técnica utiliza los casos ricos en información producto de una selección arbitraria de los mismos. Sin embargo, tal selección requiere de un conocimiento detallado de los problemas a ser evaluados y un diagnóstico profundo de los mismos, tanto a través de discusiones y entrevistas en profundidad realizadas al personal involucrado en la implementación del programa como por medio de las observaciones de los beneficiarios y su interacción con el personal que ofrecen los servicios, entrevistas e incluso a través de grupos focales.

c. Muestreo homogéneo

Esta técnica incluye casos similares con el fin de obtener una caracterización detallada de los subgrupos de la población investigada y es útil para los programas destinados a diversas categorías de beneficiarios. En este caso, las observaciones pueden ser seleccionadas de cada categoría de beneficiarios, lo que permitirá la formación de grupos focales para captar las diferencias en las percepciones de los participantes en el programa por cada categoría de beneficiarios.

d. Casos típicos

Esta técnica toma en cuenta la distribución de los casos con el fin de describir las características de los casos "normales" de estudio. Los casos típicos a seleccionar son los que tienen alta frecuencia en la distribución de la población estudiada. La selección se hace de acuerdo con los datos ofrecidos por las encuestas, el análisis demográfico o de otros análisis estadísticos, los cuales permiten la identificación de los casos promedio. Estos casos ofrecen una fuente importante de información sobre las oportunidades para describir y explicar lo que sucede en el programa.

e. Casos críticos

En las evaluaciones cualitativas de las intervenciones es posible usar muestras formadas a partir de casos críticos, ya que pueden identificar una situación dramática o drástica de la intervención, sus causas e importancia. Esta estrategia de muestreo se puede utilizar cuando no hay suficientes recursos para evaluar el programa, optando por el estudio de un caso crítico,

teniendo en cuenta que dispone de suficiente información sobre el desarrollo del programa.

Este modelo de análisis puede revelar una serie de causas que pueden explicar los fracasos de algunas intervenciones desde el inicio de su ejecución. La información obtenida mediante la identificación y el análisis de algunos casos críticos permite la evaluación de las prácticas, los contextos y procesos que condicionan y motivan la aparición de estos casos críticos. Lo anterior puede conducir a cambios en el programa con el fin de contrarrestar las causas y los efectos negativos que produjeron estos tipos de casos.

f. Muestreo en cadena (bola de nieve)

Esta técnica se basa en la selección de casos sobre la base de las recomendaciones de los sujetos investigados previamente; cada persona investigada recomienda otras personas u observaciones potenciales similares a ellos. Esta técnica se utiliza cuando no hay bases de datos con los nombres de las personas de la población objetivo y cuando el fenómeno estudiado tiene una muy baja frecuencia o una gran variabilidad.

En síntesis, en la investigación cualitativa la relación entre problemas de investigación y los casos seleccionados debe ser revisada continuamente. En este tipo de investigación se debe decidir cuándo y dónde observar, con quién conversar, así como qué información registrar y cómo hacerlo. Con este proceso estamos decidiendo no solo qué es lo relevante o no, sino también estamos extrayendo varias muestras de la información disponible.

2.10 Diseño de los instrumentos de recojo de información

El siguiente paso dentro del diseño de evaluación es la elaboración de los instrumentos. A continuación se presentan las pautas para el diseño de los principales instrumentos de evaluación cuantitativos y cualitativos.

2.10.1 Instrumentos cuantitativos

Los datos que se utilizarán en la evaluación son recopilados principalmente a través de encuestas por muestreo, censos (información de toda la población) y/o registros y bases de datos del programa. En el caso de las encuestas y censos, usualmente estos se implementarán a través de entrevistas estructuradas sobre un cuestionario previamente diseñado y validado que permita recoger de manera confiable la información necesaria para el cálculo de los indicadores de evaluación. Al respecto,

a continuación se explican los aspectos clave para desarrollar un cuestionario, así como para ejecutar una entrevista. Finalmente, en el punto 3 se brindan algunas pautas para realizar la recopilación de información estadística, la cual, como se indicó en la sección anterior, es un procedimiento que se realiza en todo proceso de evaluación.

a. Cuestionario

El cuestionario es un documento que recoge en forma organizada a través de preguntas las variables implicadas en los objetivos de la evaluación de forma que podamos obtener información confiable.

Una característica de este tipo de instrumento es que la mayor parte de las preguntas son cerradas o codificadas; es decir, las categorías de respuesta ya se han previsto. Sin embargo, algunas veces puede incluirse también un número pequeño de preguntas abiertas, las cuales pueden codificarse una vez terminado el número total de las entrevistas, buscando patrones que puedan sintetizarse.

Hay cuestionarios que son autoaplicados. En ellos los individuos sujetos de estudio completan por sí mismos la entrevista.

¿Cómo elaborar un cuestionario?

Paso 1. Definir el tipo de cuestionarios que requerimos

Una de las decisiones clave que debe tomarse se refiere al tipo de cuestionario que se utilizará para el recojo de información. Las principales alternativas son:

a) *El cuestionario para ser autoadministrado*, en el cual el entrevistado mismo responde las preguntas. Este tipo de cuestionario goza de algunas ventajas, como ser un procedimiento menos costoso, útil si es enviado por correo, o cuando se aplica a grupos numerosos de personas en el mismo lugar y tiempo, como por ejemplo: un estudio realizado en colegios a alumnos de determinado nivel y año de educación sobre determinados temas.

Este tipo de cuestionarios también es útil porque da mayor confianza al entrevistado para expresar puntos de vista, que pueden ocasionar situaciones embarazosas.

De otro lado, entre sus limitaciones está que solo pueden responder aquellos que pueden leer y escribir. Por otro lado, debido a falta de tiempo, interés o motivación pueden quedar preguntas sin respuesta.

Otros problemas que se pueden mencionar: los entrevistados pueden dejar preguntas sin responder; además, en muchos países no existe una "cultura" respecto al mecanismo de devolución de las entrevistas cuando estas se autoaplican a distancia (por ejemplo, por correo).

b) *El cuestionario de entrevista*, en el que las preguntas son hechas por un entrevistador, o entrevistadora, quien registra las respuestas.

En este tipo de cuestionario, en el que las preguntas son hechas por un entrevistador o entrevistadora, se reduce el problema de preguntas sin respuesta, y puede ser aplicado a personas de diferente nivel educativo. La desventaja es que se requiere un buen número de entrevistadores.

Paso 2. Formulación de preguntas

Aspectos preliminares a la preparación de las preguntas

- En primer lugar se debe tener clara la relación de los objetivos del estudio o evaluación y las preguntas. Del mismo modo, la forma de las preguntas y la forma de las respuestas.
- Decidir cuándo usar preguntas, guía o de observación (en las que el entrevistador/a observa y registra la respuesta) o preguntas literales, que son las que dan más libertad al entrevistador/a para que introduzca sus propias interpretaciones; así se reduce una fuente potencial de error en los resultados.
- Las preguntas pueden ser abiertas, cuando se escribe el texto de la respuesta del informante, o cerradas, cuando las posibilidades de respuesta están acotadas. Estas últimas pueden ser de selección simple, cuando solo se acepta una alternativa de respuesta, o múltiple, cuando se pueden marcar varias alternativas.
- Las preguntas también pueden hacer uso de escalas de medición para representar posiciones dentro de un rango. Las más usuales son las escalas de Likert, en las cuales se ponen posibilidades de respuesta ordenadas (por ejemplo: totalmente en desacuerdo, indiferente,

parcialmente de acuerdo, totalmente de acuerdo, frente a una pregunta sobre percepciones); otras son las escalas ordinales (por ejemplo: primaria, secundaria, superior, para el caso del nivel educativo), o las escalas numéricas (por ejemplo ante la pregunta: en un orden de 1 a 5 donde 1 significa nada importante y 5 significa muy importante, cómo califica usted la participación de los padres en el acompañamiento de las tareas dejadas a los hijos en la escuela?).

En cuanto a la forma de respuesta existe:

- La respuesta no estructurada o abierta, en la que se presenta un tópico a la persona entrevistada y se le alienta a que responda en sus propias palabras, y el entrevistador/a registra la respuesta literalmente como es dada.
- La respuesta estructurada o cerrada, en la que se diseña cierto número de respuestas posibles que se mencionan a la persona entrevistada dentro de la pregunta para que escoja una alternativa de respuesta. Es siempre recomendable dejar una alternativa "otros".
- Una variación introducida consiste en diseñar las posibles respuestas pero no leerlas dentro de la pregunta sino esperar a que el entrevistado responda y el entrevistador/a la clasifique donde corresponda. Como es probable que la persona entrevistada dé una respuesta diferente a las alternativas diseñadas, se introduce la categoría "otra respuesta" y se la registra literalmente donde se dice "especifique".

Consideraciones para la formulación de preguntas

En el proceso de redacción de las preguntas se debe tener en cuenta que se desea obtener información completa y precisa, y a la vez mantener la cooperación de la persona entrevistada. Del balance entre estos dos aspectos depende el éxito de la redacción de las preguntas. Se debe tener en consideración lo siguiente:

- El lenguaje debe ser simple, directo, en un nivel que pueda ser entendido por todos los entrevistados. No usar palabras técnicas.
- La pregunta debe ser específica y tratar una idea única.

- La pregunta no debe sugerir que una respuesta es más deseable que las otras.
- No asumir o suponer nada respecto al comportamiento, actitudes, creencias y conocimientos de la persona entrevistada.

Paso 3. Organización del cuestionario

- Hay que organizar las preguntas de modo tal que puedan ser aplicadas sin dificultad y se pueda, decidir la secuencia de las preguntas según lo tópicos o asuntos de los que se trate, considerando el comportamiento probable tanto del entrevistador/a como de la persona entrevistada.
- Tener el cuidado de ubicar las preguntas sensitivas, pues no es aconsejable dejarlas para el final; deben ser generalmente introducidas dentro del material que sea menos sensitivo y que presente el tópicos gradualmente.
- Desde el punto de vista del entrevistador/a las preguntas deben tener una secuencia clara y ordenada, con instrucciones adecuadas sobre como pasar de una sección a otra.
- Es necesario también determinar la forma física de los formatos a usarse: color y peso del papel, el tamaño del formato, tipo de letra, la identificación del cuestionario, suficiente espacio para que la respuesta sea registrada con claridad; no solo teniendo en cuenta el uso del entrevistador/a sino también las etapas posteriores al trabajo de campo.

Paso 4. Elaboración de manuales

Esta tarea consiste en la elaboración de las instrucciones y pautas metodológicas para la realización del trabajo de campo, incluida la aplicación del cuestionario y tareas específicas de los encuestadores y supervisores que se organizan en manuales.

Paso 5. Realizar las pruebas piloto

Cuando el cuestionario parezca satisfactorio, debe ser sometido a una prueba bajo condiciones similares a aquellas que serán encontradas en el trabajo mismo, en lugares similares a los de la muestra pero en hogares que no estén considerados en la muestra.

La prueba previa es útil para determinar la viabilidad y adecuación de los instrumentos y contenidos teóricos, saber si las preguntas tienen interpretaciones inequívocas, y saber cómo se puede frasear en un idioma distinto, por ejemplo el quechua, para hacer traducciones correctas. Asimismo hay que analizar si la secuencia es la mejor para no causar aburrimiento de la persona entrevistada o si la secuencia evita que algunas respuestas sean sugeridas. Por ejemplo, en algunos cuestionarios las preguntas sobre conocimientos se ponen después de las de prácticas, para no inducir las respuestas acerca de cuál es la "práctica correcta".

De esta manera se podrán llevar a cabo modificaciones técnicas y obtener finalmente el cuestionario definitivo, con el cual se capacitará al personal de trabajo de campo.

Cuando es necesario operacionalizar conceptos complejos como por ejemplo nivel de conocimientos, nivel de compromiso, actitudes, etc., generalmente se construyen conjuntos de preguntas con una puntuación para cada respuesta y luego se aplican escalas aditivas. Este tipo de instrumentos deben ser validados previamente siguiendo métodos científicos.

Para ampliar el tema tratado, revisa la siguiente lectura:

- García. T. El cuestionario como instrumento de investigación / evaluación. Sin editor. Almendralejo, 2003. Disponible en: http://cvonline.uaeh.edu.mx/Cursos/Maestria/MTE/Gen02/seminario_de_tesis/Unidad_4_anterior/Lect_El_Cuestionario.pdf

2.10.2 Instrumentos cualitativos

a. Formatos de observación

A diferencia de los formatos de recolección de información cuantitativa, en la observación cualitativa es necesario captar todo lo que se considere pertinente. Por tal motivo, el investigador debe entrenarse en técnicas psicológicas, antropológicas, sociológicas, comunicaciones, educativas y otras similares. De ahí que los formatos de observación se ajusten a las necesidades del investigador y a las dimensiones del fenómeno.

Un ejemplo del uso de los formatos de observación puede darse en una evaluación para analizar los obstáculos en la implementación de la tecnología en el ámbito escolar. El formato de observación puede tomar la siguiente forma:

b. Guías o formatos de entrevistas a profundidad

Se debe recordar que estas entrevistas se realizan persona a persona mediante una guía de temas preestablecidos con una secuencia lógica que, no obstante, permite un amplio grado de flexibilidad durante la entrevista. Este tipo de entrevista permite abordar diferentes temas de interés según estos vayan presentándose y permite explorar temas no incluidos en el formato preestablecido pero que pueden ser de interés para el evaluador.

La principal característica de este tipo de entrevistas es que permite conocer las actitudes, motivaciones y pensamientos de las personas entrevistadas con mayor detalle, precisión y profundidad. Además, pueden surgir temas de interés no anticipados por el evaluador. El grado de utilidad de esta metodología depende mucho de la habilidad del evaluador al momento de formular las preguntas y profundizar en los temas de interés, evitando el condicionamiento de las respuestas reduciendo el sesgo de la información.

Así exista un conocimiento previo sobre los aspectos que debe abordar una entrevista, no existe un formato estándar, puesto que cada experiencia de diálogo es única. A continuación se muestra un ejemplo de guía de entrevista semiestructurada sobre el clima laboral en empresas medianas:

REGISTRO – RESUMEN DE OBSERVACIÓN GENERAL	
Estudio sobre los obstáculos para la puesta en marcha de la tecnología en el ámbito escolar.	
Fecha: _____	Hora: _____ Episodio, reunión, observación _____
Participantes: _____	
Lugar: _____	
<ol style="list-style-type: none"> 1. Temas principales. Impresiones (del investigador). Resumen de los que sucede en el evento, episodio, etcétera. 2. Explicaciones o especulaciones, e hipótesis de lo que sucede en el lugar o contexto. 3. Explicaciones alternativas. Reportes de otros que experimentan o viven en la situación. 4. Sigüientes pasos en la recolección de datos. Derivado de los pasos anteriores, qué otras preguntas o indagaciones es necesario realizar. 5. Revisión, actualización. Implicaciones de las conclusiones. 	

c. Guía de tópicos de grupos focales

Para llevar a cabo un grupo focal es necesario preparar una guía con los temas a tocar durante la discusión, ordenándolos de mayor a menor grado de generalidad, incluyendo una serie de preguntas exploratorias para generar un ambiente de confianza y así profundizar en las respuestas. La discusión es conducida en forma natural pero de manera estructurada. No es necesario respetar el orden de preguntas en la guía, sino tratar de cubrir todos los temas durante la sesión.

La guía de tópicos o temática, al igual que en las entrevistas, puede ser: semiestructurada o abierta. En la guía semiestructurada se presentan tópicos que se deben tratar, aunque el conductor del grupo focal tiene la libertad de incorporar nuevos tópicos que surjan durante la sesión e incluso alterar el orden preestablecido en la guía inicial. Finalmente, en la guía abierta, se plantean temáticas generales para cubrirse con libertad durante la sesión. A continuación se muestra un ejemplo de guía para grupo focal:

UNA GUÍA DE TÓPICOS ABIERTA SOBRE ADICCIONES

1. ¿Qué tipo de drogas (estupefacientes, sustancias) consumen los jóvenes de este barrio?
2. ¿Se consumen más bien en soledad o en grupo?
3. ¿Quién o quiénes la proporcionan?, ¿se venden?
4. ¿Cuánta cantidad consume un(a) joven cada vez que lo hace?
5. ¿Por qué la consumen? (razones motivos).
6. ¿Qué tipo de sensaciones y experiencias tienen cuando se drogan?
7. ¿Cómo definen la drogadicción?
8. ¿Qué cosas buenas y malas obtienen del consumo?
9. ¿Cómo es su vida actualmente?
10. ¿Qué esperan del futuro?

2.11 Presupuesto y cronograma

El **presupuesto** debe considerar todos los gastos necesarios para el desarrollo de la evaluación y puede ser desarrollado considerando los componentes de gasto, que son los siguientes:

- **Recursos humanos.** Se debe considerar todo el personal permanente y temporal que va a ser necesario, tomando en cuenta todas las actividades que va a demandar la ejecución de la evaluación.
- **Bienes.** Se debe considerar todo lo que va a ser necesario adquirir para desarrollar la evaluación, considerando, por ejemplo, materiales de escritorio, mochilas, grabadoras, material para el procesamiento de datos, combustibles, etc.
- **Servicios.** Es necesario considerar todos los servicios que deben ser contratados para los fines de la evaluación, por ejemplo, pasajes, alojamiento, impresiones, fletes, etc.
- **Otros gastos.** Es necesario considerar los aspectos que no pudieran incorporarse en las categorías anteriores.
- **Contingencias/imprevistos.** Se debe considerar un porcentaje, generalmente no más de 5% del presupuesto total, para afrontar gastos que no pudieran haber estado considerados en los rubros considerados.

Para cada uno de los elementos del presupuesto se debe considerar lo siguiente. Se usa como ejemplo el contrato de supervisores.

- **Costo unitario.** Es el monto que se va a pagar a los supervisores por unidad de tiempo en que van a ser contratados.
- **Periodo.** Es la cantidad de unidades de tiempo que se va a contratar a los supervisores, pueden ser meses, semanas o días.
- **Unidad de medida.** Es la unidad de tiempo en que se va a pagar a los supervisores.
- **Cantidad.** Es la cantidad de supervisores que se van a contratar.
- **Total.** Es la valorización monetaria del costo de los supervisores que se obtiene multiplicando todos los datos de los conceptos antes citados.

Una buena práctica es dividir el presupuesto por actividades, por ejemplo: planeamiento, operaciones, procesamiento, análisis, publicación. Para la elaboración de los presupuestos es muy útil emplear una hoja de cálculo como por ejemplo el Excel.

El **calendario de actividades** debe comprender cada una de las actividades o tareas que va a comprender la evaluación, las unidades de tiempo que va tomar cada actividad, la fecha de inicio y la de finalización previstas, así como los hitos que deben ser alcanzados con indicación de la fecha. Este calendario de actividades debe ser presentado en forma gráfica, lo que constituye el **cronograma**. Un programa muy utilizado para planificar proyectos y su correspondiente cronograma es el MS Project; también es usual el empleo del Excel.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 2

Antes de ejecutar las actividades aplicativas, revisa tus respuestas de las preguntas de reflexión, analízalas, interprétalas y sintetízalas con el respaldo teórico que has adquirido al desarrollar las unidades 1 y 2

Ahora procede a desarrollar la actividad aplicativa de la unidad

Recuerda que el Producto de la Sección es el diseño de una evaluación que será elaborada gradualmente

Para el diseño de la evaluación se hará uso del plan, programa o proyecto seleccionado en la actividad 1 de la unidad 1 de la presente sección.

Institución:

Nombre del plan, programa o proyecto que se evaluará:

Título de la evaluación (debe ser anotado considerando el tipo de evaluación que se realizará)

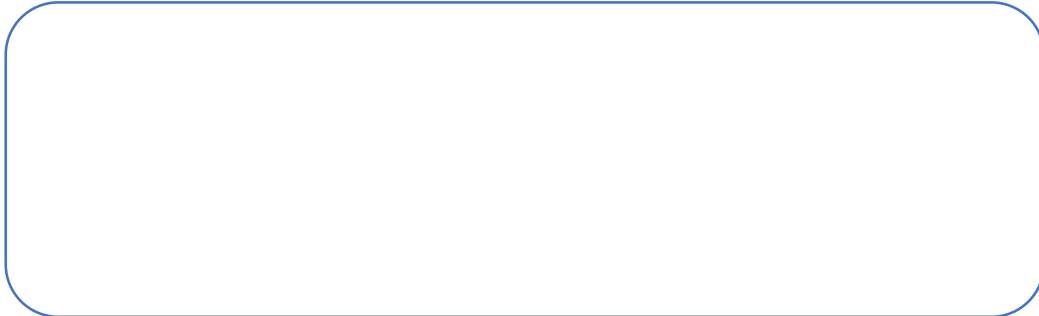
ANÁLISIS Y REFLEXIONES PREVIAS AL DISEÑO DE LA EVALUACIÓN

Antes de elaborar el diseño es necesario que intercambies con tu equipo las siguientes interrogantes:

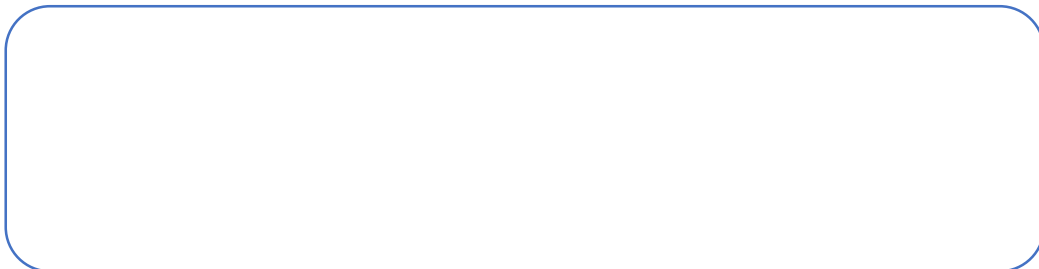
1. ¿Cuál es el objetivo principal de esta evaluación? (rendición de cuentas, aprendizajes u otro).



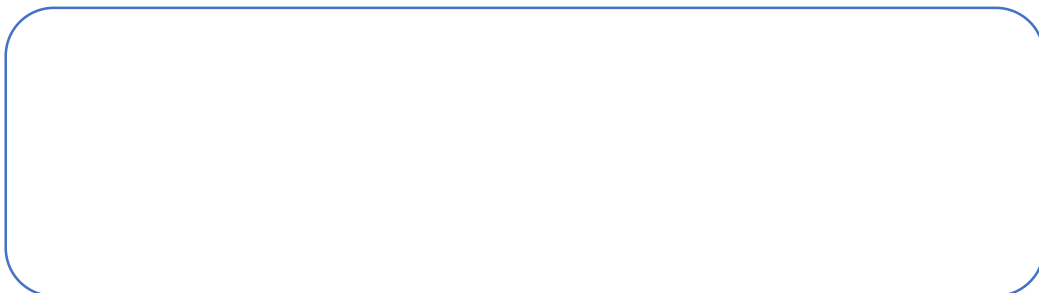
2. ¿Quiénes necesitan/están interesados en los resultados de la evaluación?



3. ¿Cuáles podrían ser las preguntas de evaluación?



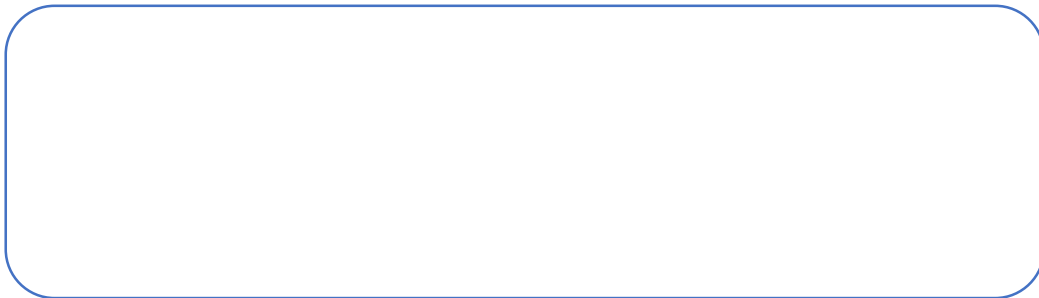
4. ¿Cuáles son los diseños que pueden responder las preguntas de evaluación?



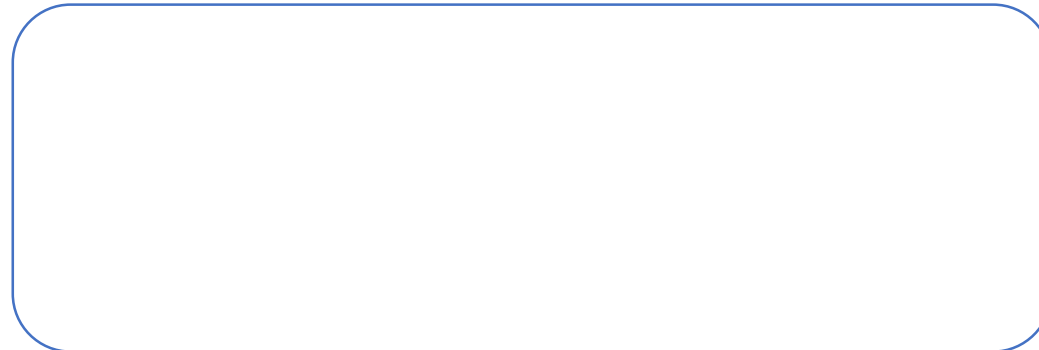
5. ¿Qué información existe?



6. ¿Cuál es el mejor diseño para obtener la información y cuáles son sus limitaciones?



7. ¿Cómo incorporar los enfoques de género, interculturalidad y participación?



Luego de esta reflexión, realiza la siguiente actividad.

ACTIVIDAD 1: Realizando el diagnóstico del contexto

1. Inicia la actividad realizando el análisis del entorno. Para ello desarrolla lo siguiente:
 - 1.1.1 Identifica las variables del entorno que puedan influir en el plan, programa o proyecto.
 - 1.1.2 Selecciona aquellas que son relevantes.
 - 1.1.3 Analiza las variables seleccionadas.
2. Solicita la tutoría respectiva.
3. Continúa con la actividad 2.

ACTIVIDAD 2: Describiendo el plan/programa/proyecto

1. Recopila y revisa minuciosamente toda la información básica del plan, programa o proyecto.
2. Puedes realizar entrevistas a personas clave y a las partes interesadas.
3. Puedes crear o elaborar herramientas que puedan identificar potenciales temas de interés para la evaluación: modelos lógicos, marcos conceptuales entre otros.
4. Elabora el informe respectivo.
5. Solicita la tutoría virtual.
6. Continúa con la actividad 3.

ACTIVIDAD 3: Identificando los objetivos/Teoría de Cambio

1. Identifica la lógica en la cual se plantearon las actividades que llevaría a cabo el plan, programa o proyecto a fin de lograr sus objetivos. Para esto aplica la Teoría de Cambio.
2. Aplica la Teoría del Cambio utilizando la siguiente matriz:

Insumos	Actividades	Productos	Resultados	Resultados finales

3. A partir de la información anotada en la matriz, analiza la lógica del plan, programa o proyecto e identifica si están adecuadamente definidos y si permiten desarrollar la verificación y/o medición en el proceso de evaluación. Si encuentras cadenas poco lógicas de cambio, por ejemplo, si los componentes del proyecto no tienen relación con el objetivo de impacto del programa, podrías cambiar el tipo de evaluación que se pretende hacer, es decir no efectuar una evaluación de impacto y hacer más bien una evaluación de desempeño o quedarte con una evaluación de la eficacia del programa.
4. Elabora el informe respectivo.
5. Solicita la tutoría virtual.
6. Continúa con la actividad 4

ACTIVIDAD 4: Identificando a los interesados

Tomando en cuenta el análisis del contexto continúa con las siguientes tareas:

1. Identifica e invita a las partes interesadas.
2. Evalúa los intereses de las partes.
3. Desarrolla estrategias para involucrar adecuadamente a los interesados.
4. Elabora el informe respectivo.
5. Solicita la tutoría virtual.
6. Continúa con la actividad 5.

ACTIVIDAD 5: Análisis de la evaluabilidad

1. Revisa el programa o proyecto seleccionado y asegúrate de contar con toda la información necesaria para realizar el análisis de evaluabilidad:
 - Diagnóstico del problema o necesidad del programa.
 - El árbol de problemas, el de medios y fines y el Marco Lógico.
 - El informe de línea de base.
 - El Plan de Monitoreo y Evaluación.

2. Define la evaluabilidad del programa o proyecto de acuerdo a la “Lista de verificación de Evaluabilidad para proyectos o programas” (UNEG), justificando su valoración en cada criterio.
3. Continúa con la actividad 6

ACTIVIDAD 6: Definiendo las preguntas de evaluación

1. Selecciona la(s) dimensión(es) que se considerará(n) en la evaluación.
2. Formula las preguntas específicas por cada dimensión (de acuerdo al tipo de evaluación).
3. Analiza las preguntas y elige las más relevantes o modifica según consideres necesario. Para ello aplica los criterios de la matriz de puntuación de las preguntas de evaluación.

La matriz debe aplicarse para cada pregunta. Ten presente que 1 significa que la pregunta cumple con lo mínimo del criterio y 5 es lo óptimo.

De acuerdo al puntaje obtenido se priorizan las preguntas.

4. Solicita tutoría virtual para analizar las preguntas formuladas y recibir la retroalimentación respectiva.
5. Continúa con la actividad 7.

MATRIZ DE PUNTUACIÓN DE LAS PREGUNTAS DE EVALUACIÓN

CRITERIOS	1	2	3	4	5
¿La pregunta responde al propósito de la evaluación?					
¿Es de interés para los principales interesados?					
¿Favorece la rendición de cuentas?					
¿La información que se obtenga permitirá tomar decisiones?					
¿Es factible responder la pregunta considerando el tiempo y el presupuesto disponible?					
¿La pregunta es clave para la credibilidad de la evaluación?					
La información o evidencias que se obtendrán al responder la pregunta, ¿permitirán proponer recomendaciones para mejorar la ejecución del plan, programa o proyecto o para determinar la factibilidad de la implementación de intervenciones similares aplicando lecciones aprendidas?					

Entrega el Producto 1: Diseño de una evaluación (Actividades 1 a 6) en la fecha indicada por el tutor.

Incorpora las recomendaciones al producto 1 y continúa con las siguientes actividades.

ACTIVIDAD 7: Definiendo el diseño de evaluación (método)

1. Define el diseño que aplicarás en la evaluación.
2. Justifica el diseño elegido.
3. De ser necesario, solicita tutoría virtual.
4. Continúa con la actividad 8.

ACTIVIDAD 8: Identificando variables e indicadores a medir/verificar y las fuentes de información

1. Identifica las variables e indicadores.
2. Verifica las fuentes de información.
3. De ser necesario, solicita tutoría virtual.
4. Continúa con la actividad 9.

ACTIVIDAD 9: Seleccionando técnicas y diseñando los instrumentos de recojo de información

1. Selecciona la (s) técnica (s) que se aplicarán en la evaluación.
2. Diseña los instrumentos de acuerdo al método definido. Asiste a la fase presencial.
3. Solicita tutoría virtual para revisar los instrumentos y recibir la retroalimentación respectiva.
4. Continúa con la actividad 10.

Entrega el producto 2: Instrumentos de recolección de información, en la fecha indicada y recibe sugerencias.

Asiste al momento presencial II llevando estos instrumentos.

ACTIVIDAD 10: Determinado la muestra

1. Participa con aportes y preguntas en el foro virtual: tamaño de efecto.
2. Identifica la población de la evaluación.
3. Determina los procedimientos de selección y tamaño de muestra.
4. Solicita tutoría virtual para revisar la muestra y recibir la retroalimentación respectiva.

ACTIVIDAD 11: Culminando el diseño de la evaluación

Elabora la matriz de consistencia, incluida más adelante, revisa y comprueba la coherencia y correspondencia entre los componentes de la matriz. Debe haber existir una relación directa, de lo contrario hay necesidad de realizar los ajustes o mejoras correspondientes.

Realiza los ajustes necesarios y elabora la versión final del diseño de la evaluación.

De ser necesario solicita tutoría virtual.

Entrega el producto 1: diseño de una evaluación en la fecha indicada.

Asimismo, ten presente que debes continuar revisando y afinando este producto, y que deberás exponer el diseño final en el momento presencial III.

MATRIZ DE CONSISTENCIA DEL DISEÑO DE LA EVALUACIÓN

Pregunta	Subpregunta	Tipo de subpreguntas (descriptiva, normativa o de causa y efecto) ⁹	Medida o indicador	Meta o estándar (Si la pregunta es normativa)	¿Tiene información de línea de base?	Fuente de información	Diseño (método)	Muestra o censo	Instrumento de recolección de datos	¿Cómo se realizará el análisis de datos?	Comentarios

Adaptado de Evaluation Design Matrix. Curso: Evaluation for Program Managers. Año 2013. USAID (Traducción no oficial). Año 2013.

⁹ **Descriptivas:** pretenden definir lo logrado por la intervención, “como una fotografía”. Usadas principalmente en evaluaciones de desempeño. Ejemplos: ¿cuáles son las actividades principales del proyecto? ¿Cuándo se ha implementado el proyecto? ¿Quiénes recibieron los servicios que brindó el proyecto?

Normativas: tratan de comparar lo alcanzado con lo planeado o con estándares. Ejemplo: ¿El proyecto alcanzó las metas planeadas? ¿El proyecto ejecutó el presupuesto planificado?

De causa y efecto: identifican si los resultados se han logrado debido a la intervención. Utilizadas en evaluación de impacto. Ejemplos: como resultado del programa de capacitación para el trabajo, ¿los participantes tienen empleos mejor pagados de los que podrían tener?

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 2

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

1. ¿Cuáles son los principales pasos para desarrollar un diseño de evaluación?

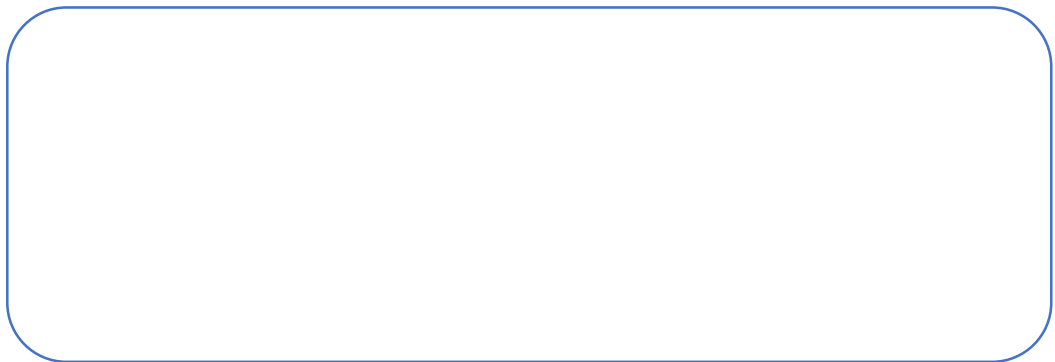
2. ¿Cuáles son los criterios para definir qué tipo de evaluación se requiere realizar?

3. ¿Cómo se define a las partes interesadas en la evaluación y qué aspectos se debe considerar para evaluar los intereses de las partes?

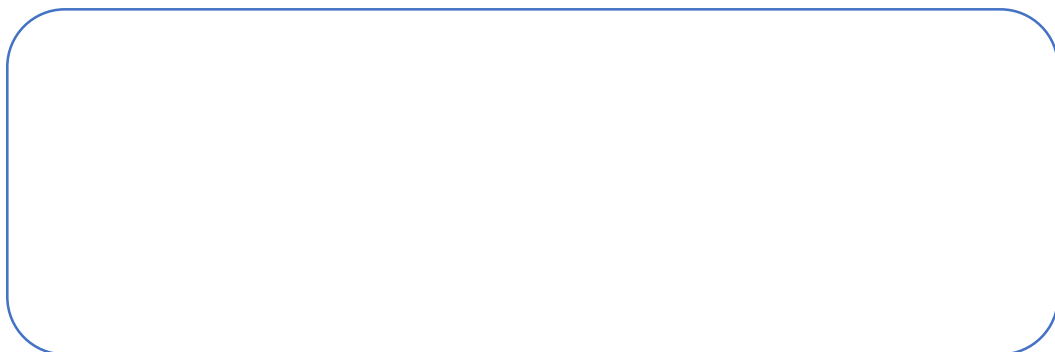
4. ¿Cuáles son los principales aspectos/información del plan, programa o proyecto que deben conocerse para iniciar su evaluación?



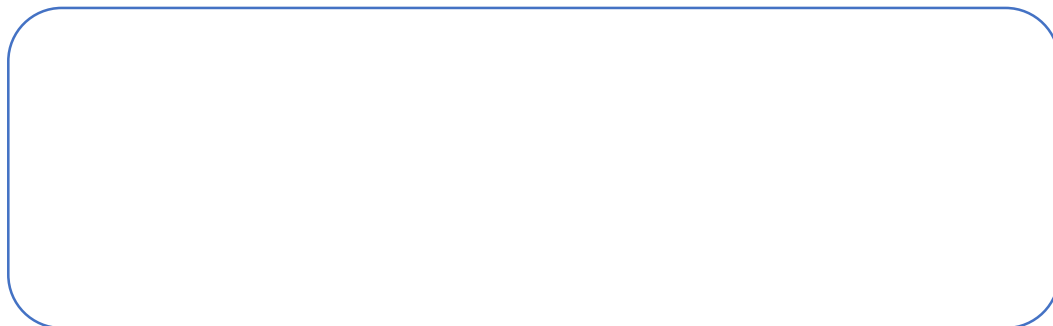
5. ¿Qué es una cadena de cambio y qué pasos se debe seguir para construirla en caso de que el plan, programa o proyecto en evaluación no la tenga explícitamente definida?



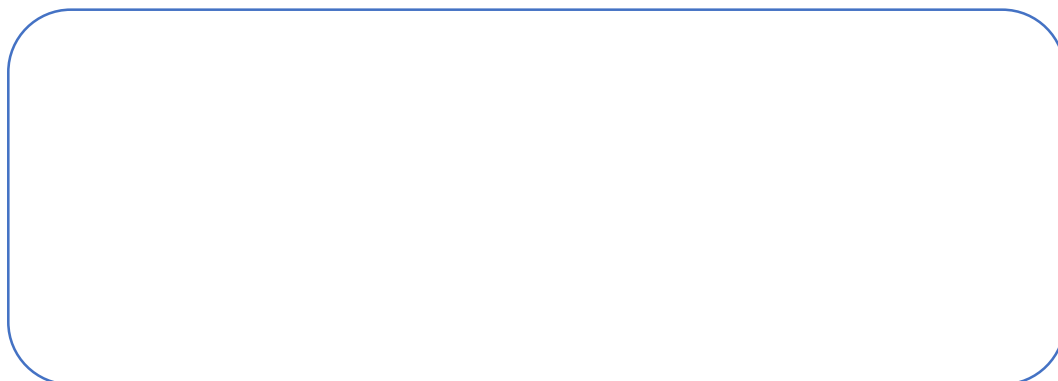
6. ¿Cuál es la principal pregunta en una evaluación de impacto y en una evaluación de desempeño?



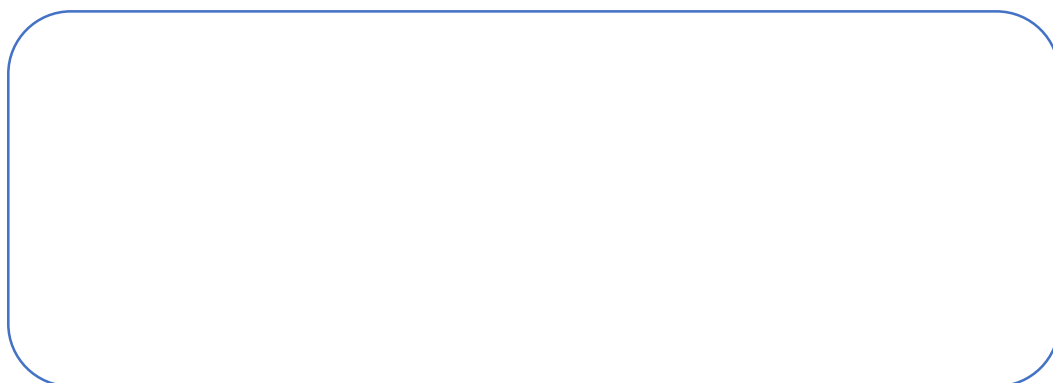
7. Explica las diferencias entre los métodos cualitativos y cuantitativos y en qué momento se definen.



8. Presenta las características de las evaluaciones experimentales y cuasiexperimentales. ¿En qué casos se utilizan y qué métodos utiliza?



9. ¿Qué es un indicador y cuáles son las principales características de un buen indicador?



10. ¿Cuáles son los principales contenidos de un plan de recolección de datos?





SECCIÓN 2

SECCIÓN 2

IMPLEMENTACIÓN DE EVALUACIONES

CAPACIDADES A LOGRAR

Aplicar técnicas de recolección de información con estándares de control y calidad.

Procesar los datos verificando su calidad.

Categorizar la información de la evaluación de acuerdo a las preguntas, objetivos y tipo de evaluación.

Identificar los hallazgos significativos en función a las preguntas y objetivos de la evaluación.

Redactar las conclusiones sintetizando y comparando las evidencias con juicios de valor y mérito.

Formular recomendaciones de acuerdo al contexto institucional, a los objetivos de evaluación e incorporando la perspectiva de los usuarios.

Formular lecciones aprendidas de acuerdo a los resultados de la evaluación.

Ejecutar el análisis crítico del informe final de una evaluación aplicando estándares.

PRODUCTOS

Informe del análisis del Informe Final de una evaluación.
(El reporte debe poner especial énfasis en el análisis de las conclusiones y de las recomendaciones de la evaluación con las propuestas de mejora)

SECCIÓN 2

IMPLEMENTACIÓN DE EVALUACIONES

CONTENIDOS

UNIDAD 1: Pautas para el desarrollo de una prueba piloto.

UNIDAD 2: Proceso de recolección de la información.

UNIDAD 3: Procesamiento de la información y consistencia de datos.

UNIDAD 4: Categorización de la información e identificación de hallazgos significativos.

UNIDAD 5: Redacción del informe de evaluación.

UNIDAD 6: ¿Cómo evaluar el informe final de una evaluación?

MATERIALES INFORMATIVOS DE APOYO

UNIDAD 1

- Ministerio de Educación. Dirección de Educación Superior Pedagógica. Sistematización del Plan Piloto de Evaluación del Desempeño Docente. Lima. Academy for Educational Development, Proyecto USAID/PER/SUMA. 2011. Págs. 24 y 25.

UNIDAD 2:

- Metodología de la investigación. Capítulo 9. Hernández Fernández Baptista, 2006.
- CERLALC. Manual de Aplicación de un Modelo de Evaluación para Identificar el Aporte de las Bibliotecas Públicas en el Desarrollo de sus Comunidades. Manual de Lecturas. Santiago de Chile. Centro Regional para el Fomento del Libro en América Latina y el Caribe (CERLALC). 2012. Capítulo 3
- Instituto de Nutrición de Centro América y Panamá INCAP. Manual de Operaciones para Recolección de Datos en un Sistema de Vigilancia Alimentario Nutricional. Guatemala, INCAP. 2012. Páginas 1 - 10

UNIDAD 3:

- Hernández, R, Fernández-Collado, C. y Baptista, P. Metodología de la Investigación. Cuarta edición. México DF. McGraw-Hill Interamericana Editores S.A de C.V. 2006. Capítulo 10.
- INE Metodología para la integración de bases de datos de la encuesta nacional de empleo e ingresos. Guatemala. Instituto Nacional de Estadística. 2003.

SECCIÓN 2

IMPLEMENTACIÓN DE EVALUACIONES

MATERIALES INFORMATIVOS DE APOYO

UNIDAD 4:

- Salavarray. L. Curso de Estadística Básica. Alto (Uruguay), Universidad de la Republica Regional Norte Sede Alto. 2000. Disponible en: http://www.cartagena99.com/recursos/matematicas/apuntes/Curso_Statistica_Basica.pdf
- Hernández R. et al. Metodología de la investigación. Capítulo 14. 2006.
- Taylor- Powell E. Y Renner M. Analyzing Qualitative Data. 2003. p. 1-12.
- Hernández R. et a Metodología de la investigación. 2006. Capítulo 14.
- Banco Mundial La evaluación del impacto en la práctica. Capítulo 13.

UNIDAD 5:

- Escorcía, O. Manual para la investigación. Guía para la formulación, desarrollo y divulgación de proyectos. Bogotá. Universidad Nacional de Colombia. 2010. p. 52-63. Disponible en: <http://www.bdigital.unal.edu.co/7551/1/manualparalainvestigaci%C3%B3n.pdf>
- Faúndez, A. Guía de Evaluación de Programas y Proyectos con Perspectiva de Género, Derechos Humanos e Interculturalidad. México. ONU Mujeres, 2014. p. 20-34.
- Comisión Europea, Dirección General de Relaciones Exteriores. Metodología de Evaluación de la Ayuda Exterior de la Comisión Europea. Francia. Comisión Europea, 2006. Volumen I Bases Metodológicas de Evaluación. p. 88-91.
- Jaramillo, Baanante & Sanz. Informe Final de Evaluación del Programa "Construyendo Perú". Ministerio de Economía y Finanzas. Lima, 2009.
- Guía Práctica para Elaborar Conclusiones. Elaboración propia de Grade (Octubre 2015).
- Ejemplo Resumen de Informe de Evaluación y Diseño de Conclusiones. Comité de Desarrollo y Propiedad Intelectual (CDIP). Noviembre 2012.
- Guía Práctica para Elaborar Recomendaciones. Elaboración propia de Grade. (Octubre 2015).
- Cómo hacer Recomendaciones Efectivas bajo la Metodología Double-Smart. APT (Association for the Prevention of Torture). Cuadernillo N°1. (Noviembre 2008).
- Banco Interamericano de Desarrollo. (Octubre 2008). Lecciones Aprendidas y Buenas Práctica. Sector de Conocimiento y Aprendizaje. p. 1-4.
- Red SIDA Perú. (2007). Marco conceptual sobre *prácticas* y lecciones aprendidas en VIH. Centro Virtual de Coordinación del Conocimiento en VIH. p. 23-30.

UNIDAD 6:

- Bamberger, M, Rugh, J y Mabry, L. Evaluación en un Mundo Real. Trabajando Bajo Restricciones Presupuestarias, Cronológicas, de Información y Políticas. Visión General. Adaptado del Capítulo 16, "Bringing it all together" de "Real World Evaluation: Working under Budget, Time, Data and Political Constraints", por Michael Bamberger, Jim Rugh y Linda Mabry. Sage Publications, 2006. p. 42-45.
- Faúndez, A. Guía de Evaluación de programas y proyectos con perspectiva de género, derechos humanos e interculturalidad. México. ONU Mujeres, 2014. p. 16-20.

Reflexionando mi práctica sobre la implementación de evaluaciones

TE INVITAMOS A REVISAR Y ANALIZAR LO SIGUIENTE:

Juan Carlos es un evaluador que tuvo el encargo de realizar la evaluación de impacto de un programa de lucha contra el analfabetismo, para lo cual definió un diseño de evaluación cuasiexperimental que requería, entre otras cosas, recoger información primaria a través de encuesta a hogares. Asimismo, antes de iniciar con la implementación de evaluación, se aseguró de contar con un plan de recojo de información en el cual había establecido el reclutamiento y capacitación a encuestadores. Hasta ese punto todo hacía prever que la información requerida para la evaluación se obtendría con éxito.

Al finalizar el diseño y con los equipo de encuestadores conformado, Juan pasó a realizar la preparación logística y seguidamente el despliegue del personal en las ciudades identificadas para la muestra; previamente se había establecido en el plan de evaluación que la selección de hogares encuestados in situ sería por conglomerados definida al azar. Sin embargo, cuando llegaron los equipos a cada ciudad, dada la fuerte etapa de capacitación al personal, Juan consideró que no sería necesario contratar equipo de supervisores. Cuando terminó el trabajo de campo y al momento del análisis de los datos, Juan decidió revisar el detalle cómo se seleccionó la muestra y descubrió que los equipos de trabajo de campo, en las ciudades seleccionadas, decidieron aplicar las encuestas a los hogares situados alrededor de la plaza central de cada ciudad y por consiguiente la información recopilada no servía para hacer inferencias por ciudad. Como los datos no servían para la evaluación, se tuvo que recoger de nuevo la información de campo.

Preguntas de reflexión:

¿Por qué es importante implementar la evaluación conforme a la metodología establecida en el diseño?

¿Cuál pudo ser el rol del plan piloto para evitar este tipo de errores en la aplicación de los instrumentos?

Ahora que culminaste las actividades de reflexión, desarrolla los contenidos y actividades de la unidad.

Si necesitas apoyo de los tutores, comunícate con ellos a través de la plataforma virtual.



UNIDAD DE APRENDIZAJE 1: PAUTAS PARA EL DESARROLLO DE UNA PRUEBA PILOTO

La prueba piloto es una etapa importante en el proceso de evaluación. Permite comprobar la utilidad y aplicabilidad de los instrumentos de recojo de información definidos en la evaluación como cuestionarios de encuestas, guías de grupos focales y fichas de entrevistas; esto considerando que el diseño de los instrumentos de evaluación en sí mismo no garantiza que estos recojan de manera fidedigna la información. En ese sentido, es necesario poner a prueba estas herramientas para mejorar su diseño e incrementar su efectividad a través de una prueba piloto.

De otro lado, la prueba piloto consiste en aplicar los diferentes instrumentos de recojo de información de la evaluación a un grupo reducido de personas lo más homogéneas y parecidas a los individuos de la muestra que serán objeto de estudio¹⁰ pero que no van a formar parte de ella. Ello permitirá hacer modificaciones a los instrumentos de tal manera que permitan incluir no solo información sobre los aspectos de la evaluación en sí, sino las posibles dificultades e inconvenientes producto de la aplicación de los mismos. De este modo, el fin de la prueba piloto es evaluar aspectos técnicos de los instrumentos y corregir y mejorar todas las deficiencias que se identifiquen durante el proceso de la aplicación de los mismos (Iaraossi, G.; 2006).

1.1. Desarrollo de la prueba piloto

La aplicación de la prueba piloto permite identificar las principales fuentes de donde pueden provenir los errores en la ejecución de las herramientas de recojo de información sobre la muestra total. Por lo general, estas fuentes son a) el ambiente de recolección, b) la situación de los entrevistados y c) los responsables de recojo de información.

En primer lugar, el ambiente de recolección constituye una de las principales fuentes que generan error en la medición y aplicación de instrumentos, puesto que de ello depende la comodidad y el estado de ánimo de los encuestados. El ambiente de recolección debe ser adecuado, puesto que factores como la temperatura, dimensiones y privacidad del mismo, así como la actitud del

¹⁰ Por lo general, el número de individuos se encuentra entre 20 y 30. En estudios de mayor envergadura, el número de individuos será un porcentaje de la muestra final, el cual está alrededor del 10%.

encuestador, pueden influir de manera significativa en los resultados de la evaluación.

En segundo lugar, una situación en la cual el entrevistado sienta hambre, cansancio, hostilidad hacia la intervención, falta de tiempo o inseguridad, son también fuentes de sesgo de la información.

Finalmente, si la persona encargada del recojo de la información no conoce bien el procedimiento de implementación de la herramienta, no considera los objetivos de la evaluación, no percibe de manera adecuada lo que los individuos tienen que transmitir o simplemente no genera empatías al momento del contacto con los individuos, es muy probable que la información que reporte no sea reflejo del objeto de estudio (Fox, J. David; 1987).

En la práctica, la información resultante de la prueba piloto puede dividirse en dos grupos:

- i) Información acerca del proceso de recojo de información, como por ejemplo: la hora más adecuada de aplicar los instrumentos, la duración de la aplicación, las reacciones que tienen los entrevistados, entre otros.
- ii) Información acerca de la herramienta de recojo, que brinda datos sobre los tipos de preguntas aplicados, el orden en el que se encuentran, la claridad con la que son formuladas las preguntas, el diseño del instrumento, etc.

A continuación se muestra el cuadro 16, donde se muestran las principales preguntas que guían el proceso de la prueba piloto.

Cuadro 16. Lista de control para el proceso de la prueba piloto

- ¿Los encuestados comprenden el objetivo de la encuesta?
- ¿Los encuestados se sienten a gusto al responder las preguntas?
- ¿La redacción de los cuestionarios es clara?
- ¿La referencia del tiempo es clara para los encuestados?
- ¿Las opciones de respuesta se relacionan con las experiencias de los encuestados en el tema de la pregunta?
- ¿Algunos de los puntos los obligan a pensar demasiado antes de responder? Si es así, ¿cuáles?
- ¿Qué puntos les producen enojo, molestia o confusión?
- ¿Algunas de las preguntas tienen influencia sobre la respuesta? Si es así, ¿cuáles?
- ¿Las respuestas recopiladas reflejan la información que necesita con respecto al objetivo de la encuesta?
- ¿Las respuestas recibidas son lo suficientemente variadas?
- ¿El cuestionario es demasiado extenso?
- De acuerdo con sus participantes de prueba, ¿se pasaron por alto otros temas importantes?

Fuente: Laraossi (2006).

Asimismo, las preguntas guía de la prueba piloto pueden ser agrupadas y analizadas con respecto a varias dimensiones que se presentan en el cuadro 17.

Cuadro 17. Dimensiones y valoración de las preguntas guía de la prueba piloto

DIMENSIONES	ASPECTOS DE VALORACIÓN
Sobre el lugar de estudio	<ul style="list-style-type: none"> • Seguridad del lugar donde se aplica el instrumento. • Ambiente para la aplicación del instrumento (comodidad, privacidad).
Autorización	<ul style="list-style-type: none"> • Autorización de las autoridades y los participantes (encuestados, entrevistados, participantes de grupos focales) para el recojo de información.
Tiempo	<ul style="list-style-type: none"> • Tiempo requerido para ejecutar los procesos de aplicación de instrumentos. • Tiempo de traslado.
Capacitación de sujetos de estudio	<ul style="list-style-type: none"> • Acceso a los sujetos de estudio (encuestados, entrevistados, participantes de grupos focales). • Porcentaje de rechazo del instrumento.

DIMENSIONES	ASPECTOS DE VALORACIÓN
Instrumentos	<ul style="list-style-type: none"> • Claridad de las preguntas de los instrumentos (lenguaje, redacción, ambigüedades). • Compresión de las preguntas.
Procedimientos	<ul style="list-style-type: none"> • Orden de las preguntas. • Orden de las opciones.
Coordinación y supervisión	<ul style="list-style-type: none"> • Claridad de las instrucciones. • Incongruencias. • Secuencia lógica. • Claridad de los participantes en cuanto a tareas y responsabilidades.

Fuente: CONALEP (1984).

1.2. Validación de los instrumentos

La validación de los instrumentos se realiza a partir del análisis de los siguientes aspectos: i) validez, ii) confiabilidad, iii) adecuada amplitud del rango, iv) sensibilidad al cambio y iv) utilidad. Como podrá reconocer, varios aspectos se responderán a partir de la información resultante de aplicar el piloto (Iaraossi; 2006):

⇒ **Validez.** Es el grado en que el instrumento mide aquello que se pretende medir. Por lo general, la validez de un instrumento es determinada por un grupo de expertos que evalúa si el instrumento tiene las cualidades deseadas. Asimismo, la validez se puede hacer sobre la base de un constructo teórico que evalúa si el instrumento refleja el concepto o fenómeno que mide.

⇒ **Confiabilidad.** Es la capacidad de los instrumentos de proveer los mismos resultados en mediciones sucesivas bajo las mismas características. Es también el nivel en que los diferentes ítems o preguntas de los instrumentos están relacionados entre sí. Ello indica el grado de acuerdo entre los mismos.

⇒ **Adecuada amplitud del rango.** Es la capacidad del instrumento para medir todas las características del fenómeno de interés. Por ejemplo, si se desea evaluar el bienestar de un individuo, es necesario incluir tanto aspectos físicos como emocionales, entre otros.

⇒ **Sensibilidad al cambio.** Es la habilidad del instrumento para encontrar las variaciones o cambios que presenta el fenómeno de interés. Para evaluar la

sensibilidad al cambio, es necesario aplicar el instrumento de manera repetida al sujeto de interés.

⇒ **Utilidad.** Es el grado de aplicabilidad del instrumento en la vida real. Es decir, es probar el instrumento en las condiciones reales en las cuales se va a recolectar la información del objeto de estudio y evaluar su pertinencia y adaptabilidad en tales situaciones.

En síntesis, la implementación de la prueba piloto permite poner a prueba el instrumental con el que se cuenta y evaluar su idoneidad para responder a los objetivos de la evaluación. A continuación se muestra un ejemplo aplicado al ámbito de la educación.

Algunas partes del cuestionario pueden contener ítems que se incorporan para la medición de cualidades no directamente observables (constructos), como por ejemplo la inteligencia, el conocimiento sobre la prevención del VIH, etc., donde los ítems consideran aspectos si no observables, estrechamente relacionados con el constructo que se trata de medir, y tienen puntuaciones que reflejan la intensidad con se presentan. Luego se aplican escalas aditivas para reflejar los grados del constructo medido. En estos casos la validación sigue un procedimiento científico que requiere conocimientos más avanzados de estadística como el Alpha de Cronbach, que es una herramienta muy usada en estos casos.

Cuando se hace uso de constructos, la validación debe hacerse por separado antes de la prueba piloto.

Poniendo a prueba los instrumentos de la Evaluación del Desempeño Docente

En los últimos años, en la Dirección de Educación Superior Pedagógica del Ministerio de Educación del Perú, se ha venido impulsado acciones de evaluación de desempeño docente a nivel piloto. Durante el año 2009 se llevó a cabo un piloto de evaluación que tuvo como objetivos: a) explorar la aplicación de algunas estrategias metodológicas derivadas de los criterios y factores establecidos en la Ley del Profesorado en lo referido a la carrera pública magisterial y su reglamento y b) registrar algunas percepciones de los docentes y directivos sobre los contenidos y alcances de la evaluación del desempeño docente previsto en la nueva carrera pública magisterial que faciliten su implementación.

En ese sentido, se aplicó una serie de herramientas dentro de las cuales destacan:

i) Ficha de autoevaluación de la práctica pedagógica. A través de la ficha de autoevaluación el docente valoraba los principales aspectos de su práctica pedagógica: planificación de la enseñanza, contenido del área curricular, estrategias metodológicas durante del desarrollo de la clase, clima para el aprendizaje y evaluación de los aprendizajes. Para los docentes participantes en el piloto, la autoevaluación ha constituido una oportunidad para revisar su trabajo y reflexionar sobre sus prácticas.

*"La autoevaluación me hizo pensar cuando iba revisando los indicadores, esto no estoy haciendo decía. Llenar la ficha te obliga a recapacitar".
(Docente evaluada, Trujillo).*

ii) Ficha de observación de la sesión de aprendizaje. Se utilizó para obtener información y comprobar el nivel de efectividad alcanzado por el docente en relación a la gestión de los procesos de enseñanza y aprendizaje en el aula. Los aspectos a observar fueron: el dominio del currículo, la metodología de enseñanza, el manejo de los procesos de evaluación y el dominio de los contenidos disciplinares y pedagógicos.

En su aplicación se definió que la observación sería inopinada y opinada a fin de tener dos referentes de información. A través de la inopinada se tendría una observación real del desempeño puesto que no se ha prevenido al docente sobre cómo y cuándo lo visitarían. De esa manera, el evaluador tendría la oportunidad de observar el comportamiento de los docentes en el aula tal como se desarrolla cotidianamente. La observación opinada le permitiría al docente prepararse para ser observado para luego poder demostrar lo mejor que sabe hacer. La observación se haría por un periodo de 45 minutos. En la práctica, la observación no siempre fue realizada solamente por el evaluador conforme se había indicado. En algunos colegios participaron además del evaluador el director o subdirector, o el padre de familia, miembro del comité. En el caso de estos últimos, su participación fue como veedores.

iii) Ficha de supervisión. A través de este instrumento, el directivo (director o subdirector) de la institución educativa supervisaba la práctica educativa en lo que se refiere a contribución a los objetivos del desarrollo institucional y las innovaciones pedagógicas que se realiza en el aula y en la institución educativa. Su evaluación se apoyaba en la verificación de algunas evidencias que demandaba el instrumento. En las instituciones educativas donde no existía el cargo de subdirector, la supervisión de la práctica pedagógica estuvo a cargo del director. Como la valoración de los indicadores tenía que sustentarse en evidencias, tuvieron algunas dificultades, dado que no siempre se contaba con ellas.

*"Las evidencias que exige el instrumento hacen el trabajo engorroso"
(director de Trujillo).*

"El instrumento no es tan funcional. Se hace muy burocrático. Se presenta como algo más administrativo. Es tedioso" (directora Trujillo).

Fuente: USAID (2011).

Para ampliar el tema tratado, revisa la siguiente lectura:

- Ministerio de Educación. Dirección de Educación Superior Pedagógica. Sistematización del Plan Piloto de Evaluación del Desempeño Docente. Lima. Academy for Educational Development, Proyecto USAID/PER/SUMA. 2011. p. 24 y 25. Disponible en: http://www.sumaeducacion.pe/pdf/plan_piloto.pdf

ACTIVIDADES APLICATIVAS DE LA UNIDAD 1

ESTAS ACTIVIDADES SON OPCIONALES
Su ejecución depende del interés del participante.



ACTIVIDAD 1: Realizando la validación de los instrumentos de evaluación

1. Identifica profesionales especialistas o expertos en el tema de la evaluación que estás realizando.
2. Solicita su opinión técnica acerca de la precisión de los contenidos, es decir, si el instrumento mide aquello que afirma medir y si están ausentes algunos temas importantes.
3. Realiza los ajustes necesarios a los instrumentos.

ACTIVIDAD 2: Planificando y organizando la prueba piloto

1. Identifica el lugar y grupo objetivo donde realizarás la prueba piloto.
2. Gestiona la autorización correspondiente para ejecutar la prueba piloto.
3. Estima el tiempo necesario (abarca el tiempo para la prueba y el traslado).
4. Identifica y organiza al equipo que realizará la prueba piloto.
5. Capacita al grupo que ejecutará la prueba piloto, incluyendo principalmente los siguientes aspectos:
 - Contenidos y uso de instrumentos.
 - Procedimientos: orden de las preguntas y orden de las opciones.

ACTIVIDAD 3: Ejecutando la prueba piloto

1. Coordina y supervisa la aplicación de los instrumentos.
2. Una vez realizada la prueba piloto en campo, recoge y organiza la información en dos grupos:

2.1 Información acerca del proceso de recojo de información:

- *Ambiente de recolección.* Por ejemplo: la privacidad, las características físicas del ambiente, otros.
- *Situación de los entrevistados.* Por ejemplo: el estado de ánimo y físico del entrevistado, los sentimientos que tenga hacia la intervención, falta de tiempo, inseguridad, reacciones de los entrevistados, entendimiento del objetivo del instrumento, otros.
- *Los responsables del recojo de la información.* Por ejemplo: el conocimiento de los procedimientos para la aplicación de los instrumentos, percepción adecuada de lo que transmiten los entrevistados, la empatía con los entrevistados, duración de la aplicación, errores más comunes cometidos, entre otros.

2.2 Información acerca de los instrumentos.

- *Preguntas aplicadas.* Por ejemplo: aceptación o incomodidad de las preguntas por parte de los entrevistados, claridad en la formulación y en la redacción, relación entre las opciones de respuestas y las experiencias de los entrevistados en el tema de la pregunta, preguntas que generan pensar demasiado al entrevistado, preguntas que generan incomodidad, enojo, molestia o confusión, si las respuestas reflejan la información que se necesita con respecto al objetivo del instrumento, preguntas que tienen influencia sobre las respuestas, etc...
- *El orden en el que se encuentran las preguntas.* Por ejemplo: si permite agilidad en la aplicación del instrumento.
- *El diseño del instrumento.* Por ejemplo: extensión del instrumento, facilidad del diligenciamiento, colores apropiados, tamaño de letra, organización por temas, espacios suficientes para las preguntas

abiertas, sistemas de codificación, etc. Analiza la información y realiza las mejoras y ajustes respectivos:

- Identifica las principales razones de los errores en la aplicación de los instrumentos.
- Procede a realizar las mejoras y los ajustes de:
 - El diseño de los instrumentos.
 - Los procesos y procedimientos para la aplicación de los instrumentos.

ACTIVIDAD 4: Verificando la validez y confiabilidad de los instrumentos

1. Considerando los resultados del juicio de expertos, verifica la validez de los instrumentos y elabora la versión final.
2. Considerando los resultados de la prueba piloto y aplicando las pruebas estadísticas respectivas verifica la confiabilidad de los instrumentos.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

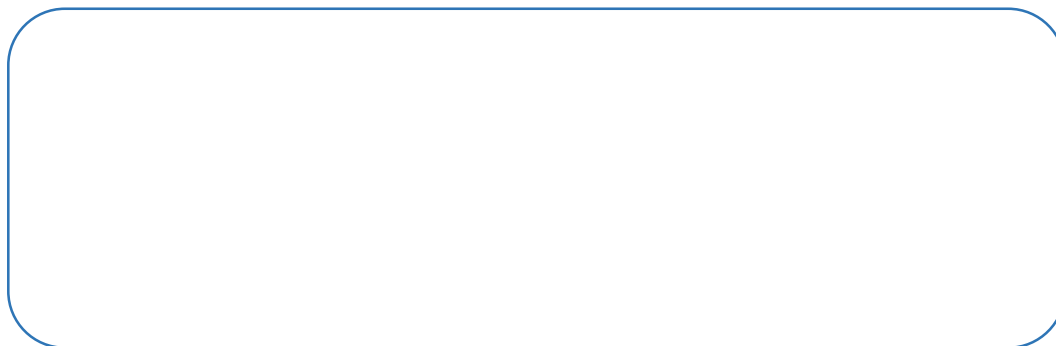
Responde lo siguiente:

1. ¿Por qué es importante la prueba piloto en un estudio de evaluación?

2. ¿Cómo se determina la validez y confiabilidad de los instrumentos luego de la aplicación de la prueba piloto?

3. Describe las principales etapas del proceso de la prueba piloto.

4. ¿Qué aspectos de la implementación de la prueba piloto no deben dejar de documentarse?





UNIDAD DE APRENDIZAJE 2: PROCESO DE RECOLECCIÓN DE LA INFORMACIÓN

Tras haber identificado y diseñado los instrumentos de recolección de información como parte del diseño de evaluación y habiendo validado los instrumentos mediante la prueba piloto, ahora se cuenta con las versiones finales de los instrumentos, con lo cual pasaremos a la siguiente etapa de la evaluación, la denominada etapa de recolección de información.

Para la implementación de esta etapa se requiere conocer tres aspectos fundamentales que nos permitan obtener la información de manera más fidedigna posible. Un primer paso en esta etapa es reclutar y capacitar al personal que ejecutará el levantamiento de información, brindar las instrucciones adecuadas, determinar la carga de trabajo y conformar los equipos. Luego de ello se debe programar a detalle las actividades diarias y dinámicas para el periodo de recojo de información según lo planificado en líneas generales durante el diseño de información y, finalmente, salir al campo.

En esta unidad se brinda criterios y pautas, así como una breve explicación de las técnicas específicas que se aplican para recoger tanto información cuantitativa como cualitativa, y que le permitirán recabar de manera confiable las características, percepciones y opiniones de la población en estudio.

Para una mejor comprensión, la presente unidad se ha desarrollado tomando como ejemplo la evaluación de impacto del programa de bibliotecas públicas en Chile, la cual tuvo como objetivo estimar los impactos de largo plazo que tienen las bibliotecas públicas en el desarrollo de sus comunidades. En esta evaluación fueron utilizados instrumentos cuantitativos y cualitativos. En la primera categoría se ejecutó una encuesta a una muestra de usuarios y no usuarios de las bibliotecas y entre los instrumentos cualitativos se aplicaron grupos focales a usuarios, entrevistas semiestructuradas a actores clave y observación participante sobre el funcionamiento habitual de la biblioteca (CERLALC; 2011). De esta experiencia se analizarán las técnicas de recolección de información aplicadas a fin de ilustrar el proceso de recolección en campo.

2.1 Plan de recolección de datos

Habiendo definido todos los aspectos básicos de la evaluación se debe elaborar un plan de recolección de toda la información cuantitativa y cualitativa. A continuación se presenta un ejemplo del contenido de un plan de recolección de datos de una evaluación nutricional.

Esquema de un plan de trabajo de campo

1. Objetivos

1.1 General

1.2 Específicos

2. Características y conceptos técnicos de la encuesta

2.1 Cobertura geográfica

2.2 Cobertura temporal

2.3 Población en estudio

2.4 Unidad de investigación

2.5 Método de recolección

2.6 El proceso de la encuesta

2.7 Definición de variables e indicadores para la encuesta

2.8 Verificación de campo

3. Organización

Organigrama de los cargos y jerarquías para el trabajo de campo

4. Funciones del personal de campo

4.1 Jefe de operación de campo

4.2 Supervisor/a

4.3 Entrevistador/a

4.4 Antropometrist/a

4.5 Auxiliar

5. Tareas de trabajo de campo

5.1 Tareas ejecutadas antes del trabajo de campo

5.1.1 Del jefe de la operación de campo

5.1.2 De la supervisor/a

5.1.3 De la entrevistador/a

5.1.4 De la antropometrista

5.1.5 De la auxiliar de antropometría

5.2 Tareas ejecutadas durante del trabajo de campo

5.2.1 Del jefe de la operación de campo

5.2.2 De la supervisor/a

5.2.3 De la entrevistado/a

5.2.4 De la antropometrista

5.2.5 De la auxiliar de antropometría

5.3 Tareas ejecutadas después del trabajo de campo

5.3.1 Del jefe de la operación de campo

5.3.2 De la supervisor/a

5.3.3 De la entrevistador/a

5.3.4 De la antropometrista

5.3.5 De la auxiliar de antropometría

6. Programación de rutas de la operación de campo

Equipo de entrevistadoras	Lugares	Fechas de permanencia	Carga de trabajo asignada
Equipo 1	Ancón		
	Huarmey		
Equipo 2	Lunahuaná		
	Quilmaná		

Para ampliar el tema tratado revisa la siguiente lectura:

- FAO. Realización de encuestas nutricionales en pequeña escala. Manual de campo. Sin fecha. Disponible en:
<https://books.google.com.pe/books?id=Tb-kRdZrgIAC&pg=PA18&lpg=PA18&dq=planificar+encuestas&source=bl&ots=buYwh-SVrf&sig=qfaNfdo8bERDruwx8DTwi8HfO9A&hl=es&sa=X&ei=rGbRVMe0J5DCsATdjoLAAQ&ved=0CEMQ6AEwBzgK#v=onepage&q=planificar%20encuestas&f=false>
 (p. 93-108).

2.2 Capacitación del personal

Las personas del equipo de investigación que van a aplicar el instrumento, se trate de entrevistadores, observadores, encuestadores o encargados de dirigir las sesiones para aplicarlo, deben recibir un entrenamiento en varios aspectos:

- La importancia de su participación y su papel en el estudio.
- El propósito del estudio.
- Las personas e instituciones que patrocinan y/o realizan la investigación.
- La calendarización y tiempos de estudio (fechas clave de entrega de información).
- La confidencialidad que deben guardar de la investigación y sobre todo de los participantes. Esta no puede romperse bajo ninguna circunstancia.
- La manera como deben vestirse si van a entrevistar o tener contacto con los participantes.
- El respeto y sensibilidad que deben tener hacia los participantes (por ejemplo, no discriminar ni utilizar lenguaje que pudiera resultar ofensivo).
- No discutir ni molestarse con los participantes.
- Establecimiento de confianza.
- Uso de credenciales y cartas que los identifiquen.

- El instrumento de medición (estructura, orden, ítems, instrucciones para el personal y para los sujetos participantes) y las condiciones de administración.

Durante el entrenamiento, el personal que habrá de aplicar y/o calificar las respuestas o mediciones debe realizar varios ensayos a la vista de todos (de entrevista, observación, aplicación de prueba, etc.), para así corregir aquellas situaciones que puedan afectar la objetividad y el correcto desempeño en la recolección de datos.

El entrenamiento puede llevarse a cabo mediante una o más sesiones, según sea la extensión del cuestionario. Durante la sesión o sesiones se motiva al personal, se escuchan sus sugerencias y se responden inquietudes (Hernández et al 2006). En la fase piloto debe programarse una cantidad suficiente de entrevistas por encuestador como para que se familiarice con el cuestionario y luego garantizar agilidad en la aplicación real.

2.3 Aplicación de técnicas de recolección de datos

Las consideraciones, pasos y/o criterios a considerarse durante el recojo de la información se diferencian ampliamente de acuerdo a la naturaleza de los instrumentos (cuantitativos y cualitativos).

2.3.1 Aplicación de instrumentos cuantitativos

Durante el proceso de recolección de información cuantitativa, que usualmente se realiza mediante la aplicación de cuestionarios de encuestas, debe tenerse en cuenta lo siguiente:

a. Selección de personas a encuestarse

En el diseño de la muestra debe haberse definido los procedimientos a seguir para la selección de los informantes. Por ejemplo, si el diseño comprende dos etapas, en una primera etapa deben ser seleccionadas manzanas con muestreo simple al azar y luego viviendas, en una segunda etapa, con el método de conglomerados compactos, entonces se debe disponer de mapas para identificar las manzanas seleccionadas; luego, al llegar a ellas, se debe aplicar el criterio establecido en el diseño de la muestra para identificar el punto de inicio de las viviendas que van a ser abordadas.

A continuación se presenta la forma como se seleccionó a los informantes en el ejemplo de las bibliotecas. En primer lugar, la selección de personas a encuestar se realiza de dos maneras: para los usuarios de las bibliotecas se hace uso de muestreo no probabilístico, en el cual se establecen las cuotas definidas para la muestra (por ejemplo: cuotas de personas según edad y/o sexo, entre otras) que se pretende analizar en el estudio. En el caso de los no usuarios se hace uso del muestreo probabilístico aleatorio simple, en el que se define el procedimiento de selección de usuarios a encuestar.

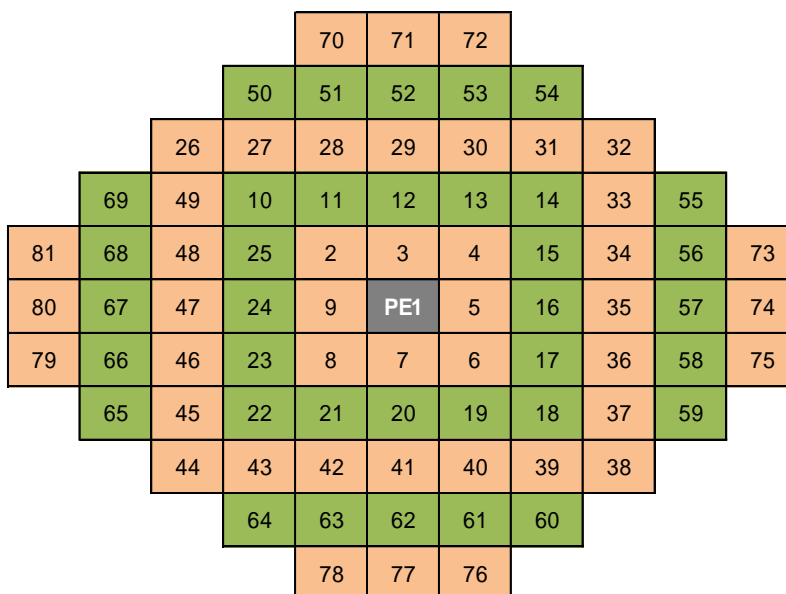
La aplicación de las encuestas es realizada respetando el tipo de muestreo: por ejemplo, en el caso de la encuesta para la evaluación de bibliotecas públicas, la aplicación de cuestionarios a la muestra de *usuarios* se realizó en los locales de las propias bibliotecas, puesto que es ahí donde se concentran estos; resultó ser un lugar cómodo y apropiado. En tanto, en el caso de la muestra de *no usuarios* las encuestas se aplicaron principalmente en hogares, tal como se aplica una gran parte de encuestas. Para estos casos se siguen los siguientes pasos como parte de la selección de personas a encuestar:

Usualmente las encuestas son aplicadas en hogares utilizando un diseño de trabajo organizado basado en un punto de referencia y sus manzanas aledañas. Para ello se siguen los siguientes pasos. A manera de ejemplo, se utilizan los datos de la encuesta aplicados a no usuarios de las bibliotecas; las letras "n1, n2,..." hacen referencia a que se dicho número se debe determinar en función de la muestra de cada estudio en particular:

- Puntos de referencia: se sortea 1 punto de referencia equivalente a una manzana base y 80 (n1) manzanas adicionales en base a un radio de 5 (n2) manzanas a la redonda. La manzana base debe mantener un distancia de alejamiento de al menos 2 (n3) cuadras de la biblioteca en donde se realizarán las encuestas.
- Se podrá realizar un máximo de 10 (n4) encuestas por manzana.
- El avance se realiza en el sentido del reloj y por anillos de manzanas. Tal como lo muestra la figura 5, cada cuadrado en la figura representa una manzana. Los números de las manzanas muestran el orden en que se debe realizar la encuesta, es decir, siguiendo el sentido de las agujas del reloj.

- Los colores identifican los anillos de manzanas. Una vez que ya se han recorrido todas las manzanas de un anillo, se continúa desde la esquina superior izquierda del anillo siguiente.
- El modelo de muestreo a utilizar es por selección aleatoria de manzanas.
- La primera selección aleatoria de la muestra corresponderá a la selección aleatoria de manzanas.
- Una vez seleccionadas las manzanas, se realizará un barrido de las viviendas de dicha manzana en busca de las personas que sean del grupo objetivo anteriormente definido.

Figura 5. Modelo para aplicar encuestas a no usuarios



PE1: Punto de Encuestación 1

Fuente: CERLALC (2011).

b. Indicaciones generales para la aplicación del cuestionario

Hay una serie de normas que el encuestador debe recordar para hacer un trabajo efectivo y eficiente. Ellas son las siguientes:

Favorecer el ambiente para la aplicación

El medio usual es acercarse a la vivienda y solicitar al jefe de hogar, o a su cónyuge –si la tiene– su cooperación al estudio accediendo a ser entrevistado. Al hacer esta solicitud, lo mejor es no decir el tiempo que se demorará; si la persona pregunta, se le puede dar un estimado como “no mucho”. Se ha comprobado que algunos entrevistados dicen no tener tiempo pero cooperan si el entrevistador/a coloca al sujeto rápidamente en situación de entrevista.

Algunos de los problemas que puede enfrentar el entrevistador/a para la aplicación de cuestionarios son los siguientes:

- El entrevistado está demasiado ocupado, entonces se debe acordar una cita y cumplirla.
- El entrevistado no se encuentra en casa: si la persona a quien se debe entrevistar no está, pero algún miembro del hogar sí, el entrevistador/a también podrá arreglar una cita; de ninguna manera reemplazar al sujeto de estudio, a menos que el diseño de la muestra considere esta posibilidad.
- Negativa a cooperar: el entrevistador/a deberá usar toda su destreza y capacidad de persuasión para aplicar el cuestionario, recalando la importancia de su participación, pero sin prometer nada a cambio. Sin embargo, si la persona insiste en la negativa, se debe retirar educadamente e informar a su supervisor.

Presentar el estudio

El punto específico que se debe cuidar es la explicación del estudio, que debe ser breve, honesta y seria, recalando la importancia de la información que pueda brindar el entrevistado.

Normas para aplicar las preguntas

- *Ajustarse a las instrucciones recibidas:* las instrucciones que se reciben al inicio del trabajo obedecen a un plan cuidadosamente elaborado. Cada

pregunta en el cuestionario tiene su razón de ser. De existir alguna duda, es importante que el encuestador durante la capacitación o luego de la prueba piloto haga todas las preguntas que sean necesarias para aclararlas antes de salir al terreno.

- *Estudiar los cuestionarios:* estudiar con detención todas las preguntas del cuestionario hasta entender claramente sus significados. Un buen encuestador debe saber lo que está preguntado y el tipo de respuesta válida para los fines del estudio. Esto se traducirá a su vez en un mejor entendimiento de las preguntas por parte del entrevistado.
- *Hablar en forma clara:* al hacer las entrevistas el encuestador debe usar un tono de voz audible, apropiado y respetuoso con el encuestado, y acorde a la situación en la que se aplica la encuesta.
- *Respetar el orden de las preguntas:* el cuestionario sigue un orden previamente diseñado que específicamente controla la posibilidad de que ciertas preguntas influyan en la respuesta de preguntas posteriores.
- *No sugerir respuestas:* el encuestador debe mantenerse neutral frente al encuestado y evitar manifestar su propia opinión, creencias o posiciones valóricas antes y durante la aplicación de la encuesta. El encuestador no debe juzgar las respuestas del entrevistado y debe propiciar que el encuestado se sienta en absoluta libertad para manifestar su opinión.
- *Ser amable y empático con el encuestado, pero no exagerar:* el encuestador debe procurar ser amable con el encuestado durante la aplicación del cuestionario y al final debe agradecerle por su tiempo y disposición para contestar. Asimismo debe mostrarse receptivo y brindar comodidad al encuestado.
- *Escribir claramente las respuestas:* las respuestas debe ser marcadas o escritas de tal manera que puedan ser identificadas o leídas claramente por los digitadores. Es importante considerar que ante cualquier duda que exista sobre una respuesta se puede anular el cuestionario.
- *Hacer todas las preguntas:* por ningún motivo una pregunta puede quedar en blanco. Es importante asegurarse de hacer todas las preguntas y de anotar todas las respuestas (o anotar que el entrevistado no quiere o no pudo responder una pregunta, si ese es el caso).

- *Reserve la confidencialidad del cuestionario:* el cuestionario no debe mostrarse al entrevistado en ningún momento, porque puede influenciar la respuesta a alguna pregunta que se haga más adelante. También es importante hacer saber al encuestado que la encuesta es *anónima*, por lo que *sus respuestas son confidenciales* y no serán conocidas por nadie más que por el equipo de investigación que analizará los datos.

Errores comunes en el registro de las respuestas

Los errores que se cometen en el registro de la respuesta dependen del tipo de pregunta. Los principales errores en el caso de respuestas abiertas son:

- *Registro incompleto de la respuesta* debido a la amplitud que pueda expresar la persona entrevistada y la presión de tiempo que tiene el entrevistador(a).
- *Registro impreciso.* El entrevistador(a) puede “mejorar” el lenguaje de la respuesta u omitir partes de la respuesta que considera inconsistente y de esta manera introduce errores serios.
- *Registro ilegible,* lo cual puede hacer perder el caso.

En el caso de preguntas con alternativas de respuesta precodificadas, el registro puede englobar dos códigos, por lo que se sabe cuál es la respuesta; puede suceder también que el entrevistador/a no encuentre la categoría apropiada que refleje la respuesta de la persona entrevistada y cometa el error de dejarla sin respuesta, o la registre según su conveniencia.

En el caso de respuestas en las que tiene que registrarse un número, el entrevistador(a) puede escribirlo de modo confuso, como por ejemplo confundir un 3 con un 5, o 1 con 4, etc. Lo que ocasionará errores en la digitación; para evitar esto, en la etapa de capacitación se indica claramente cómo deben escribirse los números.

Para finalizar la entrevista

El entrevistador debe revisar rápidamente que el cuestionario esté lleno, luego debe agradecer a la persona entrevistada por el tiempo y cooperación brindada.

2.3.2 Aplicación de instrumentos cualitativos

A continuación se presentan las pautas a considerar en la aplicación de los principales instrumentos cualitativos.

a. Entrevistas a profundidad

El entrevistador debe tener conocimiento del tema de la entrevista a fin de cumplir con los requisitos de profundización de respuestas asociados al instrumento. Idealmente el entrevistador debe ser una persona con experiencia e instrucción en la aplicación de este tipo de instrumento.

Algunas consideraciones generales para la aplicación son:

- ⇒ La entrevista semiestructurada es la más convencional de las alternativas de entrevista y se caracteriza por la preparación anticipada de un *cuestionario guía o pauta* que se sigue, en la mayoría de las ocasiones, de una forma estricta aun en su orden de formulación.
- ⇒ El cuestionario guía cumple varias funciones. Su primer papel es asegurar que el evaluador *cubra todo el tema*, en el mismo orden, para cada entrevistado, preservando de manera consistente el contexto conversacional de cada entrevista. La segunda función es cuidar el itinerario requerido para mantener la distancia con el entrevistado. La tercera función consiste en establecer los canales para la dirección y delimitación del discurso. La cuarta función es permitir al evaluador prestar toda su atención al testimonio de su entrevistado.
- ⇒ El enfoque cualitativo con que se asume este tipo de entrevista abre la oportunidad para que, con cada una de las respuestas a las preguntas del cuestionario, se exploren de manera no estructurada (esto es, no preparada de antemano, pero sí sistemática) aspectos derivados de las respuestas proporcionadas por el entrevistado.
- ⇒ Para efectos del análisis no basta solo con registrar las ideas; también se requiere examinar el contexto en que esas ideas aparecen. Por ello se recomienda grabar las entrevistas, en cuyo caso se debe pedir al informante su consentimiento para efectuar la grabación.
- ⇒ Para efectos de la aplicación *in situ*, es necesario tener en cuenta algunas consideraciones prácticas a la hora de realizar una entrevista:

- Tomar contacto por anticipado con el entrevistado para concertar una cita para la entrevista e informarle acerca del estudio y sus objetivos. Hay que tener en cuenta que la entrevista dura una hora aproximadamente, por lo que se recomienda agendarla con anticipación.
- Buscar un lugar apropiado para la realización de la entrevista. Este lugar debe ser un sitio cómodo tanto para el entrevistador como para el entrevistado, y además debe ser tranquilo, desprovisto de grandes distracciones y sin ruido, de manera tal que la conversación se pueda dar de manera fluida, pueda ser grabada y sea posible escuchar con claridad al entrevistado.
- El entrevistador debe haber leído con anterioridad la pauta de entrevista, de tal forma que se encuentre en conocimiento de las preguntas y le sea fácil darle continuidad a la entrevista y repreguntar al entrevistado. Hay que recordar que la entrevista debe fluir de manera de la manera más parecida posible a una conversación.
- El entrevistador debe procurar llevar a la entrevista los implementos necesarios, vale decir, la pauta de entrevista, lápiz y papel para tomar notas, grabadora (en caso de necesitarla) y pilas para la grabadora (estas deben ser revisadas con anterioridad para que no se corte la grabación).

b. Grupos focales

Las consideraciones generales se muestran a continuación.

⇒ La planeación de los grupos focales implica considerar, a partir del presupuesto y el tiempo disponibles, los aspectos básicos que se enuncian a continuación:

- Número de grupos que se estructurarán teniendo en cuenta que cada uno de ellos constituye una unidad de análisis en sí mismo.
- Tamaño de los grupos dentro del rango ya planteado de seis a ocho personas.
- La selección de los participantes según el diseño planteado en el plan o protocolo de investigación.
- La determinación del nivel de involucramiento del investigador como moderador.

⇒ Existen cuatro criterios para orientar las entrevistas de grupo focal en forma efectiva:

- Cubrir un rango máximo de temas relevantes según el tiempo disponible.
- Proveer datos de la manera más específica posible.
- Promover la interacción que explore los sentimientos de los participantes con una cierta profundidad.
- Tener en cuenta el contexto personal que los participantes usan para generar sus respuestas al tema explorado.
 - Tal como sucede con las entrevistas semiestructuradas, es recomendable grabar la sesión de grupos focales, independientemente de una hipotética transcripción posterior.

⇒ Para efectos de la aplicación, es necesario tener en cuenta algunas consideraciones prácticas a la hora de realizar los grupos focales:

- Tomar contacto por anticipado con los asistentes al grupo focal para concertar la fecha y hora. Hay que tener en cuenta que el grupo

focal dura una hora aproximadamente, por lo que se recomienda planearlo con anticipación.

- Buscar con antelación un lugar apropiado para la realización de la entrevista. Este lugar debe proveer un sitio cómodo tanto para los participantes como para el moderador. Debe ser un sitio tranquilo, desprovisto de grandes distracciones y sin ruido en el ambiente, de manera tal que la conversación se pueda dar de manera fluida, pueda ser grabada y sea posible escuchar con claridad lo que dicen los participantes.
- El moderador debe haber leído con anterioridad la pauta guía, de manera tal que conozca las preguntas y le sea fácil darle continuidad a la entrevista y repreguntar a los entrevistados. Hay que recordar que la sesión debe fluir de manera similar a una conversación.
- El moderador debe crear una atmósfera relajada y de conversación entre los participantes.
- El moderador debe tratar de involucrar en la conversación a todos los participantes, animando a aquellos que se encuentran más callados a contestar las preguntas y participar de la conversación, así como evitar el excesivo protagonismo de alguno de ellos.
- El moderador debe procurar llevar todos los implementos necesarios, vale decir, la pauta, lápiz y papel para tomar notas, grabadora y pilas para la grabadora (estas deben ser revisadas con anterioridad para que no se corte la grabación).

c. Observación

Algunas consideraciones prácticas para la ejecución de la observación son:

- El investigador tiene que convivir el tiempo suficiente en la comunidad que estudia para ver lo que sucede en forma reiterada. Es importante que el observador sea capaz de reproducir las estructuras significativas de la vida de los individuos, el contenido y la forma de interacción verbal, la conducta no verbal (gestos, posturas, mímicas, movimientos del cuerpo, etc.), los patrones de acción y no-acción, los registros de archivos, documentos,

artefactos y todo rastro de huellas que puedan ser observados. Para registrar las informaciones el investigador utiliza las anotaciones de campo.

- El trabajo de campo puede incluir otros instrumentos de producción de datos y pluralidad de registros, aunque destacan las formas cualitativas. Por ejemplo, conversar con los usuarios.
- El observador debe tomar una postura y actitud natural o usual a la hora de observar para afectar lo menos posible la naturalidad de la situación que se encuentra observando.
- Se recomienda al observador llevar consigo aquellos instrumentos que le ayuden a registrar las anotaciones de campo y los sucesos que ocurren (lápiz y cuaderno, grabadora y/o cámara fotográfica).
- El observador debe ser riguroso y acucioso para llevar el registro de los sucesos que ocurren. Debe procurar anotar, por ejemplo, las horas de los acontecimientos de los hechos, así como también del inicio y del fin de la observación.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Hernández Fernández Baptista. Metodología de la investigación. Capítulo 9. 2006.

2.4 Control de calidad en la obtención de datos

El control de calidad de los datos es un proceso indispensable que acompaña las diferentes etapas de la evaluación, especialmente durante el recojo de información, y depende tanto del diseño instrumental como de su adecuada aplicación en el trabajo de campo. De acuerdo a las etapas de evaluación, podemos dividir el control de calidad en:

a. *Control durante el diseño:* comprende las actividades para asegurar la calidad de los datos antes de su recolección como analizar la validez y confiabilidad de los instrumentos a partir de los resultados de la prueba piloto.

b. *Control en el recojo de información:* hace referencia al monitoreo y supervisión de la calidad de datos durante la conducción del estudio. En él se

realizan actividades como la revisión de una muestra de cuestionarios recabada cada día o la verificación de la concurrencia del personal encuestador a los hogares que debió encuestar.

c. *Control durante el procesamiento de información:* manejo y proceso de datos en todo el estudio: emisión de estadísticos descriptivos para evaluar la consistencia de las respuestas y/o evaluación el nivel de pérdida de datos por cada pregunta.

No existe un consenso claro sobre los criterios para definir las características que posee la información estadística de calidad. La amplitud del concepto dificulta dar con una definición precisa y menos aún establecer criterios uniformes para su medición (INEI 2005).

Para efectos de la presente sección nos concentraremos en la supervisión de la calidad de datos que se efectúa durante la conducción del estudio.

2.5 Criterios para la supervisión de la calidad

Es posible que los miembros del equipo de recolección de datos tengan varios niveles de competencia y experiencia y diferentes puntos fuertes y habilidades. Para asegurar una recolección de datos de calidad, es necesario nombrar a una persona o más para dirigir y supervisar al equipo (o a los equipos) de recolección de datos.

Las actividades principales en la supervisión de la calidad son las siguientes:

- a. *Control de desempeño.* Consiste en la revisión de la tarea del encuestador en cuanto a diligenciamiento oportuno de los formularios en cada uno de los encuestados y al logro de las entrevistas.
- b. *Control de calidad.* El objetivo es orientar la verificación de la calidad y consistencia de la información recolectada por los encuestadores.
- c. *Control de visitas.* Tiene como finalidad la verificación de la concurrencia directa y personal del encuestador a los encuestados que debió visitar y encuestar, así como del cumplimiento de las visitas concertadas para ubicar al informante indicado.
- d. *Control de cobertura.* Se realiza conjuntamente con el control de visitas y tiene como propósito comprobar que las unidades encuestadas correspondan

exactamente a las seleccionadas en la muestra y asegurar la efectiva determinación y clasificación de las unidades no logradas según categoría.

e. *Evaluación de la entrevista.* Tiene como objetivo verificar el manejo del cuestionario por parte del encuestador. Específicamente en el conocimiento conceptual de este, observancia de las secuencias, planteamiento de las preguntas a los informantes, determinación e inclusión de todos los miembros de hogar.

f. *Revisión de cuestionarios.* Implica examinar en gabinete todos los formularios diligenciados para detectar y corregir en el mismo terreno, si fuera necesario, las inconsistencias, omisiones, legibilidad y grado de detalle de ciertas preguntas; asimismo, implica comprender la verificación de las correcciones hechas por el encuestador en terreno generadas en la revisión de los formularios diligenciados y comprobar que las viviendas, hogares e informantes encuestados correspondan efectivamente al directorio de la selección muestral.

Para ampliar el tema tratado, revisa la siguiente lectura:

- CERALC. Manual de Aplicación de un Modelo de Evaluación para Identificar el Aporte de las Bibliotecas Públicas en el Desarrollo de sus Comunidades. Manual de Lecturas. Santiago de Chile. Centro Regional para el Fomento del Libro en América Latina y el Caribe (CERLALC). 2012. Capítulo 3.
- Instituto de Nutrición de Centro América y Panamá INCAP. Manual de Operaciones para Recolección de Datos en un Sistema de Vigilancia Alimentario Nutricional. Guatemala, INCAP. 2012. p. 1-10.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 2

Antes de ejecutar las actividades aplicativas, revisa tus respuestas de las preguntas de reflexión, analízalas, interprétalas y sintetízalas con el respaldo teórico que has adquirido al desarrollar las unidades 1 y 2

Ahora procede a desarrollar las actividades aplicativas de la unidad

ACTIVIDAD OPCIONAL DE ACUERDO AL INTERÉS EL PARTICIPANTE:

Diseñando la capacitación para el personal que aplicará los instrumentos

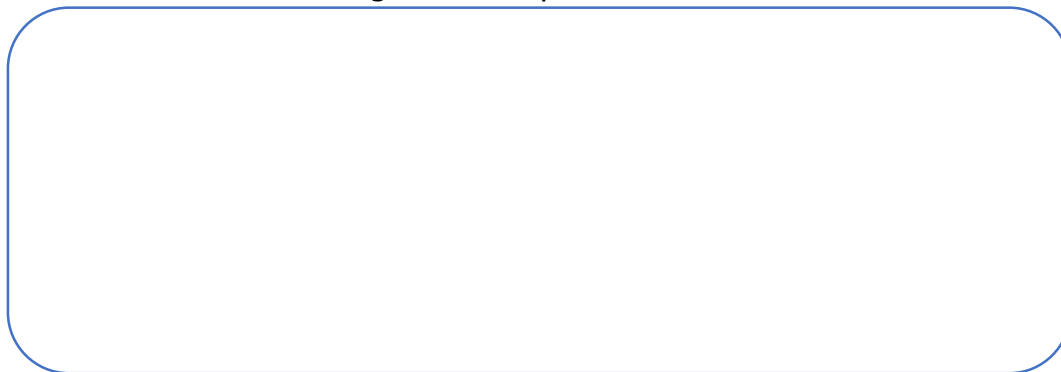
1. Elabora un perfil de equipo que deberá aplicar los instrumentos.
2. Elabora el diseño de la/las sesiones de capacitación para el personal que aplicará los instrumentos considerando los aspectos que se precisan en los ítems 2.1 y 2.2 de la unidad. Toma en cuenta las especificaciones de acuerdo al tipo de técnica que aplicaste en la prueba piloto. Considera la utilización de metodologías, técnicas participativas, materiales de apoyo.
3. De ser necesario, solicita tutoría virtual.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 2

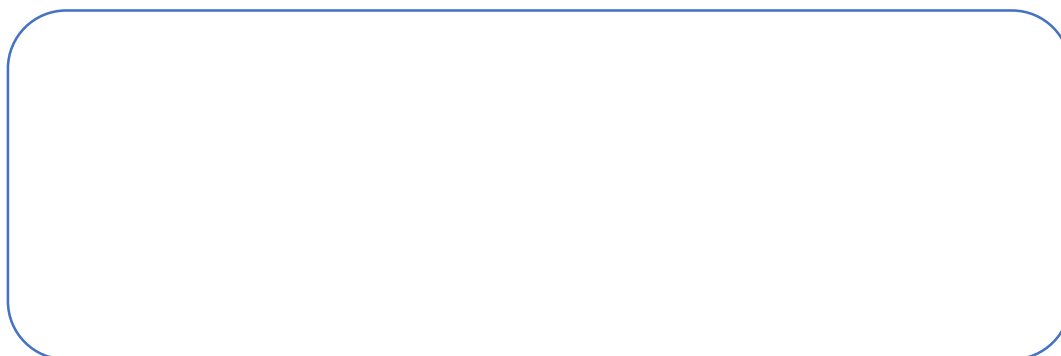
HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

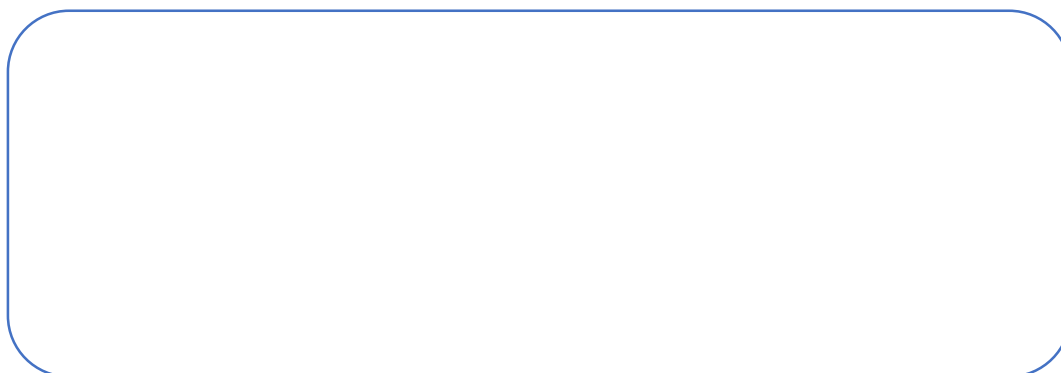
1. Suponiendo que tienes a cargo de la capacitación a encuestadores para la aplicación de una encuesta a hogares en el marco de una evaluación de impacto, anota las indicaciones generales que debes dar a los encuestadores.



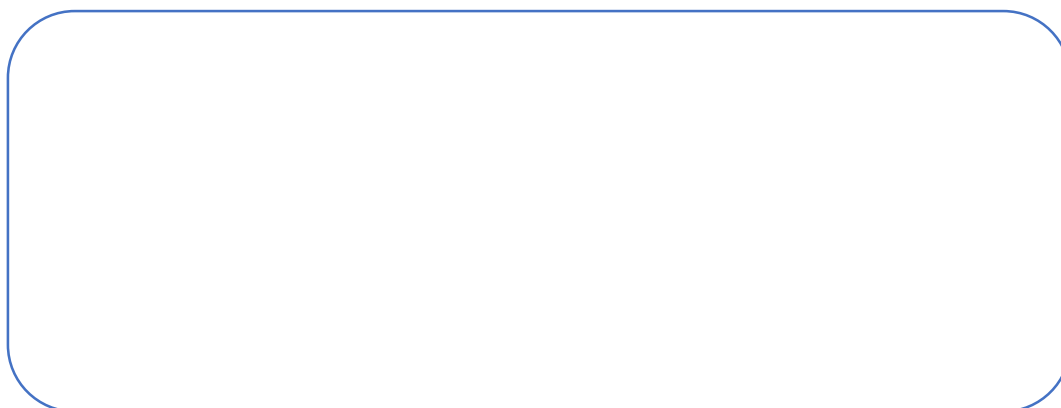
2. Explica las principales etapas para la aplicación de instrumentos cuantitativos.



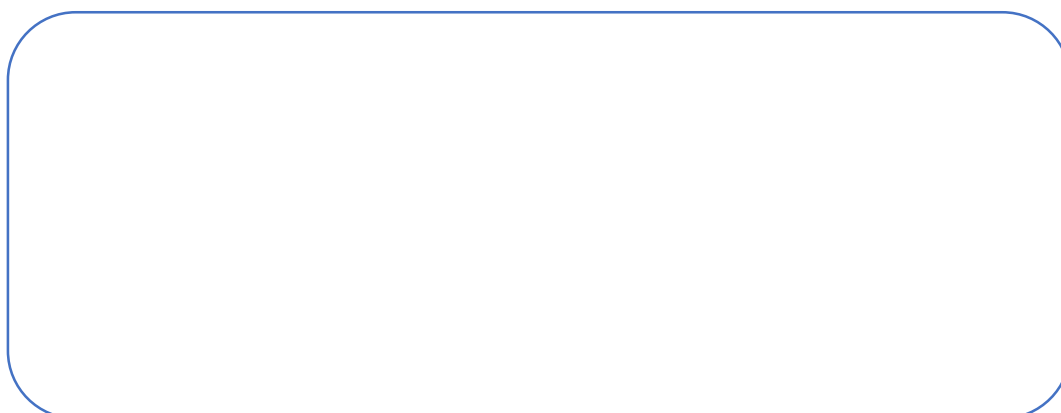
3. ¿Cuáles son las principales normas que se debe considerar en la aplicación de cuestionarios mediante la entrevista?



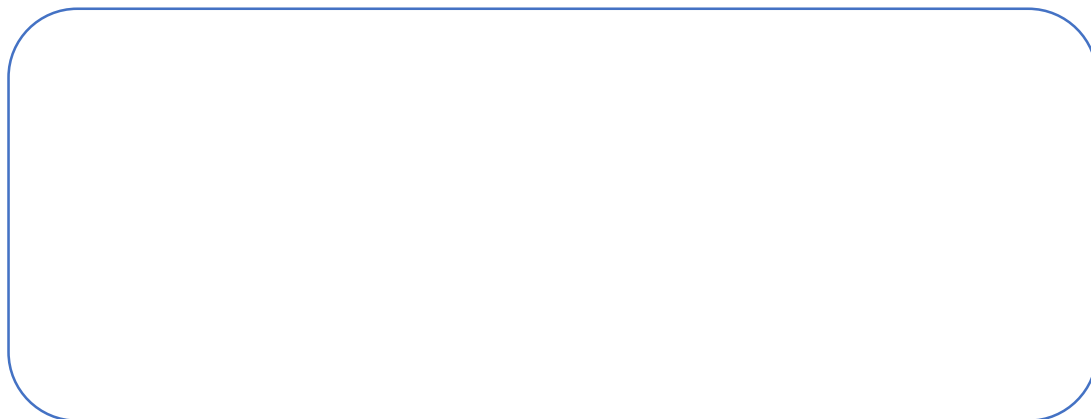
4. ¿Cuáles son los principales aspectos a considerar durante la aplicación de un grupo focal?



5. Explica el objetivo de la actividad de control antes, durante y posterior al recojo de información.



6. Describe las principales actividades del proceso de control de calidad en la recolección de datos.





UNIDAD DE APRENDIZAJE 3: PROCESAMIENTO DE LA INFORMACIÓN Y CONSISTENCIA DE DATOS

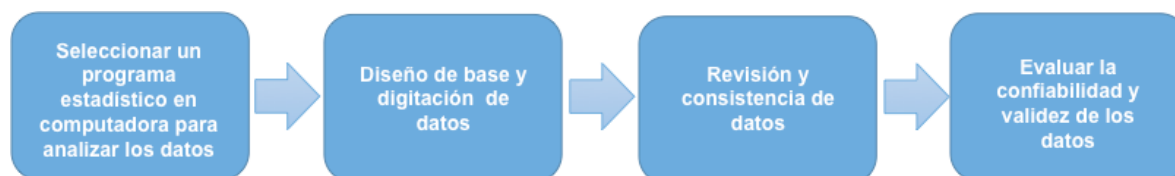
El procesamiento de información es una gran etapa dentro del proceso de evaluación. Consiste en sintetizar todos los datos cuantitativos y cualitativos recogidos mediante diferentes técnicas e instrumentos en información confiable que se constituirá en nuestro principal insumo para el análisis. Todo ello permitirá responder las preguntas de evaluación y comprobar las hipótesis.

En la actualidad, el procesamiento de datos se realiza a través de programas computarizados, los cuales se pueden clasificar de acuerdo a los tipos de datos que permiten procesar: cuantitativos y cualitativos. Por ejemplo, programas estadísticos como el SPSS y Stata nos permiten procesar datos cuantitativos; por su parte, el Atlas.ti y Etnograph son programas que nos permiten procesar datos cualitativos. La presente sección explica los procedimientos que se realizan durante el procesamiento de datos cuantitativos y cualitativos, así como los procedimientos de control de calidad y consistencia de datos que nos permitirán detectar errores y subsanarlos para garantizar la validez y confiabilidad de la información recogida. La ejecución de estas etapas valida la base de datos final que nos permitirá realizar el análisis de los datos para el informe de evaluación.

3.1. Procesamiento de datos cuantitativos

El procesamiento de los datos cuantitativos involucra una serie de pasos que van desde seleccionar el programa estadístico idóneo al tipo y cantidad de datos que requerimos procesar, el diseño de la base de datos, la digitación de información en la matriz computacional, hasta la exploración de datos a través de estadísticos descriptivos y algunos test estadísticos a fin de evaluar la confiabilidad de los datos recabados que nos permita realizar las correcciones pertinentes y contar con una base de datos final de donde se obtenga información confiable sobre los indicadores bajo medición. Ver la figura 6.

Figura 6. Proceso para efectuar el análisis estadístico



Fuente: elaboración Alejandro Bardales, 2014.

3.1.1. Selección del programa

En primer lugar se debe seleccionar el programa o paquete estadístico en el cual se van a digitar los datos contenidos en los cuestionarios. Dependiendo de la extensión y complejidad de los cuestionarios se pueden emplear distintos programas de entrada de datos; el más común es el paquete ACCESS de la familia de Microsoft; otro es el CSPRO (<http://www.census.gov/ipc/www/cspro/>), para bases de datos más complejas, y que es de uso libre. Con cuestionarios sencillos se puede emplear el EXCEL de la familia Microsoft.

Para el procesamiento existen diversos programas que en esencia tienen un funcionamiento similar dado que incluyen dos partes o segmentos: (i) definición de variables, que a su vez explican los datos (los elementos de la codificación ítem por ítem) y (ii) la otra parte, la matriz de datos. Sin embargo, es importante reconocer las potencialidades y deficiencias de los distintos tipos de programas disponibles en el mercado a fin de determinar cuál se ajusta de manera directa a los requerimientos específicos de la encuesta a procesar. Por ejemplo: el SPSS es muy amigable y cuenta con muchos comandos predeterminados que con unos conocimientos básicos del programa le permitirán obtener de manera rápida las estadísticas; en tanto, el STATA incluye una serie de comandos que, si bien no están predeterminados y deben ser programados, nos dan mayor flexibilidad para transformar variables. Otro paquete de amplio uso, sobre todo en temas de salud, es el Epi info (<http://wwwn.cdc.gov/epiinfo/7/>), que es de uso libre.

3.1.2. Diseño de base de datos

El siguiente paso, luego de haber definido el programa que se va a utilizar para la entrada de datos, consiste en realizar el diseño de esta, para lo cual se debe declarar todas las variables y categorías de cada una de las variables de la encuesta que vamos a procesar, así como la cantidad de observaciones (filas), a fin remitir la base a los digitadores para que la información pueda ser ingresada. Lo que el evaluador hace es definir los parámetros de la matriz de datos en el programa (nombre de cada variable en la matriz -que equivale a un ítem, reactivo, categoría o subcategoría de contenido u observación-, tipo de variable o ítem, ancho en dígitos, etc.) e introducir los datos en la matriz como con cualquier hoja de cálculo.

Se debe considerar que la matriz tiene columnas (variables o ítems), filas o renglones (casos) y celdas (intersección entre una columna y un renglón). Cada celda contiene un dato que significa un valor de un caso en una variable. Por ejemplo, supongamos que tenemos cuatro casos y tres variables (género, color de cabello y edad). La matriz se vería como en la tabla 1. Recordemos que en la matriz se deben ingresar los códigos y no la redacción textual de cada categoría de una pregunta.

Tabla 1. Matriz de datos con tres variables y cuatro casos

Caso	Columna 1 (género)	Columna 2 (color de pelo)	Columna 3 (edad)
1	1	1	35
2	1	1	29
3	2	1	28
4	2	4	33

Fuente: Hernández et al (2006).

La codificación (especificada en la parte de las definiciones de las variables o columnas que corresponden a los ítems) sería:

- Género (1 = masculino y 2 = femenino).
- Color de cabello (1 = negro, 2 = castaño, 3 = pelirrojo, 4 = rubio).
- Edad (dato "bruto o crudo" en años).

De esta forma, si se lee por renglón o fila (caso), de izquierda a derecha, la primera celda indica un hombre (1); la segunda, de cabello negro (1), y la tercera, de 35 años (35). En el segundo caso, un hombre de cabello negro y 29 años. La tercera, una mujer de cabello negro, con 28 años. La cuarta fila (caso número cuatro) nos señala una mujer (2), rubia (4) y de 33 años (33). No obstante, si leemos por columna o variable de arriba hacia abajo, tendríamos en la primera (género) dos hombres y dos mujeres (1,1, 2,2).

Por lo general, en la parte superior de la matriz de datos aparecen las opciones de los comandos para operar el programa de análisis estadístico como cualquier otro programa. Una vez estemos seguros de que no hay errores en la matriz, procedemos a realizar el análisis de la matriz, el análisis estadístico. En cada

programa tales opciones varían, pero en cuestiones mínimas (Hernández et. al; 2008).

3.1.3. Digitación/captura de datos

La digitación de los datos representa el primer paso para la sistematización de la información. Actualmente existen modernos sistemas de digitación que permiten el diseño de pantallas a imagen del cuestionario, la definición de criterios de validación, la verificación de información durante la digitación e incluso poder ingresar la información en línea aún durante la etapa de recojo de información. Así, los tiempos de respuesta se han acortado y la mayor velocidad y capacidad de memoria permiten reducir los plazos para la depuración y procesamiento de la información (Feres J. & F. Medina; 2001).

Para garantizar la adecuada digitación de la información es importante capacitar a los digitadores tanto en la estructura, preguntas, categorías y códigos de la propia encuesta como en el entorno de la base de datos, en tanto es importante evaluar la carga de trabajo asignada al personal en este periodo. Los controles que deben ser aplicados para evaluar la correcta digitación de la información se detallan en la sección 3.3 como parte de la revisión y consistencia de datos.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Hernández R. et al. Metodología de la investigación. Capítulo 10. 2006.

3.2. Procesamiento de datos cualitativos

El procesamiento de información cualitativa requiere que el evaluador/investigador tenga las capacidades analíticas y de síntesis, así como el conocimiento de técnicas que le permitan identificar dentro de la información las categorías que se revelan a partir de las respuestas brindadas por los participantes del estudio.

Por otro lado, el procesamiento de información cualitativa implica ciertas etapas diferenciadas. La primera es una fase de descubrimiento en progreso: identificar temas y desarrollar conceptos y proposiciones. La segunda fase, que típicamente se produce cuando los datos ya han sido recogidos, incluye la codificación de los datos y el refinamiento de la comprensión del tema de estudio. En la fase final,

el evaluador trata de relativizar sus descubrimientos, es decir, de comprender los datos en el contexto en que fueron recogidos.

3.2.1. Descubrimiento

En las evaluaciones cualitativas, los investigadores le van dando gradualmente sentido a lo que estudian combinando perspicacia e intuición y una familiaridad íntima con los datos. Con frecuencia, ese es un proceso difícil. La mayor parte de las personas sin experiencia en evaluación cualitativa tienen dificultades en reconocer las pautas que emergen de sus datos. No hay ninguna fórmula simple para identificar temas y desarrollar conceptos. Algunas recomendaciones para esta etapa son:

- ⇒ *Leer repetidamente los datos.* Se deben reunir todas las notas de campo, las transcripciones, documentos y otros materiales y leerlos cuidadosamente.
- ⇒ *Seguir la pista de temas, intuiciones, interpretaciones e ideas.* Se debe registrar toda idea importante que se tenga durante la lectura y reflexión sobre los datos. En la observación participante, los evaluadores a veces emplean los "comentarios de observador" para anotar temas y registrar interpretaciones, mientras que en las entrevistas en profundidad se puede usar con el mismo propósito el diario de apuntes. A medida que se leen los datos, también se pueden efectuar anotaciones en los márgenes.
- ⇒ *Buscar los temas emergentes.* Es preciso buscar en los datos los temas o pautas emergentes: temas de conversación, vocabulario, actividades recurrentes, significados, sentimientos, dichos y proverbios populares. Se sugiere confeccionar listas tentativas de temas en esta etapa del proceso, pero no se debe apostar a ninguna idea en particular hasta haber tenido la oportunidad de experimentarla y controlarla.
- ⇒ *Elaborar tipologías.* Las tipologías o esquemas de clasificación pueden ser ayudas útiles para identificar temas y desarrollar conceptos y teorías. Una clase de tipologías se relaciona con el modo en que las personas clasifican a los otros y con los objetos de sus vidas.

- ⇒ *Desarrollar conceptos y proposiciones teóricas.* El evaluador pasa de la descripción a la interpretación y la teoría a través de conceptos y proposiciones. Los conceptos son ideas abstractas generalizadas a partir de hechos empíricos. En el análisis cualitativo los conceptos son instrumentos que sensibilizan a los participantes.
- ⇒ *Leer el material bibliográfico.* Los evaluadores que realizan análisis cualitativo empiezan sus estudios comprometiéndose mínimamente con teorías y supuestos *a priori*. Cuando se encare el análisis intensivo, sin embargo, ya deberá haberse familiarizado con la literatura especializada y con los marcos teóricos pertinentes para la evaluación. La lectura de otros estudios con frecuencia proporciona proposiciones y conceptos fructíferos que ayudan a interpretar los datos.
- ⇒ *Desarrollar una guía de la historia.* A veces es útil desarrollar una línea guía de la historia para orientar el análisis. La guía de la historia es la hebra analítica que une e integra a los principales temas de los datos. Es la respuesta a la pregunta "¿sobre qué trata este estudio?"

3.2.2. Codificación

En la evaluación cualitativa, la codificación es un modo sistemático de desarrollar y refinar las interpretaciones de los datos. El proceso de codificación incluye la reunión y análisis de todos los datos que se refieren a temas, ideas, conceptos, interpretaciones y proposiciones. Durante esta etapa del análisis, lo que inicialmente fueron ideas e intuiciones vagas se refinan, expanden, descartan o desarrollan por completo. El siguiente es un modo de codificar los datos cualitativos.

- ⇒ *Desarrollar categorías de codificación.* Se debe empezar redactando una lista de todos los temas, conceptos, interpretaciones, tipologías y proposiciones identificados o producidos durante el análisis inicial. Al poner por escrito las ideas, se debe procurar ser lo más específico posible. Es importante tener algunas perspectivas del tipo de datos que se ajustan a cada categoría. No obstante, algunas de las ideas serán tentativas y estarán vagamente formuladas. Por ejemplo, una categoría de la codificación podría relacionarse con un tema de conversación recurrente. Tales temas también deben incluirse en la lista.

- ⇒ *Codificar todos los datos.* Es importante codificar todas las notas de campo, las transcripciones, los documentos y otros materiales, escribiendo en el margen el número asignado o la letra correspondiente a cada categoría. Hay que codificar tanto los incidentes negativos como los positivos relacionados con la categoría de que se trate.
- ⇒ *Separar los datos pertenecientes a las diversas categorías de codificación.* La separación de los datos es una operación mecánica. El evaluador reúne los datos codificados pertenecientes a cada categoría. Este procedimiento se realiza manualmente: se recortan las notas de campo, las transcripciones y otros materiales y se colocan los datos de cada categoría en carpetas de archivo.
- ⇒ *Evaluar los datos que han sobrado.* Después de haber codificado y separado todos los datos, se debe repasar el remanente de datos que no han ingresado en el análisis. Algunos de ellos probablemente se ajusten a las categorías de codificación existentes. También se pueden plantear nuevas categorías que se relacionen con las desarrolladas previamente y con la guía de la historia subyacente. Pero debe observarse que ningún estudio utiliza todos los datos recogidos. Si no se ajustan, no se debe tratar de forzar el ingreso de todos los datos en el esquema analítico.
- ⇒ *Refinar el análisis.* La codificación y separación de los datos permite comparar diferentes fragmentos relacionados con cada tema, concepto, proposición, etcétera, y en consecuencia refinar y ajustar las ideas. Se encontrará que algunos temas que parecían vagos y oscuros aparecen claramente iluminados. También es probable que algunos conceptos no se ajusten a los datos y que algunas proposiciones pierdan validez.

Sin embargo, se debe indicar que en la actualidad se han desarrollado diferentes programas -además de los procesadores de textos- que sirven para realizar análisis cualitativo pero que de ninguna manera sustituyen el análisis creativo y profundo del evaluador. A continuación se explican los nombres y principales características de los programas que más se utilizan en el análisis cualitativo.

3.2.3. Programas más utilizados en el análisis cualitativo

a. Atlas ti

Este programa permite agregar los datos o documentos primarios (que pueden ser textos, fotografías, segmentos de audio o video, diagramas, mapas y matrices) que luego son codificados de acuerdo al esquema que se haya diseñado. Las reglas de codificación son establecidas por el evaluador y el programa las aplica. En la pantalla se puede visualizar un conjunto de datos o un documento (por ejemplo, una transcripción de entrevista o las entrevistas completas si se integraron en un solo documento) y la codificación que va emergiendo en el análisis. Realiza conteos y visualiza la relación que el evaluador establezca entre las unidades, categorías, temas, memos y documentos primarios. Asimismo, el evaluador puede introducir memos y agregarlos al análisis. Ofrece diversas perspectivas o vistas de los análisis (diagramas, datos por separado, etc.).

Se puede encontrar el manual en español en el siguiente enlace:

http://atlasti.com/wp-content/uploads/2014/05/QuickTour_a7_es_05-1.pdf

b. Ethnograph

Este programa permite identificar y recuperar textos de documentos. La unidad básica es el segmento. Asimismo, codifica las unidades partiendo del esquema de categorización que haya establecido el evaluador. Los segmentos pueden ser anidados, entrelazados y yuxtapuestos en varios niveles de profundidad. Las búsquedas llegan a efectuarse sobre la base de códigos expresados en un carácter, una palabra o en palabras múltiples. Los esquemas de codificación suelen modificarse. Guarda memos, notas y comentarios, y también los incorpora al análisis.

c. QSR, QSR Nvivo

Este programa es útil para construir grandes bases de datos estructuradas jerárquicamente y puede agregar documentos para ser analizados. También, al igual que los dos anteriores, codifica unidades de contenido (texto y otros

materiales) basándose en el esquema diseñado por el evaluador. Localiza los textos por carácter, palabra, frase, tema o patrón de palabras, incluso por hojas de cálculo de variables.

d. Decisión Explorer

Es un programa que nos permite mapear categorías. El evaluador puede visualizar en diagramas relaciones entre conceptos o categorías, introduciendo las categorías y estableciendo, previamente las relaciones. Asimismo, realiza un conteo de la categoría con mayor número de relaciones con sus otros pares. Cualquier idea puede ser convertida en concepto y ser analizada. Resulta útil, también, para visualizar hipótesis y asociar los componentes más importantes de una teoría.

Es importante considerar que durante todo el procesamiento de información se debe tener en cuenta el planteamiento original del problema de investigación con la finalidad de que se no se disperse el análisis y se pueda responder con claridad las preguntas de la investigación (Hernández et. al, 2008).

e. Software libre

En internet se puede conseguir software libre como los siguientes:

- Libre QDA disponible en: <http://www.libreqda.edu.uy/>
- CDC EZ-Text disponible en:
<http://www.cdc.gov/hiv/library/software/eztext/>
- Weft QDA disponible en: <http://weft-qda.uptodown.com/>

3.2.4. Relativización de los datos

La fase final del análisis cualitativo consiste en lo que se denomina relativización de los datos: se trata de interpretarlos en el contexto en que fueron recogidos. En este sentido, todos los datos son potencialmente valiosos si se sabe evaluar su credibilidad. Todos los datos deben relativizarse. Para entenderlos, hay que detenerse en el modo en que fueron recogidos. No se descarta nada. Solo varía la interpretación de acuerdo con el contexto. Hay al respecto un cierto número de consideraciones.

- ⇒ *Datos solicitados o no solicitados.* Aunque los evaluadores que realizan análisis cualitativo por lo general tratan de permitir que las personas hablen sobre lo que tienen en mente, nunca son totalmente pasivos. Formulan ciertos tipos de preguntas y persiguen ciertos temas. Al hacerlo, solicitan datos que podría no haber.
- ⇒ *La influencia del observador sobre el escenario.* La mayor parte de los observadores participantes trata de reducir al mínimo los efectos de su presencia sobre las personas que está estudiando, hasta que ha logrado una comprensión básica del escenario.
- ⇒ *¿Quién estaba allí?* Si el observador puede influir sobre lo que un informante diga o haga, lo mismo puede suceder con muchas otras personas del escenario.
- ⇒ *Datos directos e indirectos.* Cuando analizamos nuestros datos, codificamos tanto los enunciados directos como los datos indirectos referentes a un tema, interpretación o proposición.
- ⇒ *Fuentes.* Existe el peligro de generalizar acerca de un grupo de personas sobre la base de lo que una sola o unas pocas han dicho y hecho. Algunos observadores participantes han sido tan absorbidos por “informantes clave”, han dependido tanto de ellos para recoger información, que terminaron por recoger una visión selectiva del escenario. Una persona habladora puede producir grandes cantidades de datos que aparecen a lo largo de las notas de campo o de las transcripciones.
- ⇒ *Nuestros propios Supuestos.* En el análisis cualitativo, tal como lo hemos descrito, el evaluador empieza el estudio con un mínimo de supuestos. Pese a nuestros compromisos y preconcepciones son imposibles de evitar. Los datos nunca se explican a sí mismos. Todos los evaluadores se nutren de sus propios supuestos teóricos y de sus conocimientos culturales para extraer el sentido de los datos.

3.3. Revisión y consistencia de datos de datos cualitativos y cuantitativos

Una de las etapas más importantes dentro del procesamiento de la información es la revisión y consistencia de la información obtenida¹¹, con lo cual se busca asegurar la validez, calidad y coherencia de los datos. En principio, esta etapa consiste en realizar una serie de procedimientos que permitan identificar posibles errores y su origen, para luego proceder a corregirlos. Los primeros procedimientos detallados en esta sección (críticas y codificación, así como, falta de respuesta) usualmente son evaluados antes de la captura/digitación de datos; luego de ello se realizan verificaciones para detectar errores de digitación y la consistencia de los datos en su conjunto.

3.3.1. No respuesta o rechazo

Durante el diseño de los instrumentos de recolección de información tales como los cuestionarios de encuestas, existe una serie de preguntas cerradas cuyas categorías de respuestas se encuentran claramente codificadas. Se debe contar con un libro de códigos, por ejemplo: sexo: (1) hombre, (2) mujer. Sin embargo, existe otro tipo de preguntas que se recoge de manera abierta, dado que no se conocen *ex ante* los tipos de respuestas que serán proporcionados por los entrevistados o porque existen ciertos temas en los cuales no es conveniente que se efectúe una asignación automática ya que habitualmente se cometen errores de interpretación y clasificación. Por ejemplo: cuando se investiga el tema del empleo y se pretende codificar el tipo de ocupación (principal y secundaria) que realizan los miembros del hogar que manifiestan estar ocupados durante el periodo de referencia de la encuesta, es habitual que se presenten errores en la codificación efectuada en campo. De allí que una alternativa de solución sea obtener una amplia descripción de las actividades de los entrevistados, las cuales se codificarán a partir de las alternativas consignadas en la clasificación de ocupaciones elaborada por los organismos internacionales especializados en el tema del empleo.

En dicho escenario, es indispensable que los encargados de administrar las encuestas dispongan de un grupo de técnicos con gran experiencia y profundo conocimiento de los temas investigados y de los alcances del cuestionario,

¹¹Los procedimientos de revisión y consistencia presentados a continuación se aplican usualmente a la información cuantitativa.

encargados exclusivamente de evaluar la calidad de la información recolectada y la asignación de códigos a las preguntas abiertas que requieren clasificarse durante la etapa de revisión. Así, en esta etapa se definen los códigos válidos para cada categoría de respuesta y sus correspondientes rangos, las subpoblaciones de interés para cada tema y las relaciones permitidas entre las variables.

Se debe estar consciente de que es normal que se presenten problemas de campo no previstos en los manuales y formatos de llenado. De hecho, es prácticamente imposible prever todas las situaciones y reflejarlas en la modalidad de captura, crítica y validación de los datos en campo sin quedar expuesto a que se genere una cantidad considerable de errores de clasificación sin que los programas de verificación sean capaces de detectar tales inconsistencias. Del mismo modo, es preciso reiterar una vez más que los sesgos en los datos son difíciles de evaluar y en la mayor parte de los casos son más dañinos que los errores de muestreo.

3.3.2. No respuesta o rechazo

Se debe efectuar el análisis de la no respuesta total, por causas, por entrevistador, por grupos de hogares y por zonas geográficas, a fin de evaluar posibles errores de omisión o subespecificación. Mediante este procedimiento se pueden identificar posibles sesgos u omisiones, determinar la calidad del marco y evaluar la eficiencia de la labor realizada por los enumeradores. La concentración de la no respuesta en alguna zona o grupo social específico puede distorsionar los resultados de la investigación. Un buen registro y control de la falta de respuesta es un requisito para garantizar la confiabilidad de los datos reportados.

Cuando el número de entrevistas no logradas es poco significativo o se encuentra dentro del rango estimado, y tiene visos de haberse dado de manera aleatoria en la distribución de la muestra, la eliminación de estas observaciones no afecta la precisión de los datos. Sin embargo, una concentración de la no respuesta en algunos segmentos poblacionales, sí puede tener efectos en los resultados, por lo que es necesario evaluar los métodos más adecuados para resolver este problema, como por ejemplo revisar las ponderaciones (inversos de las probabilidades de selección). Por otra parte, es también necesario prestar atención a la no respuesta parcial, la cual ocurre cuando algunas preguntas o secciones del cuestionario no son respondidas por el hogar o alguno de sus miembros. Cuando la frecuencia de ocurrencia es baja, las observaciones que

presenta esta condición pueden ser eliminadas sin que ello afecte la precisión de los resultados. Sin embargo, cuando muchas variables tienen datos faltantes, el análisis de las relaciones entre ellas puede basarse en pocos datos, por lo cual se recurre a procedimientos estadísticos para imputar valores faltantes, como por ejemplo la media entre los datos válidos o, mejor, métodos de regresión. En efecto, esta práctica permite reducir el número de preguntas sin respuesta; sin embargo, la imputación no siempre reduce los errores de estimación.

3.3.3. Evaluar errores de digitación

El proceso de digitación es sensible a errores. Por eso se registran, en muchos casos, datos faltantes o valores o códigos distintos a los que estaban registrados en el cuestionario. Para estos casos es recomendable aplicar métodos de muestreo por lotes para verificar la calidad de la información capturada, así como para evaluar el desempeño del personal encargado de la digitación. En términos prácticos, un verificador revisa para una muestra de encuestas si los códigos ingresados a la base son los mismos que se encuentran registrados en papel. La mejor práctica para controlar este error es la doble digitación, por personas distintas, y después compararlas individualmente. Luego de identificadas las diferencias se debe contrastar con los cuestionarios para determinar cuál es el dato válido.

3.3.4. Consistencia de la información

El análisis de la consistencia de los datos tiene como propósito garantizar la calidad de la información mediante la revisión sistemática de cada una de las respuestas tanto de manera individual como en su relación con otras. Esta etapa comprende al menos las dos etapas siguientes: i) análisis de consistencia interna y ii) validación de los resultados (consistencia externa) con encuestas sobre el mismo tema realizadas en el pasado, con otras encuestas que generen datos sobre temas comunes, datos macroeconómicos, estudios internacionales y registros administrativos.

Este proceso se realiza sobre bases de datos ya sin errores de digitación.

El análisis de consistencia interna comprende los siguientes pasos en el caso de encuestas.

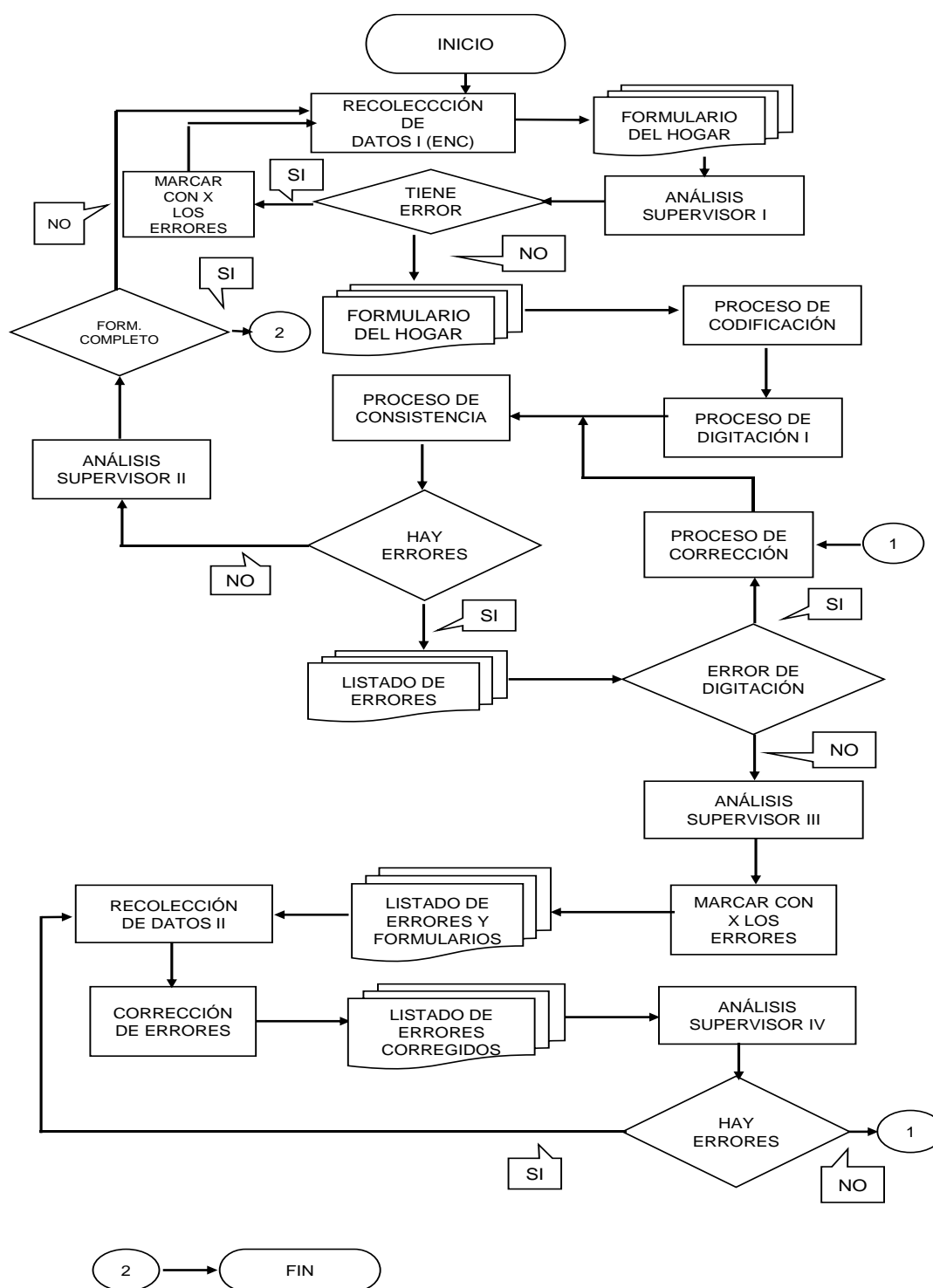
- i. Consistencia de rangos, es decir, que no aparezcan códigos que no correspondan a los establecidos en el cuestionario o valores que estén fuera de los esperados, por ejemplo que no aparezcan mujeres de nueve años con hijos.
- ii. Análisis de valores extremos en datos cuantitativos, es decir, que no aparezcan datos muy lejanos de la distribución general, por ejemplo que aparezcan ingresos de un nuevo sol por día, o de 100 000 nuevos soles en el último mes en el área rural.
- iii. Análisis cruzado de variables, que permiten ver las fallas en los flujos en las entrevistas en que se puede haber consignado datos a quienes no debían responder, por ejemplo en un estudio de fecundidad en mujeres en edad fértil que aparezcan varones con datos sobre nacimientos, o niños con estudios completos de postgrado.

Las reglas de consistencia deben estar explicitadas en un documento especial y se deben generar listados de errores, los cuales sirven para que el personal de campo pueda efectuar las revisiones y análisis que correspondan, hasta determinar el origen del error, establecer la corrección a efectuar y realizar la actualización de los archivos de bases de datos; también sirven como soporte documental de los procesos de consistencia efectuados, de manera que los monitores y directores del proyecto puedan evaluar en cualquier momento, durante el desarrollo del trabajo, los problemas de calidad que se están presentando, las soluciones que se dieron, y tomar con base en su análisis las acciones correctivas y de estandarización de procesos que correspondan.

Por otra parte, estos listados son un valioso material de consulta para los analistas de resultados y especialistas temáticos interesados en conocer al detalle el grado de cambio que el proceso consistencia de datos introdujo en los resultados.

Al emitir un listado de errores se debe verificar que los errores señalados no son producto de una mala digitación; se deben señalar las correcciones que procedan de esta fuente y entregar el listado al supervisor para su revisión. A continuación, en la figura 7, se detallan a manera de ejemplo los pasos del proceso de consistencia de datos para una encuesta de hogares sobre empleo e ingresos (Instituto Nacional de Estadística - INEI; 2003).

Figura 7. Diagrama del procedimiento de consistencia de datos



Fuente: INEI (2003).

Descripción del procedimiento de consistencia de datos de una encuesta de hogares de empleo e ingresos

1. El encuestador recoge la información del hogar seleccionado.
2. El encuestador entrega al supervisor de grupo los formularios que tienen información pendiente de procesar.
3. El supervisor revisa los formularios entregados por el encuestador, analiza si hay errores por parte de este y los marca con una "X" de color verde. Si hay errores por parte del encuestador, el supervisor regresa los formularios con error al encuestador para que los corrija directamente en los hogares.
4. El supervisor de grupo entrega los formularios al digitador para que este los capture.
5. El digitador ejecuta el procedimiento de consistencia de datos en la opción de menú que corresponda en el sistema de procesamiento de datos.
6. Si el sistema emite el listado de errores correspondientes, el digitador verifica si hay errores de "digitación". Si hay errores de digitación, el digitador graba los errores detectados en la opción de menú correspondiente.
7. El supervisor recibe del digitador el listado de errores de inconsistencia y los respectivos formularios; analiza el listado y si el formulario fue procesado sin error va a paso No. 12, si no, marca los errores a corregir con una "X" de color verde.
8. El encuestador recibe los formularios y el listado de errores para visitar de nuevo a los hogares y corregirlos. Si la corrección es un cambio en un valor, el mismo se hace a la par (en el formulario y el listado) con color azul; si no hay ningún cambio, es decir permanece el mismo valor, el encuestador hace un chequeo a la par del dato (en el formulario y el listado) correspondiente con color azul.
9. El supervisor recibe de nuevo el listado de error y formulario. Revisa que todos los errores hayan sido corregidos y que la corrección se haya anotado tanto en el formulario como en el listado.
10. El supervisor entrega al digitador el listado de errores corregido y los formularios que correspondan.
11. El digitador captura las correcciones señaladas, las aplica a los archivos correspondientes y realiza de nuevo el procedimiento a partir del paso número 5.
12. Si el formulario fue procesado completamente, el supervisor notifica verbalmente esta circunstancia al digitador y al encuestador correspondiente y almacena el formulario en el lugar que corresponda; si el formulario no está procesado completamente, debe regresarlo al encuestador para que se repita este procedimiento a partir del paso 1 y completar la información pendiente.

Fuente: INEI (2003).

En el análisis de la consistencia externa se generan corridas de las variables y se determina su coherencia con el sentido común y con otras fuentes válidas; por ejemplo, en una población general no puede ser posible que la edad promedio sea 45 años, o que el 60% de los adultos tenga estudios superiores.

Finalmente, cabe indicar que los datos generados se considerarán de buena calidad y útiles para la toma de decisiones en la medida que sus órdenes de magnitud guarden relación con el resto de cifras micro y macro del sistema nacional de información. Asimismo, se podrán considerar adecuados desde el punto de vista estadístico en la medida que el orden de su error absoluto y relativo esté en un rango aceptable, el cual fue definido al momento de establecer los objetivos de la investigación.

3.3.5. Precisión estadística de los estimadores

Un paso final para evaluar la confiabilidad de los datos es la evaluación de la calidad de la información desde la perspectiva estadística a partir del cálculo de los errores de muestreo y del efecto de diseño. Los errores de muestreo se calculan sobre los promedios, porcentajes y razones.

El cuadro 18 presenta la información de los errores de muestreo para un ejemplo donde se estima el porcentaje de mujeres en edad fértil de Ancash que en el periodo de la encuesta (2004-2006) estaba usando anticonceptivos.

Cuadro 18. Errores de muestreo

Valor estimado	0,679 (67,9%)
Error estándar:	0,025
Número de casos sin ponderar:	352
Efecto del diseño:	0,988
Efecto del diseño:	0,036 (es el coeficiente de variación)
Intervalo de confianza al 95%:	Mínimo 0,630, máximo 0,729

La información obtenida a partir del cuadro de errores muestrales nos dice lo siguiente: el 67,9% de las mujeres en edad fértil estaba usando métodos anticonceptivos, con un margen de error de +/- 4,9% (es el error muestral 0,025 multiplicado por 1,96, que es el valor correspondiente al 95% de confianza), con lo que el verdadero porcentaje se encontraba entre 63,0% y 72,9% con un 95% de confianza, que el coeficiente de variación de esta estimación era del 3,6% lo que indica que es de alta precisión, y el efecto de diseño de 0,988 es muy cercano a 1,00 lo que indica que el diseño muestral empleado ha dado la misma varianza que un muestreo aleatorio simple.

En caso que el error relativo estimado se encuentre por debajo de 0,15 (5%) se dice que la precisión es muy alta, si está entre 5 y 10% es buena, si está entre 10 y 15% es aceptable, en estos casos será posible generalizar el resultado de la variable al universo de estudio con la confianza asumida en los objetivos de la evaluación (generalmente 95%); si el error relativo o coeficiente de variación es superior a 15% no es recomendable generalizar el valor obtenido por su baja precisión.

El error estándar y el efecto de diseño son diferentes para cada variable en una misma encuesta y sus valores cambian debido a la magnitud de los valores estimados y a la cantidad de casos (muestra real) sobre los que se basa el cálculo de cada uno de ellos.

La precisión de la estimación es solo un componente de la calidad de las estimaciones que puede ser mejorado solo aumentando el tamaño de la muestra; el segundo componente de la calidad es la validez y confiabilidad de los instrumentos utilizados, y el tercero, el más importante, tiene que ver con la pulcritud y prolijidad en los procesos de la encuesta (fiel cumplimiento del diseño muestral, capacitación, entrevista, registro del datos, digitación, depuración de datos, y cálculos, etc.), lo cual es conocido como errores ajenos al muestreo.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Metodología para la implementación de bases de datos en la Encuesta Nacional de Empleo e ingresos. Instituto Nacional de Estadística Guatemala (2003). "Hacia un sistema integrado de encuestas de hogares en los países de América Latina".

ACTIVIDADES APLICATIVAS DE LA UNIDAD 3

Ahora procede a desarrollar las actividades aplicativas de la unidad 3

ACTIVIDAD 1

**Realizando un ejercicio de procesamiento de datos
ACTIVIDAD PRESENCIAL**

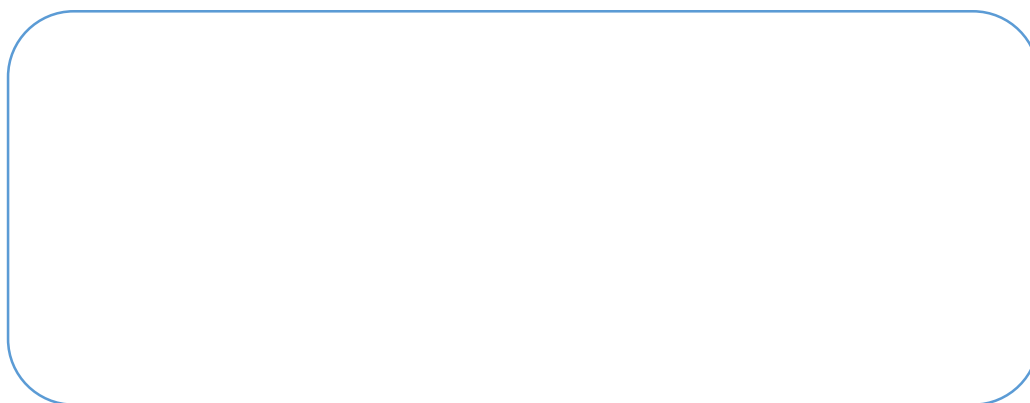
1. Asiste a la clase programada en el sílabo.
2. Participa en las siguientes actividades:
 - Manejo y reducción de base de datos.
 - Triangulación de fuentes y técnicas.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 3

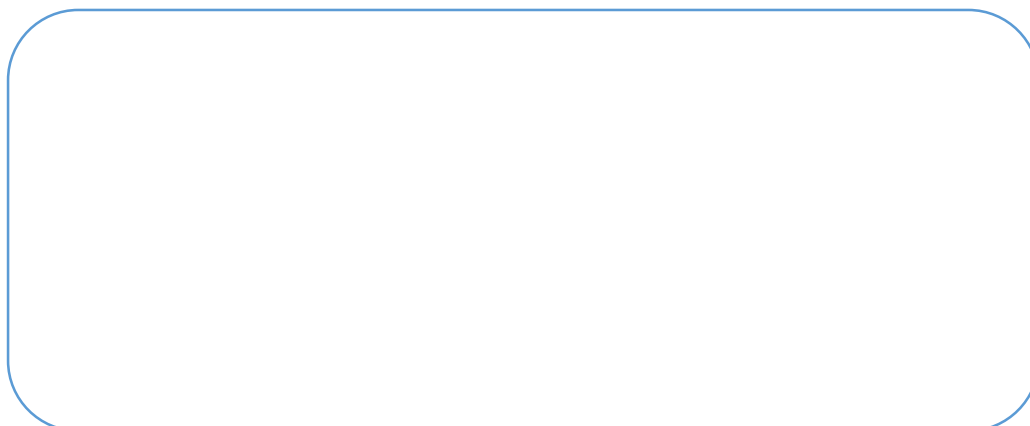
HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

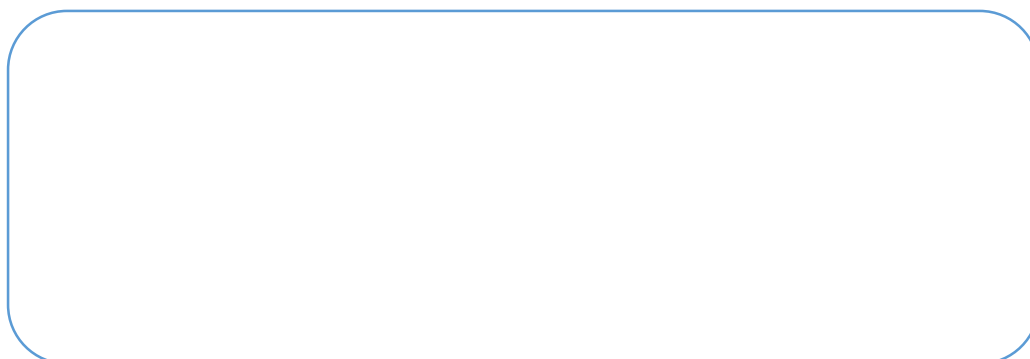
1. Explica brevemente los pasos para el diseño de una base de datos.



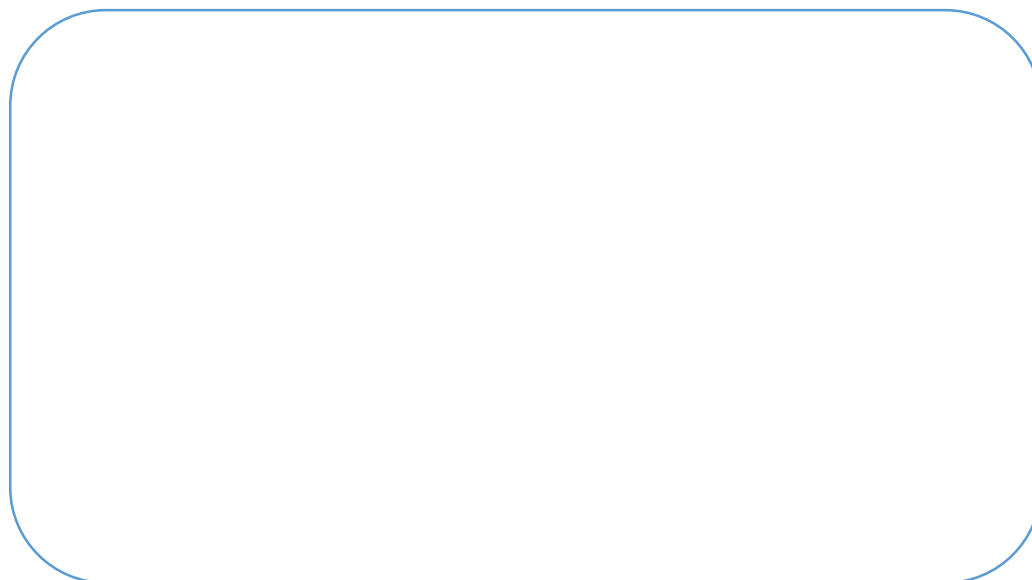
2. Explica los criterios para la consistencia de datos.



3. ¿Cómo se determina la precisión estadística de los estimadores?

A large, empty rounded rectangular box with a thin blue border, intended for the student to write their answer to question 3.

4. Describe los principales pasos durante la etapa de descubrimiento y codificación para el procesamiento de datos cualitativos.

A large, empty rounded rectangular box with a thin blue border, intended for the student to write their answer to question 4.



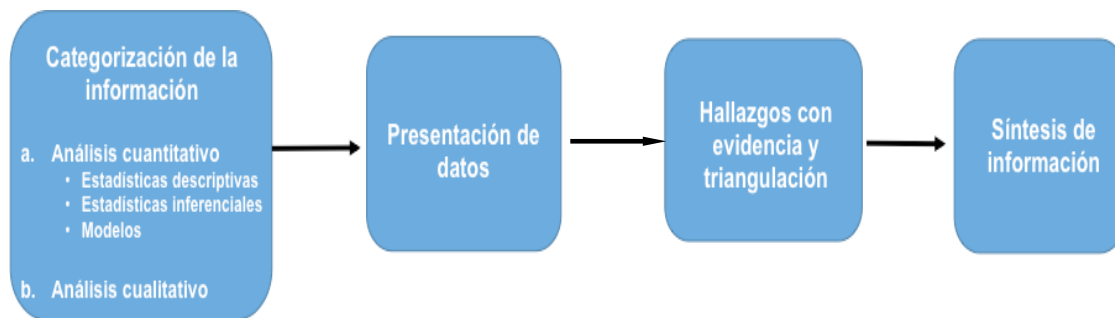
UNIDAD DE APRENDIZAJE 4: CATEGORIZACIÓN DE LA INFORMACIÓN E IDENTIFICACIÓN DE HALLAZGOS SIGNIFICATIVOS

La categorización de información es el primer paso dentro de la etapa de análisis de información luego de contar con la información cuantitativa y cualitativa, validada e ingresada en los respectivos software/programas. La categorización de información consiste en **llevar dicho conjunto de datos a unidades y/o dimensiones sintetizadas** que constituirán las evidencias sobre las que se sostiene el análisis de la evaluación. Por ejemplo: si se ha realizado una encuesta de hogares como parte de la evaluación de un programa nutricional que tiene como indicador la tasa de desnutrición crónica en niños menores de cinco años, lo que buscaremos en esta etapa es calcular dicha tasa a partir de los datos registrados en la base sobre el peso de los niños(as).

En el caso de la información cualitativa, de manera análoga lo que se buscará es identificar las grandes dimensiones de respuesta que se evidencian a partir de los datos recopilados. Por ejemplo: en una evaluación que analiza la percepción de los usuarios sobre los mecanismos de participación en un determinado programa, la información producto de los grupos focales debe permitirnos identificar las principales dimensiones de respuesta en torno al tema, así como opiniones a favor y en contra con las que podamos determinar la percepción general sobre este.

La presente unidad tiene por objetivo brindar los elementos conceptuales básicos que se requiere para implementar e interpretar la información obtenida a partir de datos cuantitativos y cualitativos. Para ello, considerando que los procedimientos de análisis de datos cuantitativos son ampliamente distintos a los de datos cualitativos, se presentan dichos procedimientos en secciones separadas. La figura 8 presenta la secuencia de desarrollo de esta unidad.

Figura 8. Secuencia de organización de contenidos de la unidad 4



El rubro 4.1 presenta los procedimientos de categorización de información en dos subsecciones: análisis cuantitativo y análisis cualitativo; seguidamente, el rubro 4.2 brinda pautas generales para la presentación de datos; el rubro 4.3 explica en qué consiste el proceso de triangulación de datos y el rubro 4.4 brinda criterios y pautas para la interpretación de la información obtenida.

Finalmente, cabe indicar que los hallazgos de información cuantitativa y cualitativa no se presentan de forma separada en una evaluación; estos se conjugan y analizan de manera conjunta a fin de permitirnos emitir conclusiones sobre el plan, programa o/proyecto en evaluación respondiendo a las preguntas y dimensiones de evaluación.

4.1. Categorización de la información

En esta etapa se elaboran las estadísticas, se interpretan los resultados de grupos focales, entrevistas abiertas y otros tipos de instrumentos cualitativos y se calculan los indicadores del programa y/o proyecto en evaluación. Para ello, en los siguientes dos ítems se presentan algunos conceptos y elementos básicos cuantitativos y cualitativos que permiten sintetizar los datos cuantitativos y cualitativos, respectivamente.

4.1.1. Análisis cuantitativo

El análisis de los datos de manera general depende de tres factores: (i) el nivel de medición de las variables, (ii) la manera como se hayan formulado las hipótesis y (iii) el interés del evaluador.

En lo operativo el evaluador busca, en primer término, **describir los datos y posteriormente efectuar análisis estadísticos** para relacionar sus variables y/o calcular indicadores. Es decir, reporta estadística descriptiva para cada una de las variables de la base (ítems) y luego para cada una de las variables del estudio (ver matriz de evaluación de la sección I-unidad 2); finalmente aplica estadística inferencial para probar sus hipótesis.

Los tipos o métodos de análisis cuantitativo o estadístico son variados, pero cabe señalar que el análisis no es indiscriminado, cada método tiene su razón de ser y un propósito específico; por ello no deben hacerse más análisis de los necesarios. Finalmente, si lo que estamos implementando es una evaluación de impacto, lo que utilizamos como herramienta son modelos cuantitativos (explicativos) para estimar los impactos y determinar la causalidad.

Cabe mencionar que el análisis cuantitativo incluye los siguientes pasos: elaboración de estadísticas descriptivas, estadísticas inferenciales para comprobación de hipótesis y, tercero, estimación de modelos (de ser el caso).

a. Estadística descriptiva

La primera tarea que tenemos es describir los datos, los valores o las puntuaciones obtenidas para cada variable. Para ello haremos uso de la estadística descriptiva. Esto se logra al describir la distribución de las puntuaciones o frecuencias de cada variable.

i. Distribución de frecuencias y gráficos.

Es un conjunto de puntuaciones ordenadas en sus respectivas categorías. Pueden completarse agregando los porcentajes de casos en cada categoría, los porcentajes válidos (excluyendo los valores perdidos) y los porcentajes acumulados (porcentaje de lo que se va acumulando en cada categoría, desde la más baja hasta la más alta). La tabla 2 muestra un ejemplo de distribución de frecuencias.

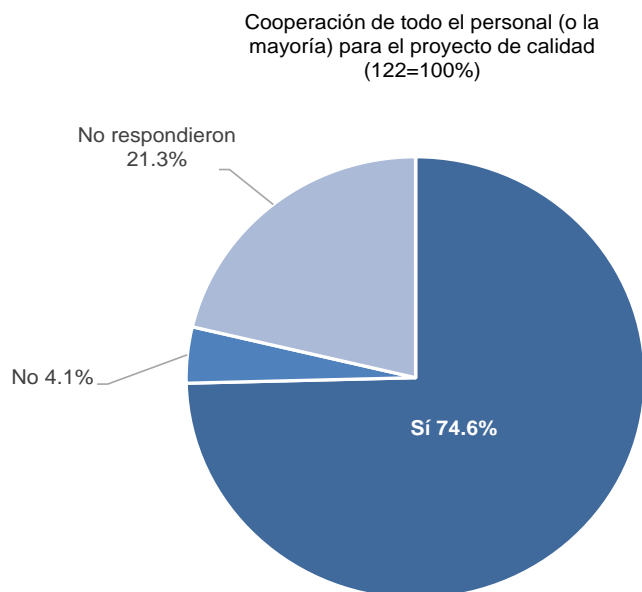
**Tabla 2. Distribución de frecuencias
¿Cómo calificaría usted el INFObras en relación a: Amigable?**

	Código	Frecuencia	%	% Acumulado
Excelente	5	18	40%	40%
Muy bueno	4	18	40%	80%
Bueno	3	7	16%	96%
Malo	2	0	0%	96%
Muy malo	1	0	0%	96%
No responde	999	2	4%	100%
Total		45	100%	

Elaboración: Alejandro Bardales, 2014.

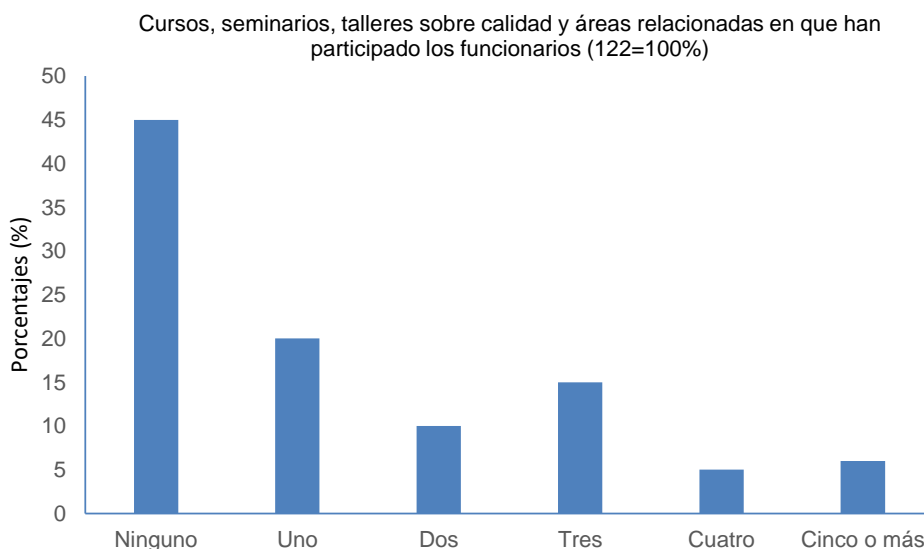
La información también puede ser mostrada en gráficos, los cuales pueden ser de diferentes tipos. A continuación se presentan dos ejemplos. El primer gráfico corresponde a un gráfico circular, y el segundo corresponde a un gráfico de barras verticales, también llamado histograma, aunque es más común usar este término cuando los valores corresponden a una variable continua.

Gráfico 1. Gráficas circulares



Fuente: Hernández et al (2006).

Gráfico 2
Histogramas



Fuente: Hernández et al (2006).

ii. Medidas de tendencia central:

Son los valores medios o centrales de una distribución. Son tres: moda, mediana y media. El nivel de medición de la variable determina cuál es la medida de tendencia central apropiada.

La **moda** es la categoría o puntuación que ocurre con mayor frecuencia. Por ejemplo, en el gráfico anterior, la moda es "Ninguno". Se utiliza tanto con variables cuantitativas como cualitativas.

La **mediana** es el valor que divide la distribución en dos grupos con igual cantidad de casos, con lo cual la mitad de ellos cae por debajo de la mediana y la otra mitad se ubica por encima. La mediana refleja la posición intermedia de la distribución. Por ejemplo, si los datos obtenidos fueran:

24 31 35 35 38 43 45 50 57

La mediana es 38, porque deja cuatro casos por encima (43, 45, 50 y 57) y cuatro casos por debajo (35, 35, 31 y 24). Parte a la distribución en dos

mitades. En general, para descubrir el caso al que corresponde la puntuación que constituye la mediana de una distribución, simplemente se aplica la fórmula: $\frac{N+1}{2}$.

En nuestro ejemplo anterior tenemos nueve casos y aplicando la fórmula $\frac{9+1}{2} = 5$, entonces buscamos el quinto valor y este es la mediana (38).

La fórmula no nos proporciona directamente el valor de la mediana, sino el número de caso en donde está la mediana. La mediana es particularmente útil cuando hay valores extremos o fuera de rango en la distribución, y solo tiene sentido su aplicación con datos cuantitativos.

La **media** es la medida de tendencia central más utilizada y puede definirse como el promedio aritmético de una distribución. Se simboliza como \bar{X} , y es la suma de todos los valores dividida entre el número de casos. Es una medida solamente aplicable a mediciones por intervalos o de razón. Carece de sentido para variables medidas en un nivel nominal u ordinal. Su fórmula es:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{N}$$

Donde N es el número de casos.

Por ejemplo, si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 8

La media sería igual a:

$$\bar{X} = \frac{8 + 7 + 6 + 4 + 3 + 2 + 6 + 9 + 8}{9} = 5,888$$

La fórmula simplificada es:

$$\bar{X} = \frac{\sum X}{N}$$

Esto significa que debe efectuarse una sumatoria, X es el símbolo de una puntuación y N es el número total de casos o puntuaciones. Para el ejemplo:

$$\bar{X} = \frac{53}{9} = 5,888$$

La media es sensible a valores extremos, los cuales pueden "atraerla". Si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 20

La media sería:

$$\bar{X} = \frac{65}{9} = 7,22$$

En este caso vemos que el valor 20, que es mucho mayor que el resto, ha empujado hacia arriba el promedio.

iii. Medidas de variabilidad

Indican la dispersión de los datos en la escala de medición y responden a la pregunta: ¿dónde están diseminadas las puntuaciones o los valores obtenidos? Las medidas de la variabilidad más utilizadas son rango, desviación estándar y varianza.

El **rango**, también llamado recorrido, es la diferencia entre la puntuación mayor y la puntuación menor. Se calcula así: $X_M - X_m$ (puntuación mayor menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

El rango será: $33-17=16$

Cuanto más grande sea el rango, mayor será la dispersión de los datos de una distribución.

La **varianza** es del promedio de desviación de las puntuaciones con respecto a la media, elevadas al cuadrado y se simboliza s^2 .

Es un concepto estadístico muy importante, ya que muchas de las pruebas de hipótesis se fundamentan en ella. Su fórmula es la siguiente:

$$s^2 = \frac{\sum(X - \bar{X})^2}{N}$$

Esto es, la desviación de cada puntuación respecto a la media se eleva al cuadrado, se suman todas las desviaciones cuadradas y se divide entre el número total de puntuaciones

Entre dos distribuciones que tienen un mismo promedio, tendrá mayor dispersión aquella que tenga la mayor varianza.

La varianza solo se utiliza en variables medidas por intervalos o de razón.

La **desviación estándar o típica** es la raíz cuadrada de la varianza. Esta media se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la distancia de los datos con respecto a la media, mayor será la desviación estándar. Se simboliza con s o la sigma minúscula σ y su fórmula esencial es:

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{N}}$$

Entre dos distribuciones que tienen un mismo promedio, tendrá mayor dispersión aquella que tenga la mayor desviación estándar.

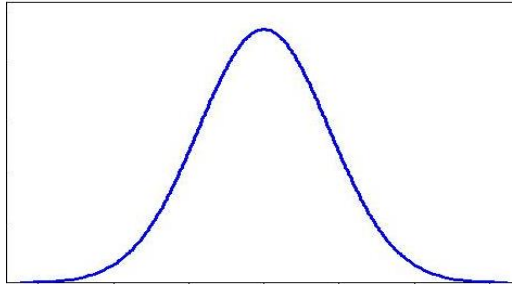
La desviación estándar solo se utiliza en variables medidas por intervalos o de razón.

iv. Asimetría

Es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada curva normal (la cual es simétrica con respecto a la media y se presenta en la figura 9 y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría=0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por

debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media).

Figura 9. Distribución normal



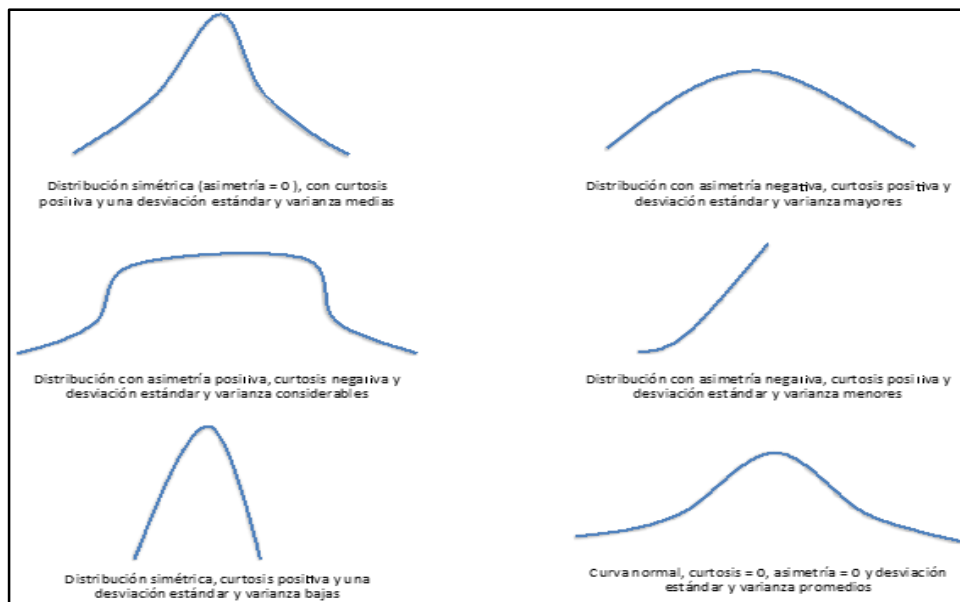
v. Curtosis

Es un indicador de lo plana o "picuda" que es una curva. Cuando es cero (curtosis = 0), significa que puede tratarse de una curva normal. Si es positiva, quiere decir que la curva, la distribución o el polígono es más picuda(o) o elevada(o). Si la curtosis es negativa, indica que es más plana la curva.

La asimetría y la curtosis requieren de un nivel de medición por intervalos.

En la figura 10 se muestran ejemplos de curvas con su interpretación.

Figura 10. Ejemplos de distribuciones, asimetría y curtosis



Fuente: Hernández et al (2006).

Para revisar con detalle el tema tratado, revisa la siguiente lectura:

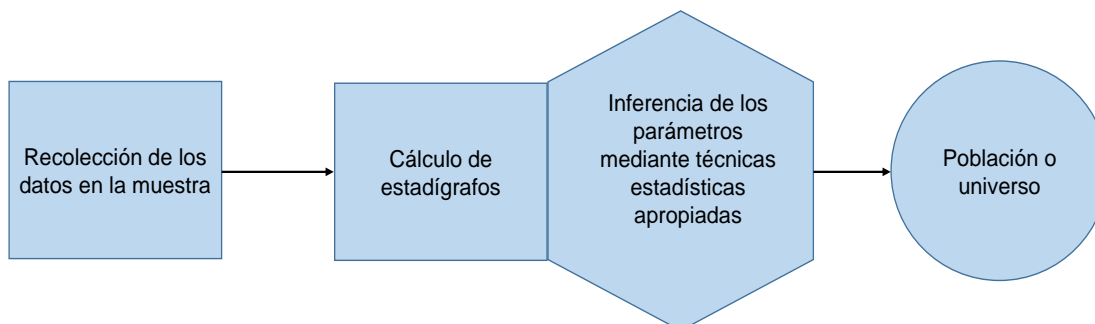
Salvarrey. L. Curso de Estadística Básica. Alto (Uruguay), Universidad de la Republica Regional Norte Sede Alto. 2000. Disponible en:

<http://guajiros.udea.edu.co/descriptiva/articulos/Curso%20de%20Estadística%20Basica.pdf>

b. Estadística inferencial

Es frecuente que el propósito de la investigación vaya más allá de describir las distribuciones de las variables: se pretende generalizar los resultados obtenidos en la muestra a la población o el universo. Los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan estadígrafos; la media o desviación estándar de la distribución de una muestra son estadígrafos. A las estadísticas de la población o al universo se les conoce como parámetros. Los parámetros no son calculados porque no se recolectan datos de toda la población, pero pueden ser inferidos de los estadígrafos, de ahí el nombre de estadística inferencial. El procedimiento se esquematiza en la figura 11.

Figura 11. Procedimiento de la estadística inferencial



Fuente: Hernández et al (2006).

La inferencia de los parámetros se lleva a cabo mediante técnicas estadísticas apropiadas y se utiliza, entre otras cosas, para probar hipótesis.

i. Intervalo de confianza

Aunque en las investigaciones por muestreo las medias obtenidas constituyen estimaciones (aproximaciones) de las medias poblacionales (parámetros que se obtendrían si se toma datos de toda la población en estudio), se debe tener en cuenta que se podrían tomar varias muestras diferentes con exactamente el mismo diseño y los resultados podrían ser diferentes; sin embargo, en la realidad se toma solo una muestra. Por esa razón se dice que el resultado de una muestra es una estimación del parámetro y que puede tener un rango de variación.

Por esta razón es necesario estimar un intervalo dentro del cual se encontraría el parámetro poblacional con un nivel de confianza establecido.

Los intervalos de confianza se obtienen con la siguiente fórmula:

$$\bar{X} \pm z\sigma$$

Donde \bar{X} es el promedio obtenido en la muestra, z es el valor correspondiente a un nivel de confianza (generalmente se usa 1,96, que corresponde a un nivel de confianza del 95%) y σ es el error muestral (desviación estándar) de la estimación. El valor $1,96\sigma$ es conocido como el

margen de error. Los paquetes de análisis estadístico proveen estas estimaciones.

Por ejemplo, veamos el caso de la exposición diaria a la televisión (en fin de semana) por parte de los niños de Lima:

$$\text{Media} = 2,9 \text{ horas}$$

$$s = 1,2 \text{ horas}$$

$$S\bar{x} = 0,0679 \text{ (error muestral)}$$

$$\text{Nivel de confianza} = 0,95 \text{ (} z = 1,96 \text{)}$$

$$\begin{aligned} \text{Intervalo de confianza} &= 2,9 \pm (1,96)(0,0679) = 2,9 \pm (0,133) \\ &= (2,767, 3,303) \end{aligned}$$

Con esta información se puede decir que en promedio los niños de Lima ven 2,9 de horas de televisión en fines de semana, que esta estimación tiene un margen de error de +/- 0,133 horas con el 95% de confianza, y que, por tanto, el verdadero promedio puede estar entre 2,8 y 3,3 horas (con valores redondeados a un decimal).

ii. Prueba de hipótesis

Una hipótesis en el contexto de la estadística inferencial es una proposición respecto a uno o varios parámetros. Lo que el investigador hace por medio de la **prueba de hipótesis** es determinar si la hipótesis es congruente con los datos obtenidos en la muestra. La hipótesis se retiene como un valor aceptable del parámetro, si es congruente con los datos. Si no lo es, se rechaza (pero los datos no se descartan). Para comprender lo que es la prueba de hipótesis en la estadística inferencial es necesario revisar el concepto de distribución muestral y nivel de significancia.

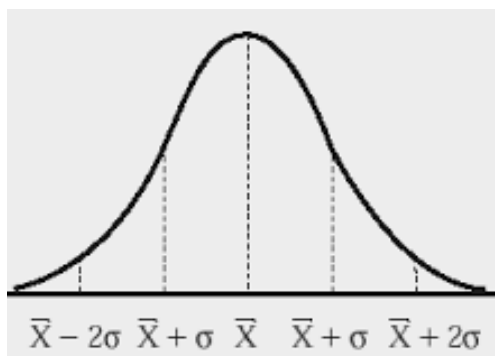
iii. Distribución muestral

Es un conjunto de valores sobre una estadística calculada de todas las muestras posibles de determinado tamaño que se podrían obtener de una población. Las distribuciones muestrales de medias son probablemente las más conocidas. Expliquemos este concepto con un ejemplo. Supongamos que nuestro universo son los hogares del sector rural y deseamos averiguar

el tiempo promedio de permanencia en el programa JUNTOS. De este universo podría extraerse una muestra representativa. Vamos a suponer que el tamaño adecuado de muestra es 512 hogares ($n=512$). Del mismo universo podrían extraerse diferentes muestras de 512 hogares.

Teóricamente, incluso podría elegirse al azar una, dos, tres, cuatro muestras, y las veces que fuera necesario hacerlo, hasta agotar todas las muestras posibles de 512 hogares del sector rural (todos los hogares serían seleccionados en varias muestras). En cada muestra se obtendría una media del tiempo de permanencia en JUNTOS. Tendríamos pues, una gran cantidad de medias, tantas como las muestras extraídas ($\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4, \bar{X}_5, \dots, \bar{X}_k$). Y con estas elaboraríamos la distribución de medias. Habría muestras que en promedio, pasarán más tiempo en JUNTOS que otras. Este concepto se representa en la figura 12.

Figura 12. Distribución normal



Si calculáramos la media de todas las medias de las muestras, obtendríamos el valor de la media poblacional.

Muy rara vez se obtiene la distribución muestral (la distribución de las medias de todas las medias posibles). Es más bien un concepto teórico definido por la estadística para los investigadores. Lo que comúnmente hacemos es extraer una sola muestra.

En el ejemplo de los beneficiarios de JUNTOS, solo una de las líneas verticales de la distribución muestral presentada en la figura 12 es la media obtenida para nuestra única muestra seleccionada de 512 personas. La pregunta es: ¿nuestra media está cerca de la media de la distribución muestral? Si está cerca podremos tener una estimación precisa de la media poblacional (el parámetro poblacional es prácticamente el mismo que el de la distribución muestral). Esto se expresa en el teorema del límite central.

Teorema del límite central

Si una población (no necesariamente normal) tiene media m y de desviación estándar s , la distribución de las medias en el muestreo aleatorio realizado en esta población tiende, al aumentar n , a una distribución normal de media m y desviación estándar $\frac{s}{\sqrt{n}}$ donde n es el tamaño de muestra. El teorema especifica que la distribución muestral tiene una media igual a la de la población, una varianza igual a la varianza de la población dividida entre el tamaño de muestra (su desviación estándar es $\frac{\sigma}{\sqrt{n}}$ y se distribuye normalmente). La desviación estándar (s) es un parámetro normalmente desconocido, aunque es posible estimarlo por la desviación estándar de la muestra.

Fuente: Hernández et al (2006).

iv. Nivel de significancia

La probabilidad de que un evento ocurra oscila entre (0) y uno (1), donde cero significa la imposibilidad de ocurrencia y uno la certeza de que el fenómeno ocurra. Al lanzar al aire una moneda no cargada, la probabilidad de que salga "sello" es de 0,50 y la probabilidad de que la moneda caiga en "cara" también es de 0,50. Con un dado, la probabilidad de obtener cualquiera de sus caras al lanzarlo es de $1/6 = 0,1667$. La suma de posibilidades siempre es de uno.

Aplicando el concepto de probabilidad a la distribución muestral, tomaremos el área de esta como 1,00; en consecuencia, cualquier área comprendida entre dos puntos de la distribución corresponderá a la probabilidad de la distribución. Para probar hipótesis inferenciales respecto a la media se debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la mitad de la distribución muestral. Si es baja, se dudará de generalizar a la población. Si es alta, se podrán realizar generalizaciones. Es

aquí donde entra el **nivel de significancia** o **nivel alfa** (α), el cual es un nivel de la probabilidad de equivocarse y se fija antes de probar hipótesis inferenciales.

Se explicará esto con un ejemplo. Si fuéramos a jugar un boleto de lotería y tuviéramos un 95% de probabilidad de ganar el premio mayor contra solo 5% de perder, resulta claro que jugaríamos, siempre y cuando nos aseguraran ese 95% en favor. O bien, si nos dieran 95 boletos de 100 las posibilidades de ganar serían mayores. No tendríamos la certeza total ya que esta no existe.

Algo similar hace el investigador. Obtiene una estadística en una muestra (por ejemplo, la media) y analiza qué porcentaje tiene de confianza en que dicha estadística se acerque al valor de la distribución muestral (que es el valor de la población o el parámetro). Busca un alto porcentaje de confianza, una probabilidad elevada, porque sabe que tal vez haya error de muestreo y, aunque la evidencia parece mostrar una aparente "cercanía" entre el valor calculado en la muestra y el parámetro, tal "cercanía" puede no ser real o deberse a errores en la selección de la muestra.

Para suponer que tal cercanía es real y no le corresponde un error de muestreo existen dos niveles convenidos:

- ⇒ *El nivel de confianza de 0,95*, el cual implica que el investigador tiene 95% de seguridad para generalizar sin equivocarse y solo 5% en contra. En términos de probabilidad, 0,95 y 0,05, respectivamente; ambos suman la unidad.
- ⇒ *El nivel de significancia de 0,01*, el cual implica que el investigador tiene 99% en su favor y 1% en contra para generalizar sin temor.

El nivel de significancia es un valor de certeza que el investigador fija *a priori*, respecto a no equivocarse. Cuando uno lee en un reporte de investigación que los resultados fueron significativos al nivel de 0,05 ($p < 0,05$), indica que existe 5% de posibilidad de error al aceptar la hipótesis, correlación o valor obtenido al aplicar una prueba estadística; o 5% de riesgo de que se rechace una hipótesis nula cuando era verdadera.

Procedimiento para aceptar o rechazar la hipótesis sobre la media poblacional

1. Establecer una hipótesis del parámetro poblacional.
2. Definir el nivel de significancia. Por ejemplo, $\alpha = 0,05$.
3. Recolectar los datos de una muestra representativa.
4. Estimar la desviación estándar de la distribución muestral de la media utilizando la siguiente fórmula:

$$S_{\bar{x}} = \frac{s}{\sqrt{n}}$$

5. Transformar la media de la muestra en una puntuación z , en el contexto de la distribución muestral, con una variación de la fórmula ya conocida para obtener puntuaciones z :

$$z = \frac{X - \bar{X}}{S_{\bar{x}}}$$

6. En la tabla de áreas bajo la curva normal buscar aquella puntuación z que deje a 2.5% por encima de ella, dado que es una distribución simétrica y se aplica para ambos lados de la media. Este valor es 1,96.
7. Comparar la media de la muestra transformada a puntuación z con el valor 1,96; si es menor, aceptar la hipótesis; si es mayor rechazarla.

Fuente: Hernández et al 2006.

v. Análisis paramétrico:

Hay dos tipos de análisis estadístico que pueden realizarse para probar hipótesis: los análisis paramétricos y los no paramétricos. Cada tipo posee características y presuposiciones que lo sustentan; la elección de qué clase de análisis efectuar depende de estas presuposiciones. De igual forma, cabe destacar que en una misma investigación es posible llevar a cabo análisis paramétricos para algunas hipótesis y variables y análisis no paramétricos para otras. Asimismo, los análisis a realizar dependen de la hipótesis que hayamos formulado y el nivel de medición de las variables que las conforman.

Para realizar análisis paramétrico debe partirse de los siguientes supuestos:

- ⇒ La distribución poblacional de la variable dependiente es normal: el universo tiene una distribución normal.
- ⇒ El nivel de medición de la variable es dependiente por intervalos o razón.
- ⇒ Cuando dos o más poblaciones son estudiadas, tienen una varianza homogénea: las poblaciones en cuestión poseen una dispersión similar en sus distribuciones.

Estos criterios pueden resultar demasiado rigurosos y algunos investigadores solo basan su análisis en el tipo de hipótesis y los niveles de medición de las variables. Existen diversas pruebas estadísticas paramétricas entre las cuales nos interesa la prueba t.

Prueba t

Es una prueba estadística para evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias. Se propone como hipótesis nula que los grupos no difieren significativamente. Los grupos pueden ser dos escuelas contrastadas en los resultados a un examen.

La comparación se realiza sobre una variable (teóricamente dependiente). Si hay diferentes variables, se efectuarán varias pruebas t (una por cada variable), y la razón que motiva la creación de los grupos puede ser una variable independiente. Por ejemplo un experimento con dos grupos, donde a uno se le aplica tratamiento y al otro no, que viene a ser el grupo control.

El valor t se obtiene en muestras grandes mediante la fórmula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

En donde \bar{X}_1 es la media del primer grupo, \bar{X}_2 la media del segundo grupo, s_1^2 representa la varianza del primero, n_1 es el tamaño del primer grupo, s_2^2 es la varianza del segundo y n_2 es el tamaño del segundo grupo. En realidad,

el denominador es el error estándar de la distribución muestral de la diferencia entre medias.

Para saber si el valor t es significativo, se aplica la fórmula y se calculan los grados de libertad. La prueba t se basa en una distribución muestral o poblacional de diferencia de medias conocida como la distribución t de Student. Esta distribución se identifica por los grados de libertad, los cuales constituyen el número de maneras en que los datos pueden variar libremente. Son determinantes, ya que nos indican qué valor debemos esperar de t , dependiendo del tamaño de los grupos que se comparan. Cuanto mayor número de grados de libertad se tenga, la distribución t de Student se acercará más a ser una distribución normal; usualmente, si los grados de libertad exceden los 120, la distribución normal sí utiliza como una aproximación adecuada de la distribución t de Student.

Los grados de libertad se calculan con la fórmula siguiente, en la que n_1 y n_2 son el tamaño de los grupos que se comparan e indican cuántos casos fueron usados para calcular en valor estadístico en particular.

$$\text{grados de libertad} = (n_1 + n_2) - 2$$

Una vez calculados el valor t y los grados de libertad, se elige el nivel de significancia y se compara el valor obtenido contra el valor que le correspondería. Si el valor calculado es igual o mayor al que aparece en la tabla, se acepta la hipótesis de investigación. Pero si es menor, se acepta la hipótesis nula.

En la tabla se busca el valor con el cual vamos a comparar el que hemos calculado, con base en el nivel de confianza elegido (0,05 o 0,01) y los grados de libertad. Los niveles de confianza adquieren el significado del que se ha hablado (el 0,05 significa 95% de que los grupos en realidad difieran significativamente entre sí y 5% de posibilidad de error).

Cuanto mayor sea el valor t calculado respecto al valor de la tabla y menor sea la posibilidad de error, mayor será la certeza en los resultados.

Los paquetes estadísticos proveen el "pvalue", que es el tamaño de la región de rechazo de la hipótesis. Cuando el "pvalue" es menor que el valor de

significancia establecido, siendo 0,05 el más usado, se acepta la hipótesis de igualdad; si el "pvalue" es mayor que 0,05, se dice que las diferencias son significativas.

vi. Análisis no paramétrico

Los análisis no paramétricos tienen menos restricciones para su aplicación y responden a las siguientes consideraciones:

- ⇒ La mayoría de estos análisis no requieren presupuestos acerca de la forma de la distribución poblacional. Aceptan distribuciones no normales.
- ⇒ Las variables no necesariamente tienen que estar medidas en un nivel por intervalos o de razón; pueden analizar datos nominales u ordinales. De hecho, si se quiere aplicar análisis no paramétricos a datos por intervalos o razón, estos se convierten a categorías discretas (a unas cuantas). Las variables deben ser categóricas. Entre las pruebas no paramétricas más utilizadas se encuentra la chi cuadrada o χ^2 .

La prueba chi cuadrada o χ^2

Es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas. La prueba *chi* cuadrado no considera relaciones causales y el nivel de medición de las variables es nominal u ordinal (o intervalos o razón reducidos a ordinales).

Se calcula por medio de una tabla de contingencia o tabulación cruzada, que es una tabla de dos dimensiones, y cada dimensión contiene una variable. A su vez, cada variable se subdivide en dos o más categorías.

La tabla 3 presenta un ejemplo de *tabla de contingencia* o tabulación cruzada. Las variables aparecen señaladas al lado izquierdo y en la parte superior de la tabla, cada una con sus dos categorías. Se dice que se trata de una tabla 2x2, donde cada dígito significa una variable y el valor de este indica el número de categorías de la variable.

Después se calculan las frecuencias esperadas para cada celda. En esencia, la *chi* cuadrada es una comparación entre la tabla de frecuencias observadas y la denominada tabla de frecuencias esperadas, la cual constituye la tabla que esperaríamos encontrar si las variables fueran estadísticamente independiente o no estuvieran relacionadas.

Tabla 3. Tabla de contingencia de población por condición de pobreza, según región de residencia

		Línea de pobreza monetaria		Total
		Pobre	No pobre	
Región	Costa	180	100	280
	Sierr	190	280	470
	Selva	170	120	290
	Total	540	500	1040

Elaboración: Alejandro Bardales, 2014.

La *chi* cuadrada es una prueba que establece como hipótesis nula “no relación entre las variables” (en el ejemplo sería “no hay relación entre la condición de pobreza y la región natural”); el investigador evalúa si en su caso esto es cierto estadísticamente o no, analiza si las frecuencias observadas son diferentes de lo que pudiera esperarse en caso de ausencia de correlación. La lógica es así: “Si no hay relación entre las variables, debe tenerse una tabla así (la de las frecuencias esperadas). Si hay relación, la tabla que obtengamos como resultado en nuestra investigación tiene que ser muy diferente respecto de la tabla de frecuencias esperadas”.

La frecuencia esperada de cada celda o casilla se calcula mediante la siguiente *fórmula aplicada a la tabla de frecuencias observadas*.

$$fe = \frac{(Total\ o\ marginal\ de\ renglón)(total\ o\ marginal\ de\ columna)}{n}$$

En donde n es el número total de observaciones. Para la primera celda (región costa y pobre por línea de pobreza monetaria) la frecuencia esperada sería:

$$fe = \frac{(280)(540)}{1040} = 145.4$$

La tabla 4 presenta todas las frecuencias esperadas calculadas.

Tabla 4. Frecuencias esperadas

145.4	134.6	280	Total de renglón
244	226	470	
150.6	139.4	290	
540	500	1040	n
Total de columnas			

Elaboración: Alejandro Bardales, 2014.

Una vez obtenidas las frecuencias esperadas, se aplica la siguiente *fórmula de chi cuadrada*:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Donde Σ significa sumatoria. (O) es la frecuencia observada en cada celda y (E) es la frecuencia esperada en cada celda. Es decir, se calcula para cada celda la diferencia entre la frecuencia observada y la esperada; esta diferencia se eleva al cuadrado y se divide entre la frecuencia esperada. Finalmente, se suman tales resultados y la sumatoria es el valor de χ^2 obtenido. El procedimiento de cálculo de la χ^2 se visualiza en la tabla 5.

Tabla 5. Procedimiento para calcular la chi cuadrada

Celda	O	E	O-E	(O-E) ²	(O-E) ² /E
Pobre/Costa	180	145.4	34.6	1197.16	8.23
Pobre/Sierra	190	244.4	-54.4	2959.36	12.11
Pobre/Selva	170	150.6	19.4	376.36	2.50
No pobre/Costa	100	134.6	-34.6	1197.16	8.89
No pobre/Sierra	280	226	54	2916	12.90
No pobre/Selva	120	139.4	-19.4	376.36	2.70
				χ^2	47.34

Elaboración: Alejandro Bardales, 2014.

El valor χ^2 para los valores observados es 47,33

Al igual que la prueba *t*, la *chi cuadrada* proviene de una distribución muestral denominada distribución χ^2 ; y para saber si un valor de χ^2 es o no significativo, debemos calcular los grados de libertad. Estos se obtienen mediante la siguiente fórmula:

$$gl = (r - 1)(c - 1)$$

En donde *r* es el número de renglones del cuadrado de contingencia y *c* el número de columnas. En nuestro caso:

$$gl = (3 - 1)(2 - 1) = 2$$

Y acudimos con los grados de libertad que corresponden a la tabla eligiendo nuestro nivel de confianza (0,05 o 0,01). Si nuestro valor calculado de χ^2 es igual o superior al de la tabla, decimos que las variables están relacionadas (χ^2 fue significativa). En el ejemplo, el valor crítico que requerimos empatar o superar al nivel de 0,05 es 5,991. El valor de χ^2 calculado por nosotros es 47,33, que es muy superior al de la tabla. Por lo tanto χ^2 resulta significativa y se rechaza la hipótesis de que no hay relación entre la condición de pobreza y la región natural.

c. Uso de modelos para el análisis cuantitativo

Existe una dimensión de evaluación que usualmente se explora en las evaluaciones de impacto y que permite responder a preguntas que evidencien relaciones de causa y efecto. Es decir: los efectos esperados que se encontraron en los beneficiarios, ¿son realmente atribuibles a la intervención? ¿O es que son otros los factores que influyen en el estado actual de los beneficiarios?

Para responder a este tipo de preguntas, es necesario ir un paso más allá en el uso de métodos estadísticos. Es necesario recurrir a diseños experimentales o cuasiexperimentales, que son los únicos que nos permiten establecer relaciones de causalidad en la evaluación. En los diseños de causalidad se usan modelos estadísticos, los cuales son representaciones simplificadas de la realidad que nos permiten por procedimientos matemáticos y estadísticos estimar la relación e influencia de una o más variables sobre otras. Estos modelos se pueden también emplear en evaluaciones que usan datos de corte transversal y que nos permiten dar indicios de variables relacionadas con las actividades o estrategias del programa y variables de resultados del programa.

Como se puede apreciar en el cuadro 19, existe una serie de modelos cuantitativos; aquí se explican los más frecuentemente utilizados.

Cuadro 19. Cuadro resumen de modelos cuantitativos más usados

Modelo	Variable dependiente	Variable independiente	Uso en la evaluación
Regresión lineal	Cuantitativa (Y= número real).	Conjunto de variables cuantitativas y/o cualitativas.	Se usa para evaluar la influencia de un conjunto de variables sobre una variable cuantitativa de interés. Utilizado para ver el cambio cuantitativo de una intervención. (¿Tuvo un efecto positivo la intervención en el ingreso de las personas? ¿De cuánto fue?).
Regresión logística	Cuantitativa discreta que puede tener dos	Conjunto de variables	Se usa para medir la probabilidad de cambio en la unidad de la variable de interés a partir de

Modelo	Variable dependiente	Variable independiente	Uso en la evaluación
	o más valores. La más común es dicotómica ($Y=0$ ó $Y=1$).	cuantitativas y/o cualitativas.	cambios de un conjunto de variables independientes (por ejemplo: probabilidad de participar en un programa social, de retirarse de uno de empleo, etc.), así como el grado de influencia de cada variable independiente.
Análisis factorial	De preferencia cuantitativa, aunque en algunos casos se usan variables cualitativas.	Conjunto de variables cuantitativas y/o cualitativas.	Es usado cuando se tiene un gran número de variables y se necesita reducir dimensiones para brindar una mejor explicación al fenómeno en cuestión.
Modelos multinivel	Cuantitativa ($Y=$ número real).	Conjunto de variables cuantitativas y/o cualitativas.	Son generalizaciones del modelo lineal y no lineal. Son usados cuando los datos tienen estructuras complejas de variabilidad, como en una muestra de varias etapas. Por ejemplo en una muestra de estudiantes, va a haber una variabilidad entre los de una misma región, pero también va a haber variabilidad entre las regiones. Los modelos de regresión utilizados en estos casos consideran a los parámetros como aleatorios, mientras que en la regresión clásica los parámetros son fijos.

Elaboración: Alejandro Bardales, 2014.

Entre los modelos presentados en el cuadro anterior, el más usado en el análisis cuantitativo que nos permite introducir relaciones de causa-efecto de manera *a priori* es el modelo de regresión lineal múltiple¹², por lo cual lo pasaremos a explicar en detalle. Dicho modelo puede ser representado a partir de la siguiente ecuación:

$$Y_i = a + bX_i + cW_i + u_i$$

Donde Y es la variable "dependiente" a la cual se intenta explicar a través de las variables X y W , las cuales se denominan "independientes". Asimismo, b y c son los parámetros de la ecuación de regresión que capturan la "intensidad" con la que X y W influyen en Y , respectivamente (el cambio en Y por cada unidad de cambio en la variable independiente). El parámetro a , captura todos aquellos efectos que no son "observables", es decir, cuando X y W son cero respectivamente. La variable u es conocida como el error aleatorio o residual y captura el efecto de todos aquellos factores que no han sido considerados en el modelo o también los errores propios de la medición de las variables. Cabe recordar que al trabajar con datos estadísticos y con ideas preconcebidas que intentan acercarse a la verdadera relación entre las variables se está constantemente en riesgo de incurrir en errores de diversos tipos, como los errores de registro, de medición, de relación entre las variables¹³, entre otros. Finalmente, el subíndice i representa una unidad observacional, la cual puede representar personas, empresas, instituciones, países, etc.

Este modelo tiene los siguientes supuestos:

- i. El error residual de las mediciones tiene una media cero.
- ii. No hay correlación serial (no hay correlación entre los residuales de dos variables cualesquiera considerados en el modelo).
- iii. Homocedasticidad (que los residuales tienen una varianza constante).

¹²El método de regresión lineal y sus distintas variantes pueden abarcar libros enteros. Aquí solo se desarrollarán sus principales características a través de un ejemplo aplicado. Para un estudio de mayor profundidad se recomienda la revisión de la bibliografía citada al final del capítulo.

¹³Cómo su nombre lo dice, el modelo de regresión lineal asume una relación lineal entre las variables, lo cual no exime el hecho de que en la realidad existan diversos tipos de influencias de las variables independientes sobre la dependiente, como relaciones cuadráticas, polinómicas, exponenciales, etc.

- iv. Las variables no están correlacionadas con el residual.
- v. No multicolinealidad (No hay relación lineal de ninguna variable en función de las otras consideradas en el modelo).

El modelo es descriptivo cuando simplemente se estima la relación entre las variables y el resultado; para estimar causalidades debe ser aplicado con datos que provienen de un diseño experimental o cuasiexperimental, es decir con un grupo control.

El modelo de regresión lineal puede ser entendido de mejor manera a través de un ejemplo en el cual se desea hallar el efecto de una intervención específica.

Ejemplo. PROJoven es un programa de capacitación laboral juvenil cuyo objetivo es mejorar el acceso al mercado laboral formal de jóvenes de 16 a 24 años de edad y de escasos recursos económicos en situación de pobreza, con bajos niveles de escolaridad y que se encuentren desempleados o subempleados. A los jóvenes que son seleccionados a través de la aplicación de una ficha de evaluación socioeconómica, se les brinda cursos de capacitación técnica básica con una duración de alrededor de tres meses, además de suscribirse convenios para que realicen pasantías en empresas y puedan aplicar los conocimientos recién adquiridos.

Como parte del sistema de evaluación se debe indagar acerca del impacto que tiene el programa en el nivel de ingresos de los jóvenes capacitados. Para ello, se definió el siguiente modelo de regresión lineal.

$$y_i = a + b_0w_i + b_1d_i + u_i$$

Donde y es el nivel de ingreso laboral (variable independiente), w representa los años de experiencia laboral, d una variable "cualitativa" que toma el valor de 1 cuando el individuo i recibió el programa (grupo de intervención) y 0 para aquellos individuos que no lo recibieron (grupo de control). Además, u es el término de error aleatorio. Finalmente, el parámetro a explica, únicamente para este ejemplo, aquellas habilidades de los jóvenes que influyen en su salario que no pueden ser medibles, como su inteligencia, proactividad, liderazgo, entre otras.

Es importante subrayar que el modelo de regresión lineal propuesto líneas arriba es útil para la introducción de ideas preconcebidas acerca de la relación entre las

variables y que puede ayudar al análisis causa-efecto, debido a que en esta ocasión se asume que el ingreso de los jóvenes es determinado solo por sus años de experiencia laboral, la capacitación que brinda el programa y por otros factores que no pueden ser determinados con la información disponible. Puede existir otro tipo de especificaciones que incluyan variables como el nivel socioeconómicos de la familia, el ámbito geográfico o la edad del beneficiario, pero ello dependerá de la concepción y conocimiento que tenga el evaluador al momento de la estimación.

Finalmente, a partir del modelo propuesto, se realizó una estimación que produjo los siguientes resultados:

$$y_i = 1,2 + 0,8w_i + 2,4d_i + u_i$$

Estos resultados se pueden interpretar como la existencia de efectos positivos del programa sobre el salario de los jóvenes beneficiarios. Ello se debe a que el valor del parámetro correspondiente a la variable "d" es positivo. Además, se debe considerar la escala de los datos, que para este ejemplo es en "cientos de soles", lo que significa que el salario de aquellos jóvenes que participaron en el programa de capacitación laboral, comparado con aquellos que no participaron, pero que tienen características similares, es mayor en S/. 240,00 aproximadamente. De otro lado, es posible observar que los años de experiencia laboral también influyen de manera positiva en el salario, dado que por cada año de experiencia laboral, un joven recibe S/ 80,0 más en promedio.

4.1.2. Análisis cualitativo

En el proceso cuantitativo, el procedimiento de análisis es bastante estandarizado; en el proceso cualitativo no sucede lo mismo. En él el procesamiento y análisis pueden ocurrir prácticamente en paralelo; además, el análisis no es estándar, ya que cada estudio requiere de un esquema propio del análisis.

Para que se pueda identificar la diferencia entre las acciones involucradas en el procesamiento y análisis cualitativo cabe recordar que en principio, durante la recolección de datos, el proceso esencial consiste en que recibimos *datos no estructurados*, a los cuales durante el procesamiento nosotros les hemos damos cierta estructura a partir de las pautas que se discutieron en el capítulo anterior.

Finalmente, el análisis cualitativo implica una exposición narrativa donde se explican los hallazgos, aunque sin incluir pormenores de cada unidad observada.

En este apartado describiremos los enfoques que se emplean fundamentalmente en el análisis cualitativo general pero que deberán ser ajustados según su requerimiento.

a. ¿Cuáles son los enfoques cualitativos?

En el análisis de información cualitativa existen diversos enfoques que permiten interpretar y sistematizar la información recogida, entre ellos: el inductivo general, teoría fundamentada, análisis de discurso y fenomenología (Ver cuadro 20). Sin embargo, entre todos estos enfoques destaca el enfoque inductivo general, el cual utiliza una lectura detallada de los datos para derivar conceptos, temas o un modelo a través de interpretaciones realizadas a partir de los datos generados por un evaluador. Su propósito fundamental es permitir que los resultados emerjan de los temas más significativos sin imponer metodologías.

Cuadro 20. Comparación de enfoques de análisis cualitativo

	Enfoque inductivo general	Análisis de discurso	Fenomenología
Estrategia analítica y preguntas	¿Cuáles son los significados evidentes más importantes del texto que son relevantes para la evaluación o los objetivos de la evaluación?	Preocupación con la charla y textos como prácticas sociales y su organización argumentativa o retórica	Trata de descubrir el significado de la vida a través de la experiencia para poder transmitirlo en palabras
Análisis de resultados	Temas o categorías más relevantes para los objetivos identificados	Múltiples significados del lenguaje y de los textos identificados y descritos	Descripción de las experiencias de la vida
Presentación de evidencia	Descripción de varios temas relevantes	Descripción de múltiples significados en el texto	Historia corriente o narrativa acerca de la experiencia

Fuente: Thomas, D., (2006).

El enfoque inductivo, en general, proporciona un conjunto de procedimientos sistemáticos para el análisis de datos cualitativos que pueden producir resultados fiables y seguros. Aunque el enfoque inductivo general no es tan fuerte como algunas otras estrategias analíticas para la teoría o modelo de desarrollo, es sencillo y directo para derivar conclusiones en el contexto de las preguntas de evaluación. Los propósitos de este enfoque son los siguientes:

- ⇒ Resumir la variedad y la extensión de la información. Establecer vínculos claros entre los objetivos de la evaluación y los resultados resumidos para garantizar que estos puedan ser demostrados por otros.
- ⇒ Desarrollar un modelo o teoría acerca de la estructura subyacente de experiencias o procesos evidentes en los datos de texto.

b. ¿Qué debo considerar durante el análisis?

- ⇒ El análisis de datos es guiado por los objetivos de la evaluación, que identifican dominios y temas a ser investigados. El análisis se lleva a cabo a través de múltiples lecturas e interpretaciones de los datos en bruto. Aunque los resultados estén influidos por los objetivos de la evaluación o por preguntas ordenadas por el evaluador, los resultados surgen directamente del análisis de los datos en bruto y no de expectativas o modelos *a priori*.
- ⇒ El principal modo de análisis es el desarrollo de categorías a partir de los datos brutos en un modelo o marco. Este modelo contiene temas clave y procesos, identificados y construidos por el evaluador durante el proceso de codificación.
- ⇒ Los hallazgos son el resultado de múltiples interpretaciones realizadas a partir de los datos en bruto por los evaluadores que codifican los datos. Inevitablemente, los hallazgos son moldeados por las suposiciones y las experiencias de los evaluadores que realizaron el estudio y la realización del análisis de datos.
- ⇒ Diferentes evaluadores pueden producir resultados que no son idénticos y que tienen componentes que no se superponen.

- ⇒ La fiabilidad de los resultados derivados del análisis inductivo, por ejemplo, se puede evaluar utilizando técnicas similares a las que se utilizan con otros tipos de análisis cualitativo.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Metodología de la investigación. Capítulo 14. Hernández R. et al (2006).
- Analyzing Qualitative Data. Págs. 1-12. Taylor- Powell E. Y Renner M. (2003).

4.2. Presentación de datos

Una vez que se obtengan los resultados de los análisis estadísticos (tablas, gráficas, cuadros, etc.), se recomienda realizar las siguientes actividades:

- ⇒ Revisar cada resultado (análisis general, análisis específico, valores resultantes, tablas, diagramas, cuadros, gráficos).
- ⇒ Organizar los resultados (primero los descriptivos, por variable; luego los resultados relativos a la confiabilidad y la validez; posteriormente los inferenciales, que se pueden ordenar por hipótesis o de acuerdo con su desarrollo).
- ⇒ Cotejar diferentes resultados: su congruencia y en caso de inconsistencia lógica volverlos a revisar. Asimismo, se debe evitar la combinación de tablas, diagramas o gráficos que repitan datos. Cuando este es el caso, debemos elegir la tabla o elemento que ilustre o refleje mejor los resultados y sea la opción que presente mayor claridad. Una buena pregunta en este momento del proceso es ¿qué valores, tablas, diagramas, cuadros o gráficos son necesarios? ¿cuáles explican mejor los resultados?
- ⇒ Reducir los datos, es decir priorizar la información más valiosa (que es en gran parte resultado de la actividad anterior), sobre todo si se van a producir reportes ejecutivos y otros más extensos. Es decidir qué datos van a ser utilizados en el análisis libre de duplicaciones o redundancias, pues puede ocurrir que varias variables puedan estar midiendo lo mismo con muy ligeras variantes.

- ⇒ Copiar y/o formatear las tablas en el programa con el cual se elaborará el reporte de la investigación (procesador de textos o uno para presentaciones). Algunos programas como SPSS y Stata permiten que se transfieran resultados (tablas, por ejemplo) directamente a otro programa. Por ello resulta conveniente usar una versión del programa de análisis que esté en el mismo idioma del programa que se empleará para describir el reporte o elaborar la presentación.
- ⇒ Comentar o describir brevemente la esencia de los análisis, valores, tablas, diagramas, gráficos.
- ⇒ Volver a revisar los resultados.
- ⇒ Finalmente, elaborar el reporte de evaluación.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Metodología de la investigación. Capítulo 14. Hernández R. et al (2006).

4.3. Hallazgos con evidencias, triangulando fuentes y técnicas

En lo regular, en la medida que el presupuesto de la evaluación lo permita, es conveniente tener varias fuentes de información y métodos para recolectar los datos. Tanto en el proceso cuantitativo como en la indagación cualitativa poseemos una mayor riqueza y profundidad en los datos si estos provienen de diferentes actores del proceso, de distintas fuentes y al utilizar una mayor variedad de formas de recolección de los datos. Por ejemplo, si queremos evaluar la intervención pública de atención a mujeres con depresión postparto de comunidades indígenas y nuestro esquema de estudio incluye:

- Observación durante la inmersión en la comunidad (contexto).
- Entrevistas con mujeres que la experimentan.
- Entrevistas con sus familiares.

- Observación inmediatamente posterior al parto (durante la convalecencia) en hospitales rurales, en sus hogares (en varias comunidades indígenas los partos se llevan a cabo en la propia “casa-habitación” de la madre).

De esta manera, el sentido de entendimiento de la intervención será mayor si únicamente llevamos a cabo entrevistas.

Al hecho de utilizar diferentes fuentes y métodos de recolección, se le denomina triangulación de datos. En la evaluación es recurrente que durante el análisis de resultados los métodos cualitativos y cuantitativos implementados hayan sido diseñados de manera complementaria a fin de brindar diversa información en torno a una misma dimensión de la evaluación, por ejemplo, sobre la eficiencia en la gestión del programa/proyecto.

Finalmente, en el proceso cualitativo debe tomarse en cuenta que los datos recolectados habrán de interpretarse. De este modo se reflejará el hecho de que los seres humanos, en sus interacciones con el mundo que los rodea, esbozan sus recursos conceptuales y los utilizan para construir los significados de sus circunstancias y de su situación.

Hay que recordar que solo se pueden generalizar los resultados cuantitativos basados en encuestas por muestreo.

4.4. Evaluación de síntesis con interpretaciones

Una vez que se tiene los resultados de la evaluación y antes de elaborar la síntesis, es necesario revisar la verosimilitud de ellos. Es común encontrar que algunos resultados no responden a lo esperado tomando en cuenta la experiencia o la forma como fue ejecutado el PPP. Es posible que se hayan cometido errores en la fase de procesamiento o inclusive mucho antes. En tales casos es necesario tomar una serie de medidas.

Se debe revisar la Teoría de Cambio considerada en el proyecto pues algunas relaciones pueden haber estado mal formuladas; se debe revisar también el Marco Lógico, el cual puede haber tenido limitaciones o indicadores no definidos adecuadamente; o puede suceder que los supuestos hayan sido mal formulados. Otros aspectos que pueden explicar hallazgos inesperados es que los instrumentos pueden haber tenido problemas que recién han sido detectados en

el momento del análisis. También es aconsejable revisar el enfoque de la evaluación pues algunos métodos pueden haber sido escogidos inadecuadamente para la naturaleza del proyecto, como por ejemplo el modelo seguido para determinar la atribución del efecto al proyecto. Puede ocurrir también que en la operación de recolección de datos primarios se haya tenido problemas en la muestra, con desvíos en la representatividad por los casos de no respuesta y no se hayan incorporado los factores de ajuste necesarios. Es común también que no se haya podido recolectar toda la información deseada que evidencie el discurrir del PPP.

En este punto de la evaluación ya se cuenta con los resultados del estudio cuantitativo y cualitativo organizados (tablas, gráficos, cuadros, diagramas, etc.) e incluso descritos, pero el proceso aún no termina. Es necesario comunicar los resultados mediante un reporte de evaluación.

Lo primero entonces es definir el tipo de reporte que es necesario elaborar. Esto depende de varias precisiones: 1) las razones por las que surgió la evaluación, 2) los usuarios de la evaluación, 3) el contexto en el cual se habrá de presentar. Por lo tanto, es necesario que antes de comenzar a desarrollar el reporte, el evaluador reflexione respecto a las siguientes preguntas: ¿cuál fue el motivo o los motivos que originaron la evaluación?, ¿cuál es el contexto en el que habrán de presentarse los resultados?, ¿quiénes son los usuarios del resultado? y ¿cuáles son las características de tales usuarios? La manera en que se presenten los resultados dependerá de las respuestas a dichas preguntas.

El informe final de la evaluación es el principal producto y se elabora después de procesar los datos. Los principales objetivos del informe de evaluación final son presentar los resultados de evaluación y responder las preguntas de evaluación que se formularon inicialmente. Asimismo, en el caso de la evaluación de impacto, el informe también debe demostrar que la evaluación se basa en estimaciones válidas del contrafactual y que los impactos estimados pueden atribuirse al programa.

El informe final de la evaluación es un informe exhaustivo que resume todo el trabajo e incluye descripciones detalladas del análisis de datos, análisis de los resultados, cuadros y apéndices.

A continuación se presentan algunas recomendaciones para la redacción de evidencias y síntesis de información:

- Las secciones del reporte deben relacionarse entre sí por un “hilo conductor” (el último párrafo de una sección con el primero de la siguiente sección).
- En el caso de los resultados cualitativos:
 - i. Los análisis deben incluir fragmentos de contenido o testimonios (unidades de análisis) expresados por los participantes (citas textuales, en su lenguaje, aunque las palabras sean incorrectas desde el punto de vista gramatical o puedan ser consideradas “impropias” por algunas personas).
 - ii. En la interpretación de resultados y la discusión: se revisan los resultados más importantes y se incluyen los puntos de vista y las reflexiones de los participantes y del evaluador respecto al significado de los datos, los resultados y el estudio en general, además de evidenciar las limitaciones de la investigación y hacer sugerencias para futuras indagaciones.

Son importantes las descripciones profundas y completas (así como su significado) del contexto, ambiente o escenario; de los participantes; los eventos y las situaciones; las categorías, temas y patrones, y de su interrelación (hipótesis y teoría). También es necesario aclarar cuál fue la posición del evaluador ante los hechos (anotaciones personales). En este sentido debe demostrarse que el evaluador valoró sus propias tendencias, creencias y otras fuentes potenciales de sesgo; que de igual forma reflexionó sobre sus experiencias y su vinculación con los participantes.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Banco Mundial. La evaluación del impacto en la práctica. Capítulo 13.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 4

Ahora procede a desarrollar las actividades aplicativas de la
unidad 4

ACTIVIDAD 1

Realizando un ejercicio de categorización y presentación de la
información

ACTIVIDAD PRESENCIAL

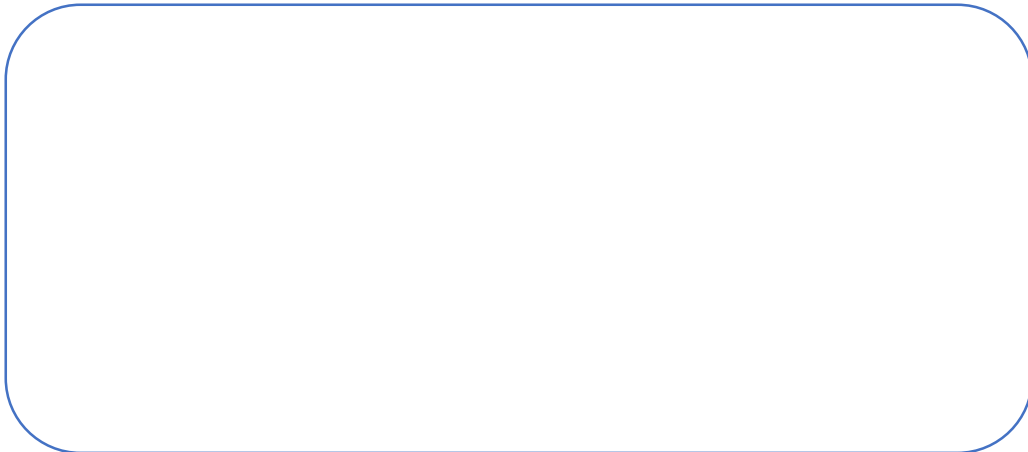
1. Asiste a la clase programada en el sílabo.
2. Participa en las siguientes actividades:
 - Categorización y presentación de la información.
 - Síntesis e interpretaciones.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 4

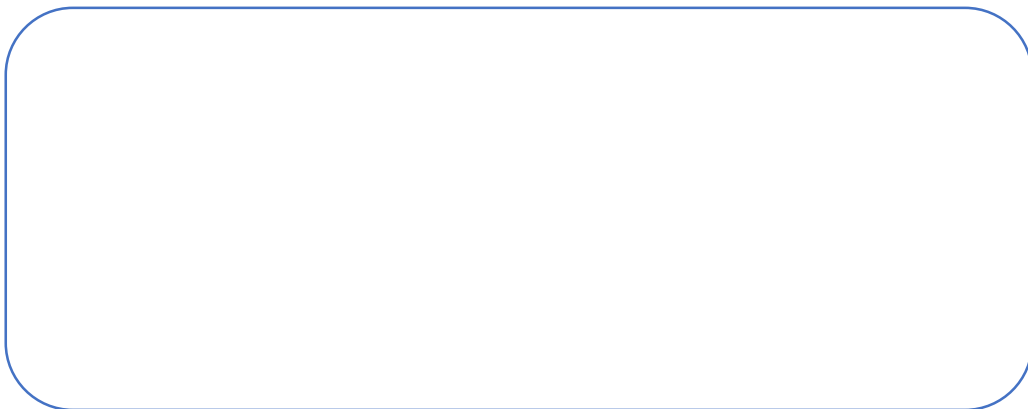
HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

1. ¿Cuáles son los Supuestos de partida del análisis paramétrico?



2. Indica las principales características del análisis cualitativo.





UNIDAD DE APRENDIZAJE 5: PREPARACIÓN Y REDACCIÓN DEL INFORME DE EVALUACIÓN

La importancia de un buen informe de evaluación

Un evaluador requiere tener grandes habilidades de comunicación escrita y oral para presentar informes que sean útiles; debe argumentar de manera lógica y contar una historia que atrape a quien lo lee o escucha. Extraer conclusiones y vincularlas a las acciones o decisiones que un gerente de programa puede implementar requiere no solo de una buena metodología sino del poder de persuasión en reuniones de trabajo, o presentaciones, o en la redacción del informe final.

En esta unidad daremos pautas para escribir un adecuado informe que contribuya a diseminar los resultados y recomendaciones de una evaluación entre aquellos que van a mejorar un programa o quieren aprender para hacer uno nuevo o de un financiador que espera saber los resultados para saber si vale la pena seguir desembolsando dinero o si valió la pena hacerlo porque los beneficiarios cambiaron para mejor.

5.1 Cómo preparar y presentar un informe de evaluación

Los informes de evaluación son el resultado final del proceso de la evaluación y son importantes porque se utilizan para tomar decisiones y mejorar los programas y proyectos. Un factor clave para usar los resultados de las evaluaciones es que los informes estén bien escritos, sean breves y comuniquen clara y rápidamente los hallazgos y conclusiones.

El cuadro 21 proporciona una orientación acerca del informe de evaluación, su estructura, contenido y los pasos que se deben de seguir para elaborar un informe. El informe de evaluación debe presentarse en un documento sólidamente elaborado, reflexivo y organizado, para mostrar objetivamente los resultados de un proyecto o programa. Los hallazgos, conclusiones y recomendaciones deben basarse en evidencia derivada de los mejores métodos disponibles, de las preguntas de evaluación y los recursos disponibles.

Los métodos de evaluación, las limitaciones y las fuentes de información deben ser documentados. Hay que proporcionar las herramientas de recolección de datos, la sintaxis de construcción de indicadores y los Términos de Referencia de la evaluación como anexos del informe principal. Las bases de datos deben

entregarse como archivos electrónicos adicionales al informe y deben ser lo suficientemente detalladas como para que cualquier otro evaluador pueda replicar los resultados.

Al escribir un informe, el equipo de evaluación debe recordar siempre cuál es su audiencia primaria: los gerentes de proyectos y programas, las autoridades y los interesados directos. El estilo de escritura debe ser fácil de entender, conciso, asegurándose de abordar las preguntas de la evaluación y cuestiones abordadas de forma precisa, con hallazgos fundamentados en datos, conclusiones sustentadas y recomendaciones prácticas.

Cuadro 21. Esquema y contenidos del informe de evaluación

Informe debería ser...	Reflexivo, bien organizado y con una evaluación objetiva de lo que funcionó, lo que no se hizo y por qué.
Resumen Ejecutivo	Un resumen ejecutivo proporciona una breve descripción general del propósito de la evaluación, los antecedentes del proyecto, las preguntas de evaluación, los métodos, los resultados y las conclusiones.
Objetivos y Preguntas de evaluación	Abordar todas las preguntas de evaluación especificadas en los Términos de Referencia.
Antecedentes acerca del programa o proyecto	Elaborar un resumen del programa que se está evaluando, incluyendo el problema que le dio origen y los cambios que ocurrieron, por ejemplo en su diseño o componentes, desde que empezó. Incluir una descripción de su población beneficiaria, el área geográfica, su lógica causal (incluir si existe el Marco Lógico o Marco de Resultados).
Métodos	<ul style="list-style-type: none"> • Explicar la metodología de evaluación en detalle: indicadores, fuentes de los datos, técnicas de recopilación usadas, muestra, modelos y controles de calidad. • Informar y discutir las limitaciones de la evaluación, especialmente aquellas asociadas a la metodología empleada (por ejemplo, el sesgo de selección, el sesgo de memoria, las diferencias no observables entre los grupos de intervenidos y de comparación, etc.). • Se deben incluir de manera explícita los aspectos relacionados al consentimiento informado u otros aspectos que tienen que ver con la ética del proceso de evaluación. <p>Nota: el resumen de la metodología se puede incluir en el cuerpo del informe, con la descripción completa como anexo.</p>

<p>Hallazgos o resultados</p>	<ul style="list-style-type: none"> • Los hallazgos deben responder a las preguntas de evaluación y deben ser presentados de manera concisa y fundamentada con datos apoyados por gráficos y tablas que faciliten la lectura. • Presentar los hallazgos como hechos analizados relacionando varias variables, con evidencias y datos apoyados por fuentes cuantitativas o cualitativas y no con anécdotas, rumores u opiniones. • Incluir los hallazgos de manera diferenciada en relación a los impactos sobre los asuntos que se refieren a poblaciones de interés de la evaluación. • Se deben priorizar los hallazgos a ser presentados en función de las consecuencias que tienen para los beneficiarios, la gestión del programa o los financiadores.
<p>Conclusiones</p>	<ul style="list-style-type: none"> • Las conclusiones sintetizan e interpretan los resultados y sirven para tomar decisiones, los cuales se fundamentan en hallazgos específicos.
<p>Recomendaciones</p>	<ul style="list-style-type: none"> • Las recomendaciones, si proceden, son acciones específicas que el equipo de evaluación propone para la gestión de los programas. Fundamentar las recomendaciones con los hallazgos y conclusiones. • Proporcionar recomendaciones orientadas a la acción que sean prácticas y específicas, y que definan quién es el responsable de la acción. Las recomendaciones deben tener la posibilidad de ser implementadas por los tomadores de decisiones.
<p>Lecciones aprendidas</p>	<p>No todas las evaluaciones generan lecciones.</p> <ul style="list-style-type: none"> • Las lecciones elaboradas representan las contribuciones al conocimiento general. • Se fundamentan en los resultados y las conclusiones de la evaluación y no son una repetición de conocimiento común. • Se presentan como lecciones que se pueden aplicar a diferentes contextos y/o de diferentes sectores, y considera las limitaciones probatorias como: la generalización de las observaciones puntuales.
<p>Los Anexos</p>	<p>Incluir los siguientes datos en forma de anexos, como mínimo:</p> <ul style="list-style-type: none"> • Términos de Referencia. • Descripción completa de los métodos de evaluación empleados. • Todas las herramientas de evaluación (cuestionarios, listas de verificación, guías de discusión, encuestas, etc.).

	<ul style="list-style-type: none"> Una lista de fuentes de información (informantes clave, los documentos examinados, otras fuentes de información). Solo si es el caso, incluir un anexo con la(s) declaración(es) de las diferencias no resueltas de opinión por parte de alguno de los interesados en la evaluación y/o miembros del equipo de evaluación. <p>Declaraciones de ausencia de conflicto de interés de parte de los integrantes del equipo de evaluación.</p>
Control de Calidad	Asegurar la calidad de los informes mediante la revisión técnica de pares.
Transparencia	Compartir las conclusiones de los informes de evaluación en la medida de lo posible con compromisos de divulgación plena y activa.
Uso	Incorporar los resultados de los informes de evaluación en la toma de decisiones acerca de estrategias, prioridades de los programas y el diseño de nuevos proyectos.

Elaborado por Alejandro Bardales, basado en USAID Evaluation Policy and ADS 203 y UNIFEM Guidance: Quality Criteria for Evaluation Reports (2009).

El informe debe presentar todas las referencias bibliográficas (asentamiento bibliográfico) que sustentan el estudio siguiendo un patrón o guía único. El procesador de textos Word tiene la opción Referencias/citas bibliográficas /insertar cita/agregar nueva fuente, que presenta la pantalla mostrada en la figura 13, en la cual se llena toda la información del documento.

Figura 13. Formato para el llenado de las referencias bibliográficas

Luego, en la sección "bibliografía", aparece el libro, revista, artículo, etc. según el estilo escogido. El software presenta doce estilos diferentes, dos de los cuales se ponen como ejemplos en las figuras 14 y 15.

Figura 14. Bibliografía con el estilo APA

Bustelo Ruesta, M. (2001). *La evaluación de las políticas públicas de igualdad de género de los gobiernos Central y Autonómicos en España: 1995 – 1999*. Madrid: Universidad Complutense de Madrid Facultad de Ciencias Políticas y Sociología.

Cohen, E., & Martínez, R. (s.f.). *Manual de formulación, evaluación y monitoreo de proyectos sociales*. División de Desarrollo Social CEPAL.

Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género. (2011). *Integración de los derechos humanos y la igualdad de género en la evaluación – hacia una guía del UNEG*. Grupo de Evaluación de las Naciones Unidas (UNEG).

Figura 15. Bibliografía con el estilo MLA

Bustelo Ruesta, María. *La evaluación de las políticas públicas de igualdad de género de los gobiernos Central y Autonómicos en España: 1995 – 1999*. Madrid: Universidad Complutense de Madrid Facultad de Ciencias Políticas y Sociología, 2001.

Cohen, Ernesto y Rodrigo Martínez. *Manual de formulación, evaluación y monitoreo de proyectos sociales*. División de Desarrollo Social CEPAL, s.f.

Grupo de Trabajo del UNEG sobre Derechos Humanos e Igualdad de Género. *Integración de los derechos humanos y la igualdad de género en la evaluación – hacia una guía del UNEG*. Grupo de Evaluación de las Naciones Unidas (UNEG), 2011.

Para ampliar el tema tratado, revisa las siguientes lecturas:

- Escorcía, O. Manual para la investigación. Guía para la formulación, desarrollo y divulgación de proyectos. Bogotá. Universidad Nacional de Colombia. 2010. Págs. 52-63. Disponible en: <http://www.bdigital.unal.edu.co/7551/1/manualparalainvestigaci%C3%B3n.pdf>
- Faúndez, A. Guía de Evaluación de Programas y Proyectos con Perspectiva de Género, Derechos Humanos e Interculturalidad. México. ONU Mujeres, 2014. Págs. 20-34.
- Comisión Europea, Dirección General de Relaciones Exteriores. Metodología de Evaluación de la Ayuda Exterior de la Comisión Europea. Francia. Comisión Europea, 2006. Volumen I Bases Metodológicas de Evaluación. Páginas 88-91.
- Jaramillo, Baanante & Sanz. Informe Final de Evaluación del Programa "Construyendo Perú". Ministerio de Economía y Finanzas. Lima, 2009.

5.2 Cómo identificar hallazgos en la evaluación

El corazón del informe de evaluación es mostrar lo que se aprendió a partir de la recolección de información, para esto tener presente lo siguiente:

- ⇒ Se debe estructurar la información de manera que se pueda responder cada una de las preguntas formuladas en el diseño de la evaluación.
- ⇒ La organización del cuerpo del informe debe orientarse siguiendo las grandes preguntas o hipótesis o asuntos que se han formulado para la evaluación.
- ⇒ El análisis de los datos se debe efectuar en relación a las implicaciones que tienen a las preguntas, hipótesis o asuntos de la evaluación. Los datos no hablan por sí solos, hay que usarlos orientándolos por un argumento. Cuente una historia ayudado por los datos. En especial si usa información cualitativa, la estructura y la fuerza del argumento deben organizarse en función de los objetivos o hipótesis emergentes del estudio. Evite las digresiones.
- ⇒ Usar tablas, mapas conceptuales y gráficos para resumir la información a fin de facilitar la lectura y comprensión de los hallazgos.
- ⇒ Evitar poner en el informe todo lo que encontró en el recojo de datos, no hay necesidad de "contarlo todo". Solo incluye lo referente a las preguntas de la evaluación.

- ⇒ Evita centrar el análisis en las cuestiones estadísticas como pruebas, regresiones, correlaciones. Recuerde que solo debe usarlos estrictamente en la medida que están relacionados a las hipótesis o preguntas de la evaluación. Estos aspectos bien explicados deben ir en un anexo especial.
- ⇒ Si trabajas con hipótesis, indica claramente cuáles de ellas no se rechazaron y cuáles sí.
- ⇒ Los datos detallados, desagregados y los procedimientos para construir índices deben ir a los anexos del informe.
- ⇒ Al priorizar los hallazgos importantes, considera cuidadosamente las consecuencias de esa selección para los beneficiarios del programa, los financiadores o la gerencia del programa. Este es un asunto más importante que enfocarse en el análisis de datos y modelos, aunque esta responsabilidad nos obliga a ser muy cuidadosos con esos análisis.

Recuerda que aun cuando puedas seleccionar los hallazgos relevantes, tendrás que tener la flexibilidad suficiente para aceptar matices o sugerencias que podrán provenir de alguno de los interesados en la evaluación.

5.3 Cómo emitir juicios de valor o mérito en una evaluación

El análisis de los resultados de una evaluación depende del enfoque elegido en el diseño. Una de las prioridades del enfoque positivista consiste en valorar o juzgar el programa. Por el contrario, las prioridades de los enfoques fenomenológico y realista hacen que valorar o juzgar un programa no sea de interés.

En el enfoque positivista el proceso para juzgar o valorar un programa consta de los siguientes pasos. Cumplirlos permite que podamos emitir juicios de valor en el informe de evaluación:

5.3.1 Seleccionar un conjunto de criterios (eficacia, eficiencia, etc.)

Recuerda que estos criterios se refieren a alguna característica predefinida del programa, por ejemplo, la eficacia se refiere al logro de objetivos, la eficiencia al uso de recursos, etc. Evidentemente estos criterios son solo un subconjunto de todos los criterios que podrían ser elegibles para juzgar o valorar un programa. Tanto es así que eventualmente cada interesado en la evaluación de un programa podría desear evaluarlo con un criterio diferente.

Una restricción a la multiplicidad e indeterminación de criterios es que al menos ellos deben ser parte de las características del programa. En un ejemplo extremo,

si un programa no ha sido diseñado para mejorar la equidad de género, entonces no tendría sentido juzgarlo sobre la base de ese criterio.

Por consiguiente se necesita que el evaluador *negocie con los interesados los criterios* sobre los cuales el programa debe ser valorado. Existen varios métodos que pueden ayudar en la negociación entre interesados para *seleccionar criterios* para evaluar un programa. Entre ellos están la argumentación panóptica y la evaluación multicriterio.

Esta fase se debe ejecutar en la etapa del *diseño del programa* y podría tener algunos ajustes en el diseño de la evaluación. También debemos elegir los indicadores que sirven para medir cada criterio. Los resultados de esta primera fase deben reflejarse en los objetivos y en las preguntas que la evaluación debe responder.

5.3.2 Establecimiento de estándares para cada uno de los criterios seleccionados

A estos estándares se les asigna una escala, por ejemplo, excelente, bueno, regular, malo. Usualmente los estándares también se establecen al momento de fijar los indicadores del programa. Cada indicador debe tener un estándar contra el cual medir el desempeño. El rol del evaluador consiste en facilitar y generar consensos acerca de los indicadores y sus estándares.

Se recuerda que *los criterios y estándares deben estar justificados* con alguna argumentación; para ello las técnicas de la evaluación multicriterio pueden apoyar estas justificaciones. Los indicadores y estándares se identifican al momento de diseñar un programa, en la etapa de preinversión.

En el siguiente ejemplo (cuadro 22) se muestran los estándares para el criterio de evaluación "cumplimiento presupuestal", que sirve al organismo rector de evaluación de México, el CONEVAL para juzgar los programas estatales.

Cuadro 22. Ejemplo de estándares para el criterio de evaluación "cumplimiento presupuestal"

Nombre de la variable: cumplimiento presupuestal 2012	
Definición:	La variable cumplimiento presupuestal reporta el avance presupuestal en 2012 de los programas; es decir, mide qué porcentaje del presupuesto autorizado de los programas fue ejercido en el periodo de análisis.
Criterios de valoración:	Porcentaje del presupuesto autorizado del programa ejercido durante el periodo de análisis.
Categorías:	
Calificación	Valoración
95-100	Cumplimiento del ejercicio presupuestal entre 95-100%
85-94	Cumplimiento del ejercicio presupuestal entre 85-94%
50-84	Cumplimiento del ejercicio presupuestal entre 50-84%
0-49	Cumplimiento del ejercicio presupuestal entre 0-49%
NA	Información no disponible
Interpretación de resultados	
El porcentaje de presupuesto autorizado ejercido se convierte en la calificación del programa; es decir, si un programa tiene un porcentaje de presupuesto ejercido contra el autorizado de 90 por ciento, se le asigna una calificación de 90 puntos.	

Fuente: <http://www.coneval.gob.mx/Paginas/principal.aspx>

Cuadro 23. Ejemplo de medición de dos criterios que se usaron para evaluar los programas sociales de México por parte de CONEVAL

DATOS DEL PROGRAMA				Cumplimiento presupuestal 2012	Resultados de la revisión de la Matriz de Indicadores para Resultados (MIR) 2012	
Dependencia	Modalidad / Clave	Nombre del programa	Descripción del programa		Presupuesto ejercido/ Presupuesto modificado	Valoración del diseño y calidad de la MIR 2012 (cumplimiento de los elementos, lógica interna, diseño de los indicadores)
Agroasemex	S001	Programa de Subsídios a la Prima del Seguro Agropecuario	El objetivo del programa consiste en ampliar la cobertura del seguro para agricultores y ganaderos subsidiando el costo de las primas que pagan los	100	Adecuado	4
Agroasemex	S172	Programa de Apoyo a los Fondos de Aseguramiento Agropecuario	El objetivo del programa es ampliar la capacidad de los Fondos y Organismos Integradores para cumplir con sus obligaciones de acuerdo con la Ley de Fondos mediante las	100	Destacado	5
Agroasemex	S199	Programa de Seguro para Contingencias Climatológicas	El programa tiene como propósito potenciar la capacidad financiera del Gobierno Federal mediante el pago de la prima para la contratación	100	Adecuado	4
CDI	S178	Programas Albergues Escolares Indígenas (PAEI)	El Programa Albergues Escolares Indígenas (PAEI) otorga hospedaje, alimentación y apoyo educativo en albergues escolares indígenas a	100	Adecuado	4
CDI	S179	Programa de Infraestructura Básica para la Atención de los Pueblos Indígenas (PIBAI)	El PIBAI tiene como objetivo general contribuir a que los habitantes de las localidades indígenas elegibles superen el aislamiento y dispongan	100	Destacado	5
CDI	S180	Programa Fondos Regionales Indígenas (PFR)	El Programa busca incrementar los ingresos de la Población indígena asociada a Fondos Regionales, que son agrupaciones de	100	Moderado	3
CDI	S181	Programa Organización Productiva para Mujeres Indígenas (POPMI)	El Programa de Organización Productiva para Mujeres Indígenas proporciona apoyos económicos a grupos de mujeres indígenas	100	Destacado	5
CDI	S182	Programa Promoción de Convenios en Materia de Justicia (PPCMJ)	El Programa de Promoción de Convenios en Materia de Justicia proporciona apoyos económicos a organizaciones sociales y/o núcleos agrarios, y a	100	Destacado	5
CDI	S183	Programa de Fomento y Desarrollo de las Culturas Indígenas (PFDCI)	Contribuir e impulsar los procesos culturales que incidan en la creatividad, fortalecimiento y rescate del patrimonio cultural de los pueblos y comunidades	100	Adecuado	4

Fuente: <http://www.coneval.gob.mx/Paginas/principal.aspx>

5.3.3 Medición del desempeño del programa

A través de alguna técnica de recolección de datos se mide el desempeño del programa en relación a los estándares mediante la asignación de un puntaje. Este puntaje mide la distancia entre el desempeño del programa y el estándar. En el ejemplo se muestra la medición de dos criterios que se usaron para evaluar los programas sociales de México por parte de CONEVAL (cuadro 23).

5.3.4 Síntesis de hallazgos

Se sintetizan los hallazgos en un juicio de valor final. La emisión de un juicio de valor se efectúa en el informe final de evaluación y requiere, de acuerdo a Hurteau, et al (2009), sujetarse a ciertas condiciones:

- Debe emitirse estrictamente en relación a la evidencia o hallazgos de la evaluación.
- Los juicios deben ser legítimos, es decir: i) deben referirse a los objetivos y preguntas de la evaluación, ii) deben estar fundamentados con algún criterio, por ejemplo, la eficacia, eficiencia, participación, etc. y iii) deben apoyarse en algún estándar.
- Los juicios deben estar justificados, lo cual incluye: i) justificación de los criterios, ii) justificación de los estándares y iii) documentación del procedimiento usado para sintetizar las informaciones en un juicio.

Nótese que para cumplir las dos últimas condiciones (juicios legítimos y justificados) estas deberán implementarse en la etapa del diseño del programa o en el diseño de la evaluación.

5.4 Cómo informar la atribución de causalidades

La redacción del informe acerca de la atribución de causalidades de una intervención debe centrarse en responder la pregunta básica de una evaluación orientada a estimar los efectos de un programa.

¿Cuál es el efecto del programa?

La evaluación causal también podría responder a una pregunta diferente

¿Cuál es la causa de que un resultado sea de cierta naturaleza?

Los métodos para averiguar causalidad en este manual se orientan a responder la primera pregunta. Como se ha dicho, son los métodos experimentales y en menor grado los cuasiexperimentales los más consistentes para responder a la pregunta clave que se establece para establecer causalidad: ¿qué hubiera pasado si el programa no hubiera existido? Al impacto logrado por el proyecto se le conoce también como su adicionalidad o su valor agregado, es decir, lo que ha logrado incorporar el PPP y que antes de la intervención no existía.

El enfoque en la redacción del informe es mostrar si existió algún cambio en los indicadores de interés y de existir dar evidencia que los cambios son atribuibles al programa. Se hará explícito el análisis y discusión de los siguientes asuntos:

- ⇒ Que la asignación entre controles y grupo de intervención fue aleatoria (solo para los diseños experimentales).
- ⇒ Que los grupos de control e intervenido son equiparables con los datos de la línea de base.
- ⇒ Que los tamaños de muestra tienen suficiente poder estadístico.
- ⇒ Señalar en el análisis, si usa algún modelo estadístico, que se han cumplido los Supuestos de ese modelo. Por ejemplo, si usa análisis de regresión.
- ⇒ Discutir si podrían existir otras razones diferentes al tratamiento como hipótesis razonables del cambio en los indicadores de interés y cómo el diseño o el análisis pudo controlar esas hipótesis.
- ⇒ Informar los hechos relevantes para el análisis que ocurrieron entre la línea de base y la evaluación final, por ejemplo si se pudieron mantener los controles fuera de toda influencia del programa o de intervenciones

similares, o si se pudieron mantener todas las unidades intervenidas a lo largo de la vida del programa.

- ⇒ Discutir la robustez de las pruebas usadas en los análisis estadísticos.
- ⇒ Si dispones de información cualitativa o proveniente de fuentes secundarias que sirve para mejorar las explicaciones de los efectos de programa, úsala.
- ⇒ Somete el informe a la revisión de expertos y consulta los análisis que has efectuado y las conclusiones a las que se llega.
- ⇒ Usa gráficas, diagramas para explicar los resultados de una manera amigable, recuerda que el informe está hecho para gerentes de programas, interesados que no son expertos en el análisis estadístico o de causalidad.

5.5 Cómo elaborar conclusiones

Las conclusiones son la síntesis de los hallazgos en relación a las preguntas de la evaluación. Las conclusiones sintetizan las respuestas a las preguntas de la evaluación y se sustentan en los datos y en juicios de valor. No es posible hacer conclusiones que no tengan sustento en la información analizada, así como tampoco es posible concluir sin hacer algún juicio de valor (revise el capítulo 5.3 de cómo hacer juicios de valor).

Para redactar las conclusiones, se recomienda tomar en cuenta los siguientes asuntos:

- ⇒ Mencionar el programa evaluado y los objetivos generales de evaluación que se plantearon.
- ⇒ Explicitar resumidamente cómo se respondieron las preguntas de la evaluación y si se cumplieron o no los objetivos.
- ⇒ Las conclusiones deben obtenerse a partir de algo más que de los simples datos registrados. De hecho, unos datos o resultados pueden tener un sentido u otro y, por tanto, pueden llevarnos a unas conclusiones u otras, dependiendo del enfoque que justifica nuestra evaluación, de la metodología seguida, de los objetivos propuestos, etc.
- ⇒ Las conclusiones no constituyen una repetición de los resultados, sino una síntesis de los resultados más importantes en relación a las preguntas de la evaluación.

- ⇒ No debe confundirse la discusión de resultados con la obtención de conclusiones, las cuales dependen tanto de los resultados y de su análisis, como de las preguntas de la evaluación y juicios de valor que puedas hacer.
- ⇒ Las conclusiones deben ser coherentes con la información analizada.
- ⇒ Las conclusiones deben derivarse de los resultados y de la discusión realizada en los apartados correspondientes.
- ⇒ Las conclusiones deben hacerse con relación a cada pregunta de evaluación. Si es necesario, haga mención a cualquier limitación que relativice los hallazgos.
- ⇒ Hay que mencionar las implicaciones que tienen los hallazgos para el programa, interesados, en especial para la población beneficiaria
- ⇒ Se resumen las explicaciones de resultados inesperados o las hipótesis que no se prueban.
- ⇒ Se deben presentar de manera interesante.

Nota: Una conclusión NO es un resumen como a veces se ha sido definida.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Guía Práctica para Elaborar Conclusiones. Elaboración propia de Grade (octubre 2015).
- Ejemplo Resumen de Informe de Evaluación y Diseño de Conclusiones. Comité de Desarrollo y Propiedad Intelectual (CDIP), noviembre 2012.

5.6 Cómo formular recomendaciones

5.6.1 Cuándo se debe formular recomendaciones

Se debe formular recomendaciones para los interesados en una evaluación, especialmente, si el evaluador está trabajando como parte de una comisión especial creada para estudiar un problema público y tiene que hacer recomendaciones para abordar ese problema. Por ejemplo, los estudios hechos para los ministerios o el congreso encomendados para investigar los problemas

del abuso de los niños, la contaminación del agua potable, el acceso a la atención médica, el progreso en mejorar la educación primaria, etc.

Cuando las recomendaciones no son obligatorias, el evaluador podría también considerar la posibilidad de ofrecer sugerencias para mejorar los proyectos o programas en lugar de recomendar. En la mayoría de los casos no hay diferencia sustancial entre las sugerencias y las recomendaciones. La diferencia que existe entre las recomendaciones y sugerencias es que las primeras son de más autoridad u obligación; por ejemplo, un informe oficial asignado para hacer frente a un serio problema público.

En el caso de comentarios que se refieran al cumplimiento de obligaciones, el vínculo entre las conclusiones y recomendaciones debe ser claro y directo. Es necesario y posible porque el cumplimiento de las conclusiones está sujeto a normas establecidas.

5.6.2 Cómo son las recomendaciones efectivas

- Las recomendaciones deben estar alineadas con los objetivos y las preguntas de la evaluación.
- Las recomendaciones deben negociarse entre la entidad que encarga la evaluación y el evaluador. Estas recomendaciones pueden estar o no destinadas a ser incluidas en la publicación del informe de evaluación.

En el caso de que el evaluador actúe de forma independiente o en representación de un grupo independiente como grupos de vigilancia ciudadana, no hay posibilidad de negociar con los funcionarios del programa evaluado; sin embargo, las recomendaciones serán más útiles solo si son realizadas después de las consultas con los funcionarios del programa y las partes interesadas.

- El evaluador, al proponer las recomendaciones, siempre tiene que pensar que tendrá que dialogar con un público más amplio que solo los funcionarios que administran el programa evaluado. Hay que recordar que los funcionarios que solicitan o reciben recomendaciones rara vez pueden actuar por sí solos, más bien necesitan que se involucren muchas más personas y organizaciones como los funcionarios de alto nivel, los órganos legislativos o grupos interesados, los cuales deben adoptar las recomendaciones antes de que se puedan implementar.

- El evaluador, al brindar recomendaciones o sugerencias, debe ofrecer varias opciones (salvo en el caso de las revisiones de conformidad), pues no hay una solución única a un problema, sino una serie de posibilidades, cada una con sus consiguientes ventajas y desventajas. Además, las decisiones rara vez son realizadas por una persona. Por lo general una combinación de decisiones de varias instituciones y niveles de funcionarios deben estar de acuerdo en una solución antes de que sea aprobada.
- Las recomendaciones deben derivar de las conclusiones y orientadas al avance de las metas, en caso las hubiera. En algunos casos, las recomendaciones derivan directamente de las conclusiones, por ejemplo, cuando se constatan incumplimientos de procedimientos (en las revisiones de conformidad). En la mayoría de casos, cuando se debe analizar eficiencia, eficacia, rentabilidad o impacto, las recomendaciones derivan indirectamente de las conclusiones, aquí la conexión entre las conclusiones y las recomendaciones es menos directa y hay mucho más espacio para discutir sobre cuán eficiente puede ser una operación. Por lo tanto, existen muchas opciones, cada una con sus ventajas y desventajas.
- Los evaluadores tienen que considerar que las recomendaciones deben ayudar a los clientes de los evaluadores a encontrar soluciones efectivas, las cuales deben estar motivadas, ser prácticas y relacionarse con los resultados de la evaluación.
- Algunos de los errores más comunes que se cometen en las recomendaciones son:

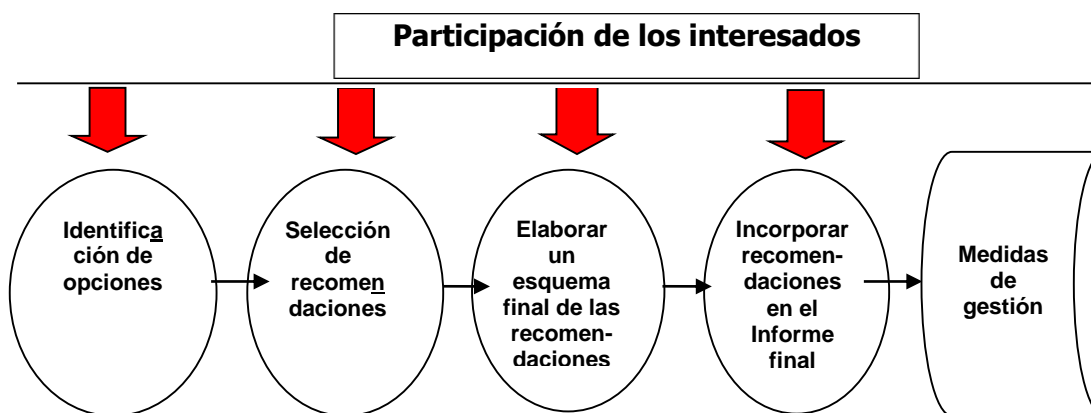
IMPORTANTE**Errores comunes al formular recomendaciones**

- Recomendar acciones que están fuera de la autoridad de la entidad evaluada.
- Recomendar acciones que son demasiado caras para aplicar (en el sentido de que su costo sería superior al presupuesto autorizado de la entidad).
- Hacer una sola recomendación cuando hay muchas más opciones disponibles.
- Recomendaciones generales, no factibles.

5.6.3 Cómo hacer recomendaciones

Hay que tener en cuenta que la participación de los interesados en las formulaciones de las recomendaciones es solo una etapa dentro de un proceso que va desde el diseño de una evaluación hasta las medidas de gestión o los acuerdos para implementarlas. Ello permite no solo conseguir ideas productivas sino también la aceptación de soluciones y alternativas útiles para el programa.

En la figura 16 se muestra el proceso de formulación de recomendaciones. El proceso se muestra desde la perspectiva del programa y no desde el equipo de evaluación. Es probable que el involucramiento del equipo de evaluación en este proceso vaya hasta la etapa donde se desarrolla el informe final de evaluación. Las medidas de gestión son directamente una responsabilidad del programa.

Figura 16. Proceso para formular recomendaciones

En cada etapa se incorporan las ideas de los interesados en la evaluación. El informe final es solo un paso en un proceso en el que se llegan a adoptar medidas para mejorar el programa antes y después de que el informe esté concluido.

a. Identificar las opciones

Hacer una lluvia de ideas con participación de involucrados clave

La mejor estrategia para resolver los problemas es comenzar por la identificación de un gran número de opciones. Este objetivo se puede alcanzar a través de una adecuada técnica y la participación de un amplio grupo de personas razonables y personas bien informadas.

Hay varias técnicas que se pueden utilizar en la sesión de lluvia de ideas para llegar a una amplia lista de recomendaciones. Tenga presente que fuera de una revisión de cumplimiento, el objetivo es encontrar soluciones que eliminen las causas o mitiguen los efectos negativos en la práctica y no creen nuevos problemas aún más graves en comparación con los que están siendo resueltos.

Las primeras ideas, por lo general, son realizadas por el equipo de evaluación, porque están familiarizados con el programa y porque se espera que tengan buen juicio, experiencia general y neutralidad. Sin embargo, es conveniente invitar a otros involucrados a la reflexión siempre que estén relacionados con el programa. Se sugiere que participen en la formulación de recomendaciones tanto los que están familiarizados como los que no conocen tanto el programa.

Para la lluvia de ideas se debe involucrar a personas que tengan las siguientes características.

- Personas con distintas perspectivas.
- Ser muy abierto a nuevas ideas.
- Ser respetuoso con los demás.

b. Selección de recomendaciones

Considera los problemas de financiación

Los problemas de financiación son uno de los aspectos más frustrantes del trabajo de los evaluadores porque se encuentran con muchos problemas sociales que no pueden resolver. Una razón para esta incapacidad es la falta de fondos disponibles para hacer frente a estos problemas. Sin embargo, hay formas de abordar el problema de la financiación. La mejor táctica es sugerir formas de ahorrar y mejorar la eficacia del programa.

Tácticas de financiación

- Sacar provecho de las opciones que no requieren grandes gastos, como los proyectos de investigación y desarrollo.
- Proporcionar autofinanciación mediante el cobro de tarifas a los usuarios.
- Requerir los fondos de contrapartida de los beneficiarios.
- Sugerir ahorros compensatorios, es decir, hacer sugerencias específicas para reducir los costos del programa.
- Utilizar fondos de incentivos, es decir, sugerir la asignación de una parte de los fondos del programa, que se va a utilizar como un fondo de incentivos, para que se distribuya entre los beneficiarios de acuerdo con el esfuerzo que hacen, la eficacia con que realizan sus programas o el éxito que tienen.

Empieza por las conclusiones

Los consumidores de los informes de evaluación esperan que las recomendaciones o sugerencias se deriven de las conclusiones. Por lo tanto, las soluciones han de abordar las causas y efectos del problema que se describe en el informe.

Evalúa las ventajas y desventajas de las opciones disponibles

Cuando se inicia el intercambio de ideas, es necesario analizar las soluciones para garantizar que se elimine la causa de los problemas o por lo menos mitigar sus efectos factibles y no crear problemas adicionales. También se podrían revelar si hay ventajas adicionales que no se han identificado. Este análisis descarta las soluciones poco prácticas y quizás proporcione una base para la clasificación o priorización para las propuestas.

Basado en: Wholey, J., Hatry, H. P. & Newcomer, K. E. Handbook of Practical Program Evaluation. Tercera edición (2010). p. 581-593.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Guía Práctica para Elaborar Recomendaciones. Elaboración propia de Grade (Octubre 2015).
- Cómo hacer Recomendaciones Efectivas bajo la Metodología Double-Smart. APT (Association for the Prevention of Torture). Cuadernillo N°1. (Noviembre 2008).

5.7 Cómo elaborar lecciones aprendidas

Para tratar el tema, revisa la siguiente lectura:

- Banco Interamericano de Desarrollo. Lecciones Aprendidas y Buenas Práctica. Sector de Conocimiento y Aprendizaje. Octubre 2008. p. 1-4.
- Red SIDA Perú. Marco conceptual sobre prácticas y lecciones aprendidas en VIH. Centro Virtual de Coordinación del Conocimiento en VIH. 2007. p. 23-30.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 5

Ahora procede a desarrollar las actividades aplicativas de la unidad 5



ACTIVIDADES

Analizando juicios de valor de evaluaciones de desempeño

LAS ACTIVIDADES DE ESTA UNIDAD SERÁN REALIZADAS DE MANERA CONJUNTA CON LAS ACTIVIDADES DE LA UNIDAD 6

ACTIVIDAD 3

Analizando y formulando conclusiones, recomendaciones y lecciones aprendidas

ACTIVIDAD PRESENCIAL

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 5

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde lo siguiente:

Responde si son verdaderas o falsas las siguientes afirmaciones:

Afirmación	Verdadera o falsa
1. Las conclusiones derivan directamente de los datos.	
2. Las conclusiones deben hacer referencia a las preguntas de la evaluación.	
3. Las conclusiones se formulan tomando en cuenta las implicancias que tienen sobre los interesados, en especial los beneficiarios.	
4. La elaboración de recomendaciones solo toma en cuenta los intereses de los financiadores y/o equipo de gestión del programa o proyecto.	



UNIDAD DE APRENDIZAJE 6: ¿COMÓ EVALUAR EL INFORME FINAL DE UNA EVALUACIÓN?

6.1 Criterios para evaluar el informe de una evaluación

En general, se debe verificar en el informe que la evaluación realizada sea concordante con los Términos de Referencia que guiaron su ejecución y que contenga las explicaciones de los aspectos no cumplidos. Asimismo, que la lógica evaluativa y la metodología empleada sean los establecidos en los TDR y que sean consistentes con los objetivos buscados; si no fuera así, deben estar explicadas las razones de los cambios introducidos. Asimismo, se debe verificar que en el informe se muestre una adecuada articulación de las evidencias con los juicios de valor o de mérito.

Los informes de evaluación constituyen el producto tangible de todo el proceso de evaluación y deben satisfacer ciertos estándares. Mostramos en el cuadro 24 criterios para evaluarlos. También orientan los contenidos básicos que debe tener. El cuadro está orientado a diseños de evaluación cuantitativa.

Cuadro 24. Criterios para evaluar el informe de una evaluación

1. Resumen ejecutivo

El resumen ejecutivo proporciona una idea clara de:

- a. ¿El problema?
- b. ¿El programa?
- c. ¿El objetivo de la evaluación?
- d. Los principales hallazgos.
- e. Las implicaciones que tienen las conclusiones para el programa.
- f. Las recomendaciones.

2. Objetivos y preguntas de la evaluación

Esta sección indica:

- a. ¿Los objetivos de la evaluación?
- b. ¿Las preguntas?

3. Antecedentes del programa

Podemos encontrar en este capítulo:

- a. ¿El problema o necesidad que dio origen al programa?
- b. ¿Su cobertura geográfica y poblacional?
- c. ¿Los cambios relevantes ocurridos desde su origen?
- d. ¿Su estructura: componentes, organización?
- e. ¿Su Marco de Resultados?

4. Metodología

La sección de la metodología describe:

- a. ¿Un diseño de evaluación?
- b. ¿Los procedimientos de muestreo?
- c. ¿Los temas?
- d. ¿Los procedimientos para la recolección de datos?
- e. ¿Los instrumentos utilizados?
- f. ¿Los modelos y pruebas estadísticas?

5. Resultados

La sección de resultados contiene:

- a. ¿Los datos relacionados al problema o el propósito de las preguntas de evaluación?
- b. ¿Tablas, gráficos para resumir los hallazgos?
- c. ¿El detalle suficiente para entender lo que se encontró?

6. Conclusiones

La sección contiene:

- a. ¿Las principales conclusiones en relación a las preguntas de evaluación?
- b. ¿Las implicaciones prácticas para la agencia, programa o práctica? ¿Las recomendaciones?
- c. ¿Las limitaciones del estudio?

7. Recomendaciones

En la sección, las recomendaciones:

- a. ¿Están alineadas con los objetivos y las preguntas de evaluación?
- b. ¿Son propuestas de acciones específicas con fines prácticos?
- c. ¿Plantean mecanismos de mejora, tanto de procesos como de implementación?
- d. ¿Son realistas (factibles), oportunas y están orientadas a la acción?
- e. ¿Toman en consideración los aportes e intereses de los grupos de interés?
- f. ¿Son fundamentadas en hallazgos?
- g. ¿Establecen un marco estructurado de diálogo para las autoridades competentes?
- h. ¿Hay diferencia entre recomendaciones generales y específicas?

8. Referencias

¿Están las referencias mencionadas en el documento de la lista?

9. Anexos

¿El informe contiene copias de los instrumentos importantes, Términos de Referencia, cartas, procedimientos de análisis, resultados detallados de la muestra lograda y formularios de las encuestas?

10. Documento general

- a. ¿Es el estilo de escritura y la profundidad apropiada al público?
- b. ¿El informe es equilibrado y razonable?
- c. ¿Hay suficiente rigor para dar la credibilidad al estudio?

Fuente: Basado en Royse, D., Thyer, B. & Padgett, D. (2010). *Program Evaluation An Introduction*.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Bamberger, M, Rugh, J y Mabry, L. Evaluación en un Mundo Real. Trabajando Bajo Restricciones Presupuestarias, Cronológicas, de Información y Políticas. Visión General. Adaptado del capítulo 16, "Bringing it all together" de "Real World Evaluation: Working under Budget, Time, Data and Political Constraints", por Michael Bamberger, Jim Rugh y Linda Mabry. Sage Publications, 2006. Págs. 42-45.
- Faúndez, A. Guía de Evaluación de programas y proyectos con perspectiva de género, derechos humanos e interculturalidad. México. ONU Mujeres, 2014. Págs. 16-20.

ACTIVIDADES APLICATIVAS DE LA UNIDAD 6

Antes de ejecutar las actividades aplicativas, revisa tus respuestas de las preguntas de reflexión, analízalas, interprétalas y sintetízalas con el respaldo teórico que has adquirido al desarrollar la unidad 6

Ahora procede a desarrollar las actividades aplicativas de la unidad
6

Recuerda que el Producto de la unidad es:

UN INFORME DEL ANÁLISIS CRÍTICO DE UN INFORME FINAL DE UNA EVALUACIÓN

ACTIVIDAD 1

Analizando el informe final de una evaluación

1. Solicitar al coordinador o tutor **el informe final de la evaluación del programa/proyecto seleccionado en la actividad 1** y realiza lo siguiente:
2. Leer cuidadosamente el informe de evaluación.
3. Luego de haberlo leído realiza el análisis crítico aplicando los criterios técnicos anotados en el manual y utilizando la matriz de la siguiente página.

4. Tener presente que además de considerar que el informe cumple o no con el criterio, es necesario hacer un análisis de los aspectos técnicos, la coherencia, trascendencia, pertinencia y calidad de lo consignado en cada rubro del informe final.

Matriz de análisis crítico

CRITERIOS PARA EL ANÁLISIS DEL INFORME DE EVALUACIÓN	ANÁLISIS CRÍTICO
<p>Resumen ejecutivo Esta sección da una idea clara del:</p> <ul style="list-style-type: none"> - ¿Problema? - ¿Programa? - ¿Los objetivos de la evaluación? - ¿La metodología? - ¿Los principales hallazgos? - ¿Las implicaciones que tienen las conclusiones para el programa y las recomendaciones? 	
<p>Objetivos y preguntas de evaluación En esta sección:</p> <ul style="list-style-type: none"> - ¿los objetivos y las preguntas están redactadas de acuerdo al tipo de evaluación? - ¿Responden al Propósito de la evaluación? - ¿Las preguntas son factibles de ser respondidas? 	
<p>Antecedentes del PPP Podemos encontrar en esta sección:</p> <ul style="list-style-type: none"> - ¿El problema o necesidad que dio origen al programa? - ¿Su cobertura geográfica y poblacional? - ¿Los cambios relevantes ocurridos desde su origen? - ¿Su estructura: componentes, organización? - ¿Su marco de resultados? 	
<p>Metodología La sección de la metodología es coherente con los objetivos y preguntas y describe:</p> <ul style="list-style-type: none"> - ¿Un diseño de evaluación? - ¿Los procedimientos de muestreo? - ¿Los temas? - ¿Los procedimientos para la recolección de datos? - ¿Los instrumentos utilizados? - ¿Los modelos y pruebas estadísticas? 	
<p>Resultados La sección de resultados contiene:</p> <ul style="list-style-type: none"> - ¿Los datos relacionados al problema o el propósito de las preguntas de evaluación? - ¿Tablas, gráficos para resumir los hallazgos? 	

CRITERIOS PARA EL ANÁLISIS DEL INFORME DE EVALUACIÓN	ANÁLISIS CRÍTICO
<ul style="list-style-type: none"> - ¿El detalle suficiente para entender lo que se encontró? 	
<p>Conclusiones La sección contiene:</p> <ul style="list-style-type: none"> - ¿Las principales conclusiones en relación a las preguntas de evaluación? - ¿Las implicaciones prácticas para la agencia, programa o práctica?, ¿Las recomendaciones? - ¿Las limitaciones del estudio? 	
<p>Recomendaciones La sección recomendaciones:</p> <ul style="list-style-type: none"> - ¿Están alineadas con los objetivos y las preguntas de evaluación? - ¿Son propuestas de acciones específicas con fines prácticos? - ¿Plantean mecanismos de mejora, tanto de procesos como de implementación? - ¿Son realistas (factibles), oportunas y están orientadas a la acción? - ¿Toman en consideración los aportes e intereses de los grupos de interés? - ¿Son fundamentadas en hallazgos? - ¿Establecen un marco estructurado de diálogo para las autoridades competentes? - ¿Hay diferencia entre recomendaciones generales y específicas? 	
<p>Referencias</p> <ul style="list-style-type: none"> - ¿Están las referencias mencionadas en el documento de la lista? - ¿Aplican criterios técnicos de asentamiento bibliográfico? 	
<p>Anexos ¿El informe contiene copias de los instrumentos importantes, Términos de Referencia, cartas, procedimientos de análisis, resultados detallados de la muestra lograda y formularios de las encuestas?</p>	
<p>Documento general</p> <ul style="list-style-type: none"> - ¿Es el estilo de escritura y la profundidad apropiada al público? - ¿El informe es equilibrado y razonable? - ¿Hay suficiente rigor para dar la credibilidad al estudio? 	

5. Hacer una síntesis del análisis crítico de cada una de las secciones.
6. Si hubiera criterios adicionales, por ejemplo adaptaciones hechas para evaluaciones de desempeño o cualitativas a considerar, menciónelos y justifíquelos.
7. Escribir el resultado del análisis en un documento no mayor a cuatro páginas y entregar en la fecha indicada.

ACTIVIDAD 2

ESTA ACTIVIDAD SERÁ REALIZADA EN EL MOMENTO PRESENCIAL III

Analizando conclusiones, recomendaciones y lecciones aprendidas

La actividad será realizada tomando en cuenta el informe final seleccionado por el docente:

1. Leer el documento **(sin incluir las secciones de conclusiones y recomendaciones)** en no más de veinte minutos (puede distribuirse el número de páginas entre los participantes del grupo). Esta lectura es para tomar conocimiento acerca de por lo menos:
 - Los objetivos y preguntas de evaluación.
 - Antecedentes del PPP.
 - Métodos aplicados.
 - Hallazgos.
2. Redactar tres conclusiones y tres recomendaciones apropiadas que se deriven del informe. En este punto ustedes deben aplicar los criterios técnicos desarrollados en el módulo.
3. Leer las conclusiones y recomendaciones originales del documento. ¿Qué tanto se diferencian las efectuadas por los evaluadores de las efectuadas por ustedes, los participantes? ¿Qué lecciones se puede extraer de este contraste?

Para este análisis tenga en cuenta los siguientes criterios:

- ¿Las conclusiones responden a las preguntas de evaluación?
- ¿Las conclusiones aluden a los estándares u objetivos del programa?
- ¿Las conclusiones aluden a las limitaciones metodológicas del estudio?
- ¿Las recomendaciones derivan de las conclusiones?
- ¿Se ofrecen y discuten varias alternativas de recomendaciones?
- ¿Las recomendaciones están hechas para que el equipo de gestión del programa/financiador/formulador de políticas pueda tomar alguna decisión?
- ¿Se incluyen lecciones aprendidas? ¿Estas cumplen con los criterios de redacciones de lecciones aprendidas?
- ¿Las lecciones aprendidas están dirigidas para el usuario pertinente?

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 6

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Señala si son verdaderas o falsas cada una de las siguientes afirmaciones:

Afirmación	Verdadera o falsa
1. Una lista de cotejo para evaluar el diseño es útil para los que practican la evaluación, así como para aquellos que la gestionan.	
2. Se puede deducir del cuadro 21 que la elección de los métodos a elegir depende de las preguntas de evaluación.	
3. Un informe de evaluación debe ser objetivo, es decir, evitar mostrar los valores e intereses de los involucrados en una evaluación.	
4. El informe de evaluación debe ser riguroso pero entendible por los interesados en la evaluación.	



SECCIÓN 3

SECCIÓN 3

ELABORACIÓN DE TÉRMINOS DE REFERENCIA

CAPACIDAD A LOGRAR

Elaborar Términos de Referencia acordes con diseño de la evaluación requerida.

PRODUCTO

Términos de Referencia para una evaluación.

CONTENIDOS

Unidad 1: ¿Cómo elaborar Términos de Referencia para una evaluación?

MATERIALES INFORMATIVOS DE APOYO

UNIDAD 1

- TDR Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Programa de Patrimonio Arqueológico. EDEPS MEF, 2010.
- TDR Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Proyecto Construyendo Perú. EDEPS MEF, 2009.
- Medina, A. Guía para la elaboración de Términos de Referencia Orientados a la Contratación de Servicios de Consultoría para la Evaluación de Programas y Proyectos Gubernamentales. Centro Latinoamericano de Administración para el Desarrollo. Banco Mundial, 2009. Disponible en: <http://old.clad.org/documentos/otros-documentos/guia-para-la-elaboracion-de-terminos-de-referencia-orientados-a-la-contratacion-de-servicios-de-consultoria-para-el-monitoreo-y-evaluacion-de-programas-y-proyectos-gubernamentales>

Reflexionando mi práctica sobre la elaboración de Términos de Referencia y planes de evaluación

TE INVITAMOS A REVISAR Y ANALIZAR LO SIGUIENTE:

Existe un crecimiento importante de las asignaciones presupuestales para programas sociales. El Ministerio de Economía y Finanzas anunció que para el 2014 se otorgaría 4 880 millones de nuevos soles para esos fines; esta cifra representa un crecimiento de poco más del 9% con respecto al 2013. El monto invertido en esos programas el 2013 era a su vez 10% más que lo asignado el 2012.

Cuna Más tendrá en el 2014 un presupuesto de S/ 264 millones; Pensión 65, S/ 578 millones; Qali Warma, S/ 1268 millones; Juntos, S/1028 millones; y Trabaja Perú, con S/ 64 millones son los programas con los presupuestos más altos. Las cifras presupuestales son elevadas y desde la perspectiva de la evaluación es necesario que se pueda garantizar y medir su rendimiento social (cumplimiento de objetivos, impactos, desempeño, etc.).

¿Cuántos de estos programas han planificado de manera sistemática el conjunto de sus actividades de evaluación a lo largo de su ciclo de vida? Planificar las evaluaciones de un programa permite anticipar qué se va a evaluar y las estrategias para lograrlo. Por el contrario, intentar solucionar los asuntos de la evaluación conforme van apareciendo solo lleva a perder oportunidades para i) tener diseños consistentes de evaluación que permitan, por ejemplo, saber el impacto de un programa, ii) identificar estándares para evaluar el desempeño del programa, iii) desconocer cuánto nos costarán los estudios de evaluación y en qué momento se llevarán a cabo y iv) recibir cuestionamientos porque determinados asuntos clave del programa no se evaluaron.

Analiza el caso y responde lo siguiente:

- ¿Por qué es necesario que las instituciones cuenten con planes de evaluación?
- En nuestro país, ¿existe la cultura de elaborar planes de evaluaciones anuales o multianuales?

Ahora que culminaste las actividades de reflexión, desarrolla los contenidos y actividades de las unidades. Si necesitas apoyo de los tutores, comunícate con ellos a través de la plataforma virtual.

UNIDAD DE APRENDIZAJE 1: ¿CÓMO ELABORAR TÉRMINOS DE REFERENCIA PARA UNA EVALUACIÓN?

1.1 ¿Por qué es necesario elaborar Términos de Referencia (TDR)?

Es necesaria la elaboración de los Términos de Referencia para estandarizar los criterios de una evaluación y para presentar una estructura básica que facilite a los servidores públicos a elaborar los Términos de Referencia para la contratación de evaluaciones. Los Términos de Referencia definen los objetivos, alcances y la calidad de estudio a preparar, presentando un listado ordenado y específico de los aspectos más importantes a ser considerados por la consultoría, elementos que servirán como fuente de información para facilitar la evaluación del trabajo y posteriormente para la toma de decisiones sobre la viabilidad y conveniencia de implementar el programa o proyecto de inversión.

1.2 ¿Cuál es el proceso de elaboración de los TDR para una evaluación?

Las etapas para elaborar los Términos de Referencia para una evaluación, según el manual del Banco Mundial (2009), son:

1.2.1 Identificar las necesidades de la contratación:

Michael Hammer (2001), acuñó el término "Mayor Valor Añadido" (MVA), es decir, el interés de los consumidores no se detiene en la compra del bien o servicio que generan las empresas, lo que realmente les interesa es solucionar problemas a través no de uno, sino de un conjunto integrado de bienes y servicios. El principio MVA se puede visualizar como una escalera con los productos en el primer escalón y la solución al problema del cliente en el último.

Este principio se puede aplicar en las entidades gubernamentales que contratan servicios de evaluación para PPP gubernamentales, pensando en entregar mayor valor añadido al prestar servicios públicos. Concretamente, se debe tener presente: ¿para qué se contrata el estudio de evaluación?, ¿qué problema o conjunto de necesidades va a solucionar el programa o la política para la que este se inscribe?, ¿con qué otros estudios de evaluación pueden combinarse e integrarse?, ¿qué otras evaluaciones son necesarias y no se han llevado a cabo?, ¿cuál es la consecuencia buscada con la contratación de los servicios de evaluación?, es decir, ¿cuál es la motivación principal que impulsa el estudio?

Al elaborar los Términos de Referencia se debe lograr que la necesidad que da origen a los requerimientos de evaluación de los programas sea explícita y clara para tener la certeza de la utilidad del estudio para su resolución o atención. Esto es muy importante, ya que existe una variedad de posibilidades metodológicas para realizar evaluaciones y debe seleccionarse aquella que responda concretamente a las necesidades identificadas y los problemas planteados, de no ser así, podrán realizarse evaluaciones ineficientes y de bajo impacto.

Además, si la necesidad de los estudios de evaluación de programas no es claramente identificada y las preguntas que estos deben responder para los principales actores involucrados no son formuladas adecuadamente, la evaluación puede ser completamente inútil.

Hay muchas opciones metodológicas de evaluación, pero lo más importante es tener en cuenta la relación entre la necesidad que se identifica y la orientación del estudio a contratar. Por ejemplo, a través de una evaluación del diseño del programa puede conocerse si el programa está articulado a un plan mayor, si existe alguna duplicidad con otros programas, si los objetivos de diverso nivel son pertinentes y congruentes entre sí, entre otros aspectos. Si lo que se ha identificado es la necesidad de conocer el impacto del programa en las condiciones de bienestar de su población objetivo y las características de los beneficiarios, lo más adecuado es una contratación de evaluación de impacto.

1.2.2 Identificar el diseño básico de la evaluación:

El diseño básico de las evaluaciones de programas y proyectos gubernamentales a ser contratadas puede variar en dimensiones o atributos conformando una amplia diversidad de tipos de evaluación.

a) Identificar al agente que va a realizar la evaluación:

Considerando al tipo de agente que va a realizar las evaluaciones, se tiene: i) evaluaciones internas, cuando es realizada por especialistas de la propia institución, ii) evaluaciones externas, cuando es realizada por un agente ajeno a la institución que desarrolla el PPP, iii) evaluaciones mixtas, cuando participan evaluadores internos y externos y iv) evaluaciones participativas, cuando se involucra a los beneficiarios primarios del PPP.

b) Especificar el diseño de la evaluación que se ha considerado

El diseño básico de la evaluación debe incluirse en este acápite de los TDR. Se mencionarán:

- El objeto de evaluación, es decir, si va a evaluar un proyecto, un programa, un plan, o una política.
- Los objetivos de la evaluación.
- La validación de la Teoría de Cambio.
- Las preguntas o hipótesis de la evaluación
- Criterios de evaluación
- Los lineamientos metodológicos del estudio: los niveles de inferencia de la muestra, precisión y confianza.
- Las técnicas de recolección de información: encuestas, grupos de enfoque, entrevistas, etc.

1.2.3 El análisis de mercado:

Cuando ya se tiene la claridad en la necesidad, se requiere que los responsables del área solicitante se apoyen en la experiencia y conocimientos técnicos acerca de las distintas opciones de evaluación, para definir qué tipo de servicio se va a contratar. También se debe llevar a cabo una investigación de mercado que dará información sobre si existe un número suficiente y con capacidad técnica y financiera de evaluadores externos para realizar la evaluación con base en los Términos de Referencia. Además un estudio de mercado permite estimar el precio del servicio a contratar en relación con el precio de mercado de servicios similares.

Los TDR son un documento que usualmente requiere de distintas versiones y ajustes, así como de la opinión de un grupo multidisciplinario de expertos para lograr la precisión requerida en su elaboración.

Cuando se tiene una variedad de aspectos técnicos a considerar en los TDR para algunos estudios muy complejos, es recomendable contratar a especialistas para apoyar su elaboración. Los especialistas, ya sean internos o externos, facilitan y apoyan en la preparación de los Términos de Referencia. Entre más claros,

precisos y detallados sean los TDR, sin duda menores serán los costos de transacción derivados de su implementación.

Además, es importante señalar que los especialistas externos involucrados en la elaboración de los TDR no podrán competir para la ejecución de los trabajos, ya que se estaría incurriendo en una situación clara de conflicto de interés. Estos especialistas podrían servir de apoyo en la calificación de las propuestas presentadas por los licitantes participantes.

Es indispensable que tanto la entidad contratante que administra el proceso de contratación como el área técnica solicitante tengan en conjunto los conocimientos técnicos y administrativos suficientes para poder constituirse en contrapartes adecuadas y realizar la supervisión que se requiere durante la ejecución de los servicios.

1.3 Contenidos de los Términos de Referencia para una evaluación

Los contenidos de los TDR de una evaluación y una lista de verificación de la calidad de los Términos de Referencia se describen en el cuadro 25.

Cuadro 25. Contenidos de los TDR, conceptos y preguntas de verificación

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
<p>1. Antecedentes</p>	<ul style="list-style-type: none"> • Realizar una breve descripción del programa o proyecto gubernamental que se pretende evaluar y de las necesidades a satisfacer. • Incluir una síntesis de la problemática que el programa aborda y los supuestos de su intervención. • Incorporar en la descripción los aspectos generales del programa: objetivo principal, tiempo de implementación, población objetivo, cobertura geográfica, metas y logros de la intervención, principales hallazgos de evaluación previa, contexto institucional en el que se ubica el programa, aspectos organizativos y de gestión del programa, asignación presupuestal, etc. • Identificar las fuentes de financiamiento de la consultoría. <p>Incluir definiciones o términos para evitar interpretaciones ambiguas sobre el funcionamiento del programa. Se deben explicar dentro de marco conceptual o también en un glosario que forma parte de los anexos.</p>	<ol style="list-style-type: none"> 1. ¿Está claramente identificada la problemática que justifica la intervención? 2. ¿Se proporciona la información necesaria y suficiente para comprender los orígenes y fines últimos que persigue el servicio a contratar? 3. ¿Están claramente descritos los antecedentes del programa a evaluar? 4. En caso de que se requiera precisión conceptual, ¿los antecedentes incluyen un marco conceptual?
<p>2. Objetivo general y objetivo específico de la consultoría</p>	<ul style="list-style-type: none"> • En el objetivo general se debe describir el propósito esencial que busca los servicios de consultoría a contratar. • En los objetivos específicos se deben señalar con precisión y detalle los propósitos concretos que se pretende alcanzar, como redefinición de metas, definición y ajuste de mecanismos de focalización, elaboración de indicadores, integración de 	<ol style="list-style-type: none"> 5. ¿Se describe claramente el propósito que persigue la consultoría? 6. ¿Existe solo un objetivo general? 7. ¿Los objetivos específicos describen con precisión y detalle los propósitos concretos que

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p>fuentes de información, diseño de sistemas automatizados de indicadores, etc.</p> <ul style="list-style-type: none"> • Los objetivos sirven como una guía para el consultor en la ejecución de los trabajos designados y una pauta de supervisión para la entidad contratante. 	<p>deben alcanzarse mediante los servicios a contratar?</p> <p>8. ¿Definen los objetivos de manera clara los beneficios que el área técnica y/o la entidad contratante pretenden obtener de los servicios adjudicados?</p> <p>9. ¿Pueden estar los objetivos específicos principalmente bajo el control del consultor y la entidad contratante?</p> <p>10. ¿Está el objetivo general y los específicos claramente alineados a la problemática que da origen a la intervención?</p>
<p>3. Alcance, enfoque de la evaluación, actores implicados y sus responsabilidades</p>	<ul style="list-style-type: none"> • Definir la amplitud y el enfoque de la evaluación es fundamental al redactar los TDR en el campo de monitoreo y evaluación de programas gubernamentales. • Incorporar con precisión y claridad la lista de preguntas clave que deben ser resueltas con la evaluación. • Identificar los atributos más importantes que faciliten la identificación de la modalidad de evaluación requerida como el agente que la realizará (interno, externo o combinación de estos), aspecto o aspectos gerenciales a evaluar (diseño, procesos, satisfacción de beneficiarios, información y/o impacto), el ciclo de vida en que se ubica la evaluación (<i>ex</i> 	<p>11. ¿Se ha definido con claridad un conjunto básico de preguntas clave que responderá la evaluación y que permita definir su modalidad?</p> <p>12. ¿Han sido identificadas las dimensiones o atributos principales de la evaluación, tales como tipo de agente evaluador, aspecto gerencial a evaluar, metodología a utilizar, etc.?</p> <p>13. ¿Está la orientación metodológica claramente definida, incluyendo técnicas a aplicar, o bien criterios para evaluarla si es el proveedor quien la desarrollará?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p><i>ante</i>, concurrente y de medio término, <i>ex post</i>), y otras dimensiones o atributos de la evaluación antes descritos</p> <ul style="list-style-type: none"> • Definir el periodo temporal que comprende la evaluación. • Explicar los criterios de evaluación que se asocian a los atributos: pertinencia, validez del diseño, eficacia, eficiencia, productividad, sostenibilidad, impacto, externalidades y otros. • Incorporar una matriz que defina brevemente los principales actores comprendidos en la evaluación, así como su rol, responsabilidades, nivel de participación, y su importancia. 	
4. Metodología	<ul style="list-style-type: none"> • Incluir elementos conceptuales básicos. • Aplicar técnicas cuyo empleo aumentará la probabilidad de los resultados esperados. • Incorporar métodos y técnicas sistematizadas para que sean empleadas por el consultor. • Realizar estrategias de desarrollo de la consultoría. 	<p>14. Si la metodología viene definida en los TDR, ¿está debidamente definida la metodología por la cual la entidad contratante desea que se procese la información?</p> <p>15. Si el estudio incluye una encuesta por muestreo, ¿se solicitan aspectos de diseño de la muestra, tamaño de muestra y factores de expansión, así como otros elementos de documentación estadística?</p>
5. Actividades a realizar	<ul style="list-style-type: none"> • Describir con precisión las tareas y principales actividades a ser ejecutadas por la consultoría, definiendo su secuencia y articulación, así como los productos y resultados que se esperan obtener, y que permitan lograr el objetivo general y los objetivos específicos de la consultoría a contratar. 	<p>16. ¿Están claramente definidas las actividades y tareas que se espera desarrollará el consultor?</p> <p>17. ¿Están las actividades y tareas claramente relacionadas con los objetivos?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<ul style="list-style-type: none"> • En los TDR no se debe establecer una excesiva, dispersa y desconectada variedad de actividades porque las tareas de seguimiento y supervisión se hacen más complejas y resulta difícil identificar su contribución al logro del objetivo. • Incluir en las actividades de campo el trabajo de gabinete o de oficina y los soportes con los que se respalda la información recolectada. Además todas las actividades deben estar contenidas en un cronograma de trabajo, señalando las fechas que deben ser presentados los escritos a la entidad contratante. • Las tareas, según su naturaleza, deben precisar los requerimientos organizativos y técnicos para que puedan llevarse a cabo. Por ejemplo: las modalidades y técnicas de levantamiento de datos, metodologías a utilizar, necesidades especiales de software, georeferenciación de datos, etc. 	<p>18. ¿Existe una descomposición del trabajo de manera jerárquica que vincule objetivos, productos, actividades y tareas?</p> <p>19. En caso de requerirse transferencia de conocimiento, ¿se proveen detalles específicos de las características de la transferencia y se solicita a los consultores que presenten sus planes y métodos de entrenamiento?</p>
<p>6. Productos e informes a entregar</p>	<ul style="list-style-type: none"> • Definir los productos con la mayor precisión posible. Su presentación y grado de avance se da en tres informes: informe inicial, de avance y final. • Determinar la frecuencia de los informes periódicos de avances y las características de los contenidos. • Proporcionar la extensión, el número de ejemplares a entregar, tipo y tamaño de letra, márgenes, tablas, anexos, tipo de soporte en el que se entregarán el impreso y/o magnético, exigencia de un resumen ejecutivo, idioma o 	<p>20. ¿Se definen con precisión y tiempos los entregables que el consultor deberá presentar a la entidad contratante?</p> <p>21. ¿Se definen los informes a presentar, su secuencia, oportunidad y características?</p> <p>22. ¿Se indica el plazo que tendrá la entidad contratante para revisar, aprobar o rechazar los informes o productos?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p>idiomas en que serán presentados, nombre y cargo de las personas que recibirán los informes, etc.</p> <ul style="list-style-type: none"> • En las evaluaciones que involucren la realización de operativos estadísticos, los requerimientos deben ser muy precisos e incluir la documentación requerida como respaldo, objetivos de la operación estadística, instrumento de colecta de información, diseño y tamaño de la muestra, trabajo de campo, etc. 	
<p>7. Duración del servicio</p>	<ul style="list-style-type: none"> • Estimar el periodo para la ejecución del servicio de la consultoría. • Señalar con sumo cuidado la estimación del plazo máximo. Lo más recomendable es que se haga una descomposición estructurada del trabajo. Esta técnica ayuda a distribuir el trabajo en paquetes o bloques facilitando la administración y el seguimiento del alcance del trabajo; además constituyen la base para la asignación de recursos al proyecto y para el análisis de riesgos, costos y elaboración del cronograma. • Incluir un calendario y un cronograma tentativo de ejecución de la consultoría, especificando el periodo en el cual se desarrollará la consultoría (fecha de inicio estimada), las actividades a desarrollar y las fechas parciales y finales previstas. • Si a la entidad contratante no le es posible hacer una estimación confiable en términos de la precisión en calendario, 	<p>23. ¿Se establece el plazo de ejecución del servicio?</p> <p>24. ¿Se indica si el plazo de ejecución fijado es un máximo o es indicativo?</p> <p>25. Si el plazo es máximo, ¿la estimación realizada tiene fundamento en alguna metodología aplicada consistentemente?</p> <p>26. ¿Está la duración de la consultoría sujeta a tiempos legales de entrega de productos a la institución contratante? Si esto es así, ¿se encuentra esta situación claramente definida en los TDR?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	podrá establecerse un plazo indicativo que podrá ser excedido o reducido justificadamente por el licitante en su propuesta.	
<p>8. Recursos y facilidades a ser provistas por la entidad contratante</p>	<ul style="list-style-type: none"> • Realizar una lista de recursos y facilidades que la institución contratante proveerá al licitante que resulte adjudicado. Esto debe incluir la información, bienes y servicios que el área técnica considere necesario para llevar a cabo la consultoría, por ejemplo: espacio de oficinas, vehículo, equipo de computación, equipo de impresión, sistema de telecomunicación, etc. En el caso de evaluaciones que involucran trabajo de campo, estos recursos y facilidades constituyen una parte fundamental del trabajo de la consultoría. • Indicar el horario laboral que tendrá el personal que la entidad contratante ponga a disposición del consultor. • Incluir también los documentos que deben ser entregados al consultor y aquellos a los que puede recurrir, dónde los puede encontrar y quién se la puede suministrar. • En algunos casos, para poder realizar la consultoría se requiere visitar propiedad privada o estatal, o consultar información que no es propiedad de la entidad contratante; por ejemplo, la consulta de bancos de información en oficinas de estadística que están protegidos por el principio de confidencialidad o anonimato, etc. En estos casos, la unidad contratante se debe comprometer a solicitar los permisos o 	<p>27. ¿Se indican claramente los insumos que la entidad contratante suministrará al evaluador externo?</p> <p>28. ¿Se indica el personal de la entidad contratante que actuará como contraparte del evaluador externo que resulte adjudicado?</p> <p>29. ¿Se indica el personal de apoyo que la entidad contratante pondrá a la disposición del evaluador externo?</p> <p>30. ¿Se indican los documentos e información que la entidad contratante pondrá a disposición del evaluador externo?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p>autorizaciones necesarias para tal fin, y esta responsabilidad debe ser explícitamente definida en los TDR.</p> <ul style="list-style-type: none"> • Cuando los trabajos vayan a realizarse en las instalaciones de la entidad contratante, esta podría comprometerse a suministrar el espacio necesario, y en algunos casos las facilidades de comunicación, reproducción y otros servicios generales. 	
<p>9. Costos y remuneración</p>	<ul style="list-style-type: none"> • Indicar las reglas aplicables a la determinación de los honorarios, gastos reembolsables y otros costos. • Incluir la forma de cálculo de honorarios, en qué moneda se realizará el pago, formas de pagos en base a la cual el consultor prepara su propuesta. • Considerar los impuestos y tasas locales, ya que constituyen un costo para los consultores. Estos costos normalmente tienen dos componentes: los honorarios o remuneración de los consultores y los denominados gastos reembolsables. • Establecer qué elementos constituyen la remuneración, ya que varían en función a la legislación, industria y sector. Además, la remuneración de los consultores se debe distinguir entre locales y extranjeros, personal profesional, personal de apoyo o administrativo, personal de campo o de oficina, las tarifas debe ser sobre la base de riesgo del país, complejidad del servicio y localización. • Incluir en los costos los gastos imprevistos en un rango que oscila entre un 5% y un 8%, aunque los multilaterales como 	<p>31. ¿Se establece con claridad la modalidad de pago; por ejemplo, suma alzada, pago de remuneración y gastos reembolsables, etc.?</p> <p>32. ¿Se establece de manera precisa la forma de pago y la distribución porcentual de cada pago parcial?</p> <p>33. ¿Se indica el tipo de moneda en que se deberá presentar la oferta económica?</p> <p>34. ¿Se precisa cuáles serán los gastos reembolsables?</p> <p>35. ¿Se especifica, de ser necesario, el mecanismo para el pago de impuestos por parte del consultor?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p>el BID y el BM manejan porcentajes entre el 10 y el 15% para las consultorías en las que, dada su naturaleza, es difícil definir con precisión y detallar las actividades.</p> <ul style="list-style-type: none"> • Estimar adecuadamente los costos es muy importante ya que un desfase entre el costo estimado y las actividades previstas puede generar ofertas inadecuadas y problemas durante el transcurso de la implementación. 	
<p>10. Perfil del evaluador externo</p>	<ul style="list-style-type: none"> • La entidad contratante debe indicar los requerimientos mínimos profesionales, divididos en dos partes, formación y experiencia; en lo referente a la formación debe definirse el nivel mínimo (licenciatura, posgrado, doctorado); por su parte, la experiencia profesional se refiere a la experiencia previa en la realización de estudios similares. • Precisar la especialidad que tiene el consultor (consultoría individual), o la del consultor y los miembros de su equipo (firmas de consultoría). • Recaltar el país de residencia, ya que a veces hay fondos multilaterales para la evaluación que exigen la nacionalidad. • En el caso de los consultores extranjeros es requisito indispensable tener conocimiento del país, región o regiones afines, excepto cuando se quiere hacer una integración económica y social de otros países. • Especificar las habilidades requeridas, excepto el número o la composición del equipo que se requiere. 	<p>36. ¿Se indican las características profesionales requeridas y experiencia para el líder del proyecto y los miembros clave del equipo si se trata de una firma, o, en su caso, del consultor individual?</p> <p>37. ¿Se especifica la necesidad de viajar y requerimientos de nacionalidad?</p> <p>38. ¿Se especifican todas aquellas situaciones que pueden constituirse en conflicto de interés, y cómo manejarlas?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<ul style="list-style-type: none"> Indicar la capacidad operativa y financiera requerida de las firmas consultoras porque constituye un criterio muy importante para la evaluación. <p>Especificar las limitaciones y prohibiciones a los evaluadores externos. De no respetarse pueden constituir conflictos de interés, por ejemplo, no tener procesos legales pendientes con la entidad contratante, no haber colaborado recientemente como funcionario o consultor de la entidad contratante.</p>	
<p>11. Criterio de selección/evaluación</p>	<ul style="list-style-type: none"> Proporcionar al evaluador externo interesado la posibilidad de realizar un "ejercicio de autoevaluación" y juzgar sobre la base de los resultados, si vale la pena presentarse y participar. Incorporar en los TDR un apartado de criterios de selección o evaluación que permitan a los evaluadores externos que, al autoevaluarse, perciban que tienen pocas posibilidades, ya sea autodescartarse o bien fortalecer aquellas áreas en las que tienen menor calificación. Considerar los siguientes criterios de evaluación. Experiencia específica en relación con la metodología de evaluación solicitada y el campo. Puntaje 0-10. Calidad de la metodología-plan de trabajo. Puntaje 20-50. Calificaciones personal clave. Puntaje 30-60. Transferencia de conocimientos (de ser necesario). Puntaje 0-10. 	<p>39. ¿Es posible que los evaluadores externos puedan realizar una autoevaluación objetiva sobre criterios específicos para juzgar si vale la pena participar en el proceso?</p> <p>40. ¿Los criterios consideran, como mínimo, experiencia, metodología/plan de trabajo, transferencia de tecnología -si aplica- y calidad del personal?</p> <p>41. ¿Se especifica el valor y la ponderación de cada criterio y subcriterio definido?</p>

Contenidos	Conceptos de los contenidos de los TDR	Preguntas para verificar los contenidos de los TDR
	<p>Otros (participación de personal nacional). Puntaje 0-10.</p> <ul style="list-style-type: none"> • Determinar cuál es la mejor oferta que cumple con lo previsto en los TDR para prestar el servicio con la calidad y estándar requerido. Se recomienda que sea el área técnica que identifique la necesidad y se involucre en la redacción de los TDR. Además de liderar el proceso de identificación de criterios de evaluación. 	
12. Confidencialidad de la información	<ul style="list-style-type: none"> • Especificar el grado de privacidad de la información de los estudios, tanto en la información que se entrega para la realización de los estudios como la que se genera durante la realización de las actividades. • Entregar a los consultores toda la información relativa a las normas vigentes en relación con la diseminación estadística. 	<p>42. ¿Se incluyen los párrafos adecuados para salvaguardar la información confidencial?</p> <p>43. ¿Se hace referencia, de ser el caso, a la normativa vigente relativa a los temas de confidencialidad de la información estadística?</p>
13. Anexos	<p>Incluir los mapas, síntesis de evaluaciones anteriores, información gerencial del programa o proyecto (diagnóstico, Marco Lógico, informes de avances), listados de beneficiarios (padrones), información de los sistemas informáticos administrativos y de apoyo del programa, información administrativa y de ejecución presupuestal, organigramas, información relativa a la entidad financiadora, referencias a las estrategias del Plan Nacional de Desarrollo a las que se articula el programa, informes financieros, glosarios, etc.</p>	<p>44. ¿Se listan los anexos que acompañan a los TDR, es decir, cumplimiento del servicio, términos de confidencialidad, etc.?</p> <p>45. De ser necesario, ¿están definidos en un glosario los términos técnicos o poco usuales?</p>

Fuente: Banco Mundial (2009) Guía para la elaboración de Términos de Referencia orientados a la contratación de servicios de consultoría para la evaluación de programas y proyectos gubernamentales.

Ejemplos de Términos de Referencia (TDR):

- TDR Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Programa de Patrimonio Arqueológico. EDEPS MEF, 2010.
- TDR Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Proyecto Construyendo Perú. EDEPS MEF, 2009.

1.4 Diseño de la supervisión y monitoreo de una evaluación

Supervisar evaluaciones es una actividad que un evaluador siempre tendrá que realizar, en especial si está a cargo de alguna unidad de Monitoreo y Evaluación de un programa o proyecto. En el cuadro 26 se presenta una lista de cotejo de los contenidos y acciones para la supervisión y seguimiento de una evaluación.

Cuadro 26. Contenidos y acciones para la supervisión y seguimiento de una evaluación

Contenido	Acción
1. Diseño	Verifica que los indicadores/dimensiones sirvan para responder las preguntas de interés.
	Verifica que la muestra cumpla con los estándares establecidos: niveles de inferencia, tamaño, procedimientos de selección.
2. Instrumentos	Verifica que los instrumentos servirán para recoger toda la información necesaria para construir los indicadores o caracterizar las dimensiones.
	Verifica que los instrumentos han sido validados.
3. Personal	Verifica que el personal cumple con los estándares establecidos.
4. Capacitación	Verifica que la capacitación se lleva a cabo cumpliendo los contenidos de un plan de capacitación.
	Verifica que el tiempo de la capacitación es suficiente para el logro de competencias.
5. Trabajo de campo	Verifica que se cumplen las cargas de trabajo asignadas al personal.
	Verifica que se cumplen los procedimientos de selección de la muestra.

Contenido	Acción
	Verifica que existan manuales de supervisión y trabajo de campo.
	Verifica que el recojo se ejecuta de manera idónea.
	Verifica que se supervise el trabajo de campo.
	Verifica la organización del trabajo de campo.
6. Ingreso de datos	Verifica la existencia de manuales de consistencia e ingreso de datos.
	Verifica la existencia de procedimientos de control en el ingreso de datos.
	Verifica que el software de ingreso de datos funcione sin errores.
	Verifica que se realicen procesos de consistencia de datos.
7. Procesamiento de datos	Verifica que exista un plan de análisis de datos.
	Verifica que modelos y pruebas se desarrollen conforme a los supuestos establecidos.
	Verifica la aceptación o rechazo de hipótesis se haga de acuerdo a estándares (para evaluaciones de impacto).
	Verifica que los análisis de desempeño cumplan los requerimientos de idoneidad (para evaluaciones de desempeño).
8. Informes	Verifica que se cumplan los criterios de calidad de un informe de evaluación.

Elaborado por Alejandro Bardales, 2014.

Para ampliar el tema tratado, revisa la siguiente lectura:

- Medina, A. Guía para la elaboración de Términos de Referencia Orientados a la Contratación de Servicios de Consultoría para la Evaluación de Programas y Proyectos Gubernamentales. Centro Latinoamericano de Administración para el Desarrollo. Banco Mundial. 2009. Disponible en: <http://old.clad.org/documentos/otros-documentos/guia-para-la-elaboracion-de-terminos-de-referencia-orientados-a-la-contratacion-de-servicios-de-consultoria-para-el-monitoreo-y-evaluacion-de-programas-y-proyectos-gubernamentales>

ACTIVIDADES APLICATIVAS DE LA UNIDAD 2

Ahora procede a desarrollar las actividades aplicativas de la unidad

Recuerda que el Producto de la sección es el: TDR para una evaluación

ACTIVIDAD: Elaborando los TDR para una evaluación

1. Considerando el diseño de la evaluación que has elaborado en la sección 1 del módulo, elabora los TDR para ejecutar dicha evaluación.
2. Realiza las consultas que consideres necesario.
3. Presenta el producto en la fecha anotada en el sílabo.

ACTIVIDADES DE AUTOEVALUACIÓN DE LA UNIDAD 1

HAS CULMINADO LA PRESENTE UNIDAD, TE INVITAMOS A REALIZAR LA AUTOEVALUACIÓN DE LOS CONTENIDOS:

Responde si las siguientes afirmaciones son verdaderas, falsas o inciertas

Pregunta	Verdadero (V) o falso (F) o incierto (I)
Los costos estimados de una evaluación se deben establecer sobre la base de costos de mercado.	
El diseño de la evaluación es una etapa previa a la confección de TDR	
Los TDR de una evaluación deben señalar los responsables de darles seguimiento.	
Los TDR de una evaluación deben indicar los tiempos de revisión de los productos.	



BIBLIOGRAFÍA DEL MÓDULO**Sección 1.****Unidad 1. Conceptos básicos sobre evaluación**

EuropeAid Oficina de Cooperación. Metodología de Evaluación de la Ayuda Exterior de la Comisión Europea. Francia. Comisión Europea, 2006. p. 42-53.

Gertler, P. y otros. La Evaluación de Impacto en la Práctica, Washington DC. The International Bank for Reconstruction and Development, 2011. p. 1-47.

OECD. Glosario de los Principales Términos sobre Evaluación y Gestión Basada en Resultados. Paris. OECD.

UNEG. Integración de los Derechos Humanos y la Igualdad de Género en la Evaluación- Hacia una guía de la UNEG. United Nations Evaluation Group. 2011. p. 13-16.

USAID. PPT para el curso Evaluation for Programs Managers. 2013.

Unidad 2. Diseño de evaluaciones

Canudas, R, coordinadora. Curso de Indicadores para el Desarrollo. Washington DC. Instituto Interamericano para el Desarrollo Social (INDES), Banco Interamericano de Desarrollo, 2008. Módulo 2: Propiedades y Construcción de Indicadores.

Gertler, P. y otros. La Evaluación de Impacto en la Práctica, Washington DC. The International Bank for Reconstruction and Development, 2011. Capítulos 3, 4, 7 y 11.

Holden, D and Zimmerman, M. Editors. A Practical Guide to Program Evaluation Planning. Theory and Case Examples. California. Sage Publications, Inc. 2009. Chapter II.

W. K. Kellogg Foundation. Logic Model Development Guide. Michigan. The W.K. Kellogg Foundation. 2004.

Sección 2.

Unidad 1. Pautas para el desarrollo de la Prueba Piloto del Cuestionario

Ministerio de Educación. Dirección de Educación Superior Pedagógica. Sistematización del Plan Piloto de Evaluación del Desempeño Docente. Lima. Academy for Educational Development, Proyecto USAID/PER/SUMA. 2011. p. 24 y 25.

Unidad 2. Proceso de recolección de la información

CERLALC. Manual de Aplicación de un Modelo de Evaluación para Identificar el Aporte de las Bibliotecas Públicas en el Desarrollo de sus Comunidades. Manual de Lecturas. Santiago de Chile. Centro Regional para el Fomento del Libro en América Latina y el Caribe (CERLALC). 2012. Capítulo 3

Hernández, R, Fernández-Collado, C. y Baptista, P. Metodología de la Investigación. Cuarta edición. México DF. McGraw-Hill Interamericana Editores S.A de C.V. 2006. Capítulo 10 y Capítulo 14.

Instituto de Nutrición de Centro América y Panamá INCAP. Manual de Operaciones para Recolección de Datos en un Sistema de Vigilancia Alimentario Nutricional. Guatemala, INCAP. 2012. p. 1-10

Unidad 3.

Feres, J. y Medina, F. Control de Calidad en el Diseño y Ejecución de Encuestas de Hogares, en Taller 1 Planificación y Desarrollo de Encuestas de Hogares para la Medición de las Condiciones de Vida, Programa de Mejoramiento de las

Encuestas y la Medición de las Condiciones de Vida en América Latina y el Caribe.. Santiago de Chile. Comisión Económica para América Latina y el Caribe (CEPAL). 1998.

Hernández, R, Fernández-Collado, C. y Baptista, P. Metodología de la Investigación. Cuarta edición. México DF. McGraw-Hill Interamericana Editores S.A de C.V. 2006. Capítulo 10 y Capítulo 14.

Instituto Nacional de Estadística (INE). Metodología para la integración de bases de datos de la encuesta nacional de empleo e ingresos. Guatemala, Instituto Nacional de Estadística. 2003. Pág. 1-10. Disponible en: www.cepal.org/deype/mecovi/docs/.../12.pdf

Taylor, S. y Bogdan, R. Introducción a los métodos cualitativos. . Tercera Edición. Barcelona, Ediciones Paidós Ibérica. 2000. Capítulo 6. Disponible en: <http://asodea.files.wordpress.com/2009/09/taylor-s-j-bogdan-r-metodologia-cualitativa.pdf>

Unidad 4. Categorización de la información e identificación de hallazgos significativos

Hernández, R, Fernández-Collado, C. y Baptista, P. Metodología de la Investigación. Cuarta edición. México DF. McGraw-Hill Interamericana Editores S.A de C.V. 2006. Capítulos 10, 14 y 16.

Gertler, P. y otros. La Evaluación de Impacto en la Práctica, Washington DC. The International Bank for Reconstruction and Development, 2011. Capítulo 13.

Salvarrey. L. Curso de Estadística Básica. Alto (Uruguay), Universidad de la Republica Regional Norte Sede Alto. 2000. Disponible en: <http://guajiros.udea.edu.co/descriptiva/articulos/Curso%20de%20EstadIstica%20Basica.pdf>

Taylor-Power, E. and Renner, M Analyzing Qualitative Data. Madison, University of Wisconsin- Extension. 2003.

Unidad 5. Preparación y redacción del informe de evaluación

Association for the Prevention of Torture (APT). Cómo hacer Recomendaciones Efectivas bajo la Metodología Double-Smart. Cuadernillo N°1. Noviembre 2008.

Banco Interamericano de Desarrollo (BID). Notas de Lecciones Aprendidas, publicado en Sector de Conocimiento y Aprendizaje, Octubre 2008. Banco Interamericano de Desarrollo. 2008. p. 1-4. Disponible en:
<http://boru.pbworks.com/f/Lecciones+Aprendidas+11-18-08.pdf>

Centro virtual de coordinación del conocimiento en VIH. Marco Conceptual sobre prácticas y lecciones aprendidas en VIH. Lima. Red SIDA Perú. p. 23-30. Disponible en: <http://www.cvcc-vih.com/comunicaciones/aids-2012-hub-peru/programacion-en-lima/tb-servicios-de-salud-resumen/item/283-marco-conceptual-sobre-pr%C3%A1cticas-y-lecciones-aprendidas>

Comisión Europea, Dirección General de Relaciones Exteriores. Metodología de Evaluación de la Ayuda Exterior de la Comisión Europea. Francia. Comisión Europea, 2006. Volumen I Bases Metodológicas de Evaluación. p. 88-91.

Comité de Desarrollo y Propiedad Intelectual (CDIP). Ejemplo Resumen de Informe de Evaluación Y Diseño de Conclusiones. Noviembre 2012.

Escorcía, O. Manual para la investigación. Guía para la formulación, desarrollo y divulgación de proyectos. Bogotá. Universidad Nacional de Colombia. 2010. p. 52-63. Disponible en:
<http://www.bdigital.unal.edu.co/7551/1/manualparalainvestigaci%C3%B3n.pdf>

Faúndez, A. Guía de Evaluación de Programas y Proyectos con Perspectiva de Género, Derechos Humanos e Interculturalidad. México. ONU Mujeres, 2014. p. 20-34.

Grade. Guía Práctica para Elaborar Conclusiones. Octubre 2015.

Grade. Guía Práctica para Elaborar Recomendaciones. Octubre 2015.

Jaramillo, Baanante & Sanz. Inform Final de Evaluación del Programa "Construyendo Perú" Ministerio de Economía y Finanzas. Lima, 2009.

Unidad 6. ¿Cómo evaluar el informe final de una evaluación?

Bamberger, M, Rugh, J y Mabry, L. Evaluación en un Mundo Real. Trabajando Bajo Restricciones Presupuestarias, Cronológicas, de Información y Políticas. Visión General. Adaptado del Capítulo 16, "Bringing it all together" de "Real World Evaluation: Working under Budget, Time, Data and Political Constraints", por Michael Bamberger, Jim Rugh y Linda Mabry. Sage Publications, 2006. p. 42-45.

Faúndez, A. Guía de Evaluación de programas y proyectos con perspectiva de género, derechos humanos e interculturalidad. México. ONU Mujeres, 2014. p. 16-20.

Sección 3.

Unidad 1. ¿Cómo elaborar términos de referencias para una evaluación?

Medina, A. Guía para la elaboración de Términos de Referencia Orientados a la Contratación de Servicios de Consultoría para la Evaluación de Programas y Proyectos Gubernamentales. Centro Latinoamericano de Administración para el Desarrollo. Banco Mundial. 2009. Disponible en:
<http://old.clad.org/documentos/otros-documentos/guia-para-la-elaboracion-de-terminos-de-referencia-orientados-a-la-contratacion-de-servicios-de-consultoria-para-el-monitoreo-y-evaluacion-de-programas-y-proyectos-gubernamentales>

Términos de Referencia Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Programa de Patrimonio Arqueológico. EDEPS MEF 2010.

Términos de Referencia Consultoría para la Evaluación de Diseño y Ejecución de Presupuestos Públicos. Proyecto Construyendo Perú. EDEPS MEF 2009.

ANEXOS

Anexo 1: Respuestas de preguntas de autoevaluación

SECCIÓN 1: Unidad 1

Pregunta	Respuesta
¿El monitoreo se halla incluido en cuál de los siguientes tipos de evaluación?	b
Estimar los beneficios sociales del programa en un estudio de preinversión es una típica...	a
¿Con cuál enfoque de evaluación relacionaría la siguiente afirmación: "el principal desafío para llevar a cabo evaluaciones eficaces de impacto es identificar la relación causal entre el proyecto, el programa o la política y los resultados de interés?"	c

SECCION 1: Unidad 2

Pregunta	Respuesta
1. ¿Cuáles son principales pasos para desarrollar un diseño de evaluación?	<p>La respuesta será correcta si presenta los once pasos del diseño.</p> <p>Existen once pasos necesarios para desarrollar el diseño de la evaluación:</p> <p>Paso 1. Diagnóstico del contexto.</p> <p>Paso 2. Identificación de los interesados</p> <p>Paso 3. Descripción del plan, programa o proyecto.</p> <p>Paso 4. Identificación de objetivos/cadenas de cambio.</p> <p>Paso 5. Definición de las preguntas de evaluación.</p> <p>Paso 6. Definición de los métodos: cuantitativos, cualitativos o mixtos.</p> <p>Paso 7. Identificación de las dimensiones, variables e indicadores a medir/verificar y las fuentes de información.</p> <p>Paso 8. Determinación del tamaño de efecto (impacto).</p> <p>Paso 9. Diseño de los instrumentos de recojo de información.</p> <p>Paso 10. Determinación de la muestra.</p> <p>Paso 11. Preparación del plan de recolección y análisis.</p>
2. ¿Cuáles son los criterios para definir qué tipo de evaluación se requiere realizar?	<p>La respuesta es correcta si al menos considera lo siguiente:</p> <ul style="list-style-type: none"> ➤ Resultados del PPP. ➤ Tiempos de la evaluación. ➤ La fase en la que se encuentra el proyecto. ➤ Momentos en los que se midieron los indicadores. ➤ Disponibilidad de información. ➤ Objetivos de la evaluación. ➤ Necesidad de información acerca del programa.

Pregunta	Respuesta
3. ¿Cómo se define a las partes interesadas en la evaluación y qué aspectos se deben considerar para evaluar los intereses de las partes?	<p>La respuesta será correcta, si presenta ideas clave relacionadas con lo siguiente:</p> <p>Las partes interesadas se definen mediante:</p> <ul style="list-style-type: none"> • La identificación de los principales usuarios de la evaluación: instituciones que financian el PPP, los patrocinadores de la evaluación, los tomadores de decisiones o gerentes del PPP, los gestores del programa y el personal en campo y/o los beneficiarios del programa. <p>Luego se identifica a los interesados y/o afectados por la evaluación.</p> <p>Para analizar los intereses de las partes se debe tener en cuenta:</p> <ul style="list-style-type: none"> • El rol del actor. • Las relaciones entre los actores. • Los puntos de acuerdo y de conflicto. • Las implicancias de la evaluación en corto y/o largo plazo.
4. ¿Cuáles son los principales aspectos/información del plan, programa o proyecto que deben conocerse para iniciar su evaluación?	<p>La respuesta será correcta, si presenta ideas clave que se muestran a continuación:</p> <p>Información necesaria para el inicio de la evaluación:</p> <ul style="list-style-type: none"> ➤ Información del diseño y planificación del PPP (documentos iniciales del PPP, documentos contractuales, modelo lógico del PPP). ➤ Información del proceso del PPP (informes de monitoreo, planes de trabajo). ➤ Información sobre evaluaciones realizadas (línea de base, evaluación de procesos, evaluación intermedia).
5. ¿Qué es una cadena de cambio y qué pasos se deben seguir para construirla en caso que el plan, programa o proyecto en evaluación no la tiene explícitamente definida?	<p>La respuesta será correcta si presenta ideas clave del siguiente contenido:</p> <p>La cadena de cambio es la esquematización de la Teoría del Cambio, que brinda una definición lógica de cómo la secuencia de insumos, actividades y productos relacionados directamente con el PPP interactúan y establecen las vías por las que se logran los impactos.</p> <p>Para la construcción de la cadena de cambio se debe clarificar con el personal responsable de gestionar el PPP, los resultados finales que contribuye el PPP, así como los resultados esperados del PPP, los productos, las actividades y los recursos que dispone el PPP.</p>
6. ¿Cuál es la principal pregunta en una evaluación de impacto y en una evaluación de desempeño?	<p>La respuesta será correcta, si presenta las ideas clave del siguiente contenido:</p> <p>A nivel de impacto: <i>¿Cuál es el impacto (o efecto causal) de un programa sobre un resultado de interés?</i></p> <p>A nivel de evaluación de desempeño: <i>¿Cuánto se ha logrado? ¿Cuán bien lo hemos hecho? ¿Cómo podríamos mejorar?</i></p>

Pregunta	Respuesta				
<p>7. Explica las diferencias entre los métodos cualitativos y cuantitativos y en qué momento se definen.</p>	<p>La respuesta será correcta, si por lo menos presenta 3 ideas clave de cada metodología:</p> <table border="1" data-bbox="623 365 1369 1020"> <thead> <tr> <th data-bbox="623 365 1005 403">Metodología cuantitativa</th> <th data-bbox="1005 365 1369 403">Metodología cualitativa</th> </tr> </thead> <tbody> <tr> <td data-bbox="623 403 1005 1020"> <ul style="list-style-type: none"> • Describen aspectos de la intervención en términos numéricos, cuantificables. • Predominan las preguntas cerradas. • Generan información numérica que favorecen las interpretaciones similares por los evaluadores. • Utilizan algún método estadístico (aleatorio) para la selección de la muestra. • Los resultados pueden generalizarse si la muestra es representativa. </td> <td data-bbox="1005 403 1369 1020"> <ul style="list-style-type: none"> • Capturan la realidad desde la perspectiva de las personas involucradas en el PPP. • Predominan las preguntas abiertas. • Generan información sobre valores, percepciones y creencias que pueden ser interpretadas de manera distinta por evaluadores. • Muestras limitadas con criterios de selección dirigidas a discreción del investigador. • Los resultados no pueden generalizarse, son indicativos. </td> </tr> </tbody> </table>	Metodología cuantitativa	Metodología cualitativa	<ul style="list-style-type: none"> • Describen aspectos de la intervención en términos numéricos, cuantificables. • Predominan las preguntas cerradas. • Generan información numérica que favorecen las interpretaciones similares por los evaluadores. • Utilizan algún método estadístico (aleatorio) para la selección de la muestra. • Los resultados pueden generalizarse si la muestra es representativa. 	<ul style="list-style-type: none"> • Capturan la realidad desde la perspectiva de las personas involucradas en el PPP. • Predominan las preguntas abiertas. • Generan información sobre valores, percepciones y creencias que pueden ser interpretadas de manera distinta por evaluadores. • Muestras limitadas con criterios de selección dirigidas a discreción del investigador. • Los resultados no pueden generalizarse, son indicativos.
Metodología cuantitativa	Metodología cualitativa				
<ul style="list-style-type: none"> • Describen aspectos de la intervención en términos numéricos, cuantificables. • Predominan las preguntas cerradas. • Generan información numérica que favorecen las interpretaciones similares por los evaluadores. • Utilizan algún método estadístico (aleatorio) para la selección de la muestra. • Los resultados pueden generalizarse si la muestra es representativa. 	<ul style="list-style-type: none"> • Capturan la realidad desde la perspectiva de las personas involucradas en el PPP. • Predominan las preguntas abiertas. • Generan información sobre valores, percepciones y creencias que pueden ser interpretadas de manera distinta por evaluadores. • Muestras limitadas con criterios de selección dirigidas a discreción del investigador. • Los resultados no pueden generalizarse, son indicativos. 				
<p>8. Presenta las características de las evaluaciones experimentales y cuasiexperimentales, en qué casos se utilizan y qué métodos utiliza.</p>	<p>La respuesta será correcta, si por lo menos presenta dos ideas clave de los tipos de evaluaciones:</p> <table border="1" data-bbox="613 1171 1390 1602"> <thead> <tr> <th data-bbox="613 1171 1008 1247">Evaluaciones experimentales</th> <th data-bbox="1008 1171 1390 1247">Evaluaciones cuasiexperimentales</th> </tr> </thead> <tbody> <tr> <td data-bbox="613 1247 1008 1602"> <ul style="list-style-type: none"> • Presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y control). • Se realiza en dos grupos al menos (tratamiento - control) • Se usa para probar la eficacia de las intervenciones en implementaciones masivas. </td> <td data-bbox="1008 1247 1390 1602"> <ul style="list-style-type: none"> • No presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y grupo control). • Son más rápidos y menos costosos de implementar. • Es un método menos sólido estadísticamente, por ello se reduce la confiabilidad de los resultados. </td> </tr> </tbody> </table>	Evaluaciones experimentales	Evaluaciones cuasiexperimentales	<ul style="list-style-type: none"> • Presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y control). • Se realiza en dos grupos al menos (tratamiento - control) • Se usa para probar la eficacia de las intervenciones en implementaciones masivas. 	<ul style="list-style-type: none"> • No presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y grupo control). • Son más rápidos y menos costosos de implementar. • Es un método menos sólido estadísticamente, por ello se reduce la confiabilidad de los resultados.
Evaluaciones experimentales	Evaluaciones cuasiexperimentales				
<ul style="list-style-type: none"> • Presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y control). • Se realiza en dos grupos al menos (tratamiento - control) • Se usa para probar la eficacia de las intervenciones en implementaciones masivas. 	<ul style="list-style-type: none"> • No presentan asignación aleatoria en la selección de los participantes (grupo tratamiento y grupo control). • Son más rápidos y menos costosos de implementar. • Es un método menos sólido estadísticamente, por ello se reduce la confiabilidad de los resultados. 				
<p>9. ¿Qué es un indicador y cuáles son las principales características de un buen indicador de desarrollo?</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <p>Indicador: Es la medición que sirve para describir un comportamiento, concepto o fenómeno. En la evaluación es importante ya que a través de la información recopilada sobre ellos que el (los) evaluador (res) podrá(n) emitir juicios sobre los avances y logros de un PPP.</p>				

Pregunta	Respuesta
	<p>Características de un buen indicador de desarrollo:</p> <p>Deben tener/ ser:</p> <ul style="list-style-type: none"> - Medidas directas. - Capaces de mostrar variaciones de un grupo a otro. - Relación directa con las intervenciones. - Confiables. - Incluir una definición completa. - Consistentes con la fuente de datos existentes. - Costo-efectivos.
10. ¿Cuáles son los principales contenidos de un plan de recolección de datos?	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <p>Un plan de recolección de datos debe contener:</p> <ul style="list-style-type: none"> ➤ Objetivos. ➤ Características y conceptos técnicos de la encuesta. 1. Población en estudio. 2. Unidad de investigación. 3. Método de recolección. 4. El proceso de la encuesta. 5. Definición de variables e indicadores para la encuesta. 6. Verificación de campo. ➤ Organización del trabajo. ➤ Funciones del personal de campo. ➤ Tareas de trabajo de campo. ➤ Programación de rutas de la operación de campo. ➤ Plan de contingencias.

SECCIÓN 2: Unidad 1

Pregunta	Respuesta
1. ¿Por qué es importante la prueba piloto en un estudio de evaluación?	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <ul style="list-style-type: none"> • La prueba piloto permite identificar las principales fuentes de donde pueden provenir los errores en la ejecución de las herramientas de recojo de información. Por lo general, estas fuentes son: a) el ambiente de recolección, b) la situación de los entrevistados y c) los responsables de recojo de información. • La prueba piloto permite realizar los ajustes necesarios para llevar de manera efectiva la recolección de los datos.

Pregunta	Respuesta
<p>2. ¿Cómo se determina la validez y confiabilidad de los instrumentos luego de la aplicación de la prueba piloto?</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <p>La implementación de la prueba piloto permite poner a prueba el instrumental con el que se cuenta y evaluar su idoneidad para responder a los objetivos de la evaluación. La validez se determinará si el instrumento logra medir lo que pretende. La confiabilidad se establecerá en la medida de la relación de los ítems relacionados entre sí, es decir la calidad del instrumento.</p>
<p>3. Describe las principales etapas del proceso de la prueba piloto</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <ol style="list-style-type: none"> 1. Planificación y organización. 2. Ejecución. <ul style="list-style-type: none"> • Aplicación e instrumentos. • Reporte del trabajo de campo: incluye (i) información acerca del proceso de recojo de datos (ii) información acerca de los instrumentos. 3. Ajustes a los instrumentos. 4. Elaboración del informe.
<p>4. ¿Qué aspectos de la implementación de la prueba piloto no deben dejar de documentarse?</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <p>Sobre el lugar de estudio:</p> <ul style="list-style-type: none"> • Ambiente de aplicación (comodidad, privacidad, seguridad). • Accesibilidad. <p>Sobre la aplicación de los instrumentos:</p> <ul style="list-style-type: none"> • Claridad y comprensión de las preguntas. • Tiempos de aplicación. • Adecuación cultural. <p>Sobre los instrumentos propiamente:</p> <ul style="list-style-type: none"> • Orden de las preguntas, opciones y tipos. • Tiempos. <p>Coordinación y supervisión:</p> <ul style="list-style-type: none"> • Coordinación de las instrucciones. • Secuencia lógica. • Claridad de los participantes en cuanto a tareas y responsabilidades.

SECCIÓN 2: Unidad 2

Pregunta	Respuesta
<p>1. Suponiendo que tienes a cargo la capacitación a encuestadores para la aplicación de una encuesta a hogares en el marco de una evaluación de impacto, presenta las indicaciones generales que debes dar a los encuestadores.</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <ul style="list-style-type: none"> ➤ Indicar con claridad el propósito de la evaluación de impacto. ➤ Resaltar la importancia de la participación de los encuestadores y el rol que tienen en el estudio. ➤ Destacar la confidencialidad que debe guardar para el desarrollo de su labor. ➤ Acentuar en la metodología de recolección de datos, las características del encuestados y del hogar según protocolo establecido ➤ Hacer hincapié en la forma de relacionarse con los participantes, mostrando respeto y sensibilidad que deben tener hacia los participantes evitando discusiones y/o valoraciones. ➤ Señalar el procedimiento de presentación al encuestado (el uso de las credenciales y cartas de identificación).
<p>2. Explica las principales etapas para la aplicación de instrumentos cuantitativos.</p>	<ul style="list-style-type: none"> • Selección y capacitación del personal encuestador. En esta etapa se requiere capacitar al personal sobre la metodología a utilizarse de acuerdo al tipo de muestreo, así como también las variables e indicadores clave del estudio. • Favorecer el ambiente para la aplicación. • Presentar los objetivos del estudio. • Aplicación de las preguntas de forma clara. • Registro de las respuestas (completo, preciso y legible). • Agradecer al encuestado para finalizar la encuesta. • Procesar los datos y establecer el valor de los indicadores.
<p>3. ¿Cuáles son las principales normas que se debe considerar en la aplicación de cuestionarios mediante la entrevista?</p>	<ul style="list-style-type: none"> • Ajustarse a los protocolos y normas del manual de campo recibido. • Estudiar los temas, subtemas y preguntas de los cuestionarios. • Buscar un espacio favorable a la aplicación. • Presentar el estudio, aclarar la confidencialidad y pedir permiso para grabar. • Ser amable y empático con el encuestado. • Hablar de forma clara. • Respetar el orden de las preguntas.

Pregunta	Respuesta
	<ul style="list-style-type: none"> • Orientar la entrevista a los temas del cuestionario. Usar el lenguaje adecuado (contextualizado) según sea entrevistado • No sugerir u orientar respuestas. • Anotar claramente las ideas principales de las respuestas. • Hacer todas las preguntas. • Reservar la confidencialidad del cuestionario.
<p>4. ¿Cuáles son los principales aspectos que se deben considerar durante la aplicación de un grupo focal?</p>	<p>La respuesta será correcta, si presenta al menos cuatro ideas clave que se muestran a continuación:</p> <ul style="list-style-type: none"> • La determinación del número de grupos y el tamaño de cada grupo. • Los criterios de selección de los participantes. • La elección del moderador. • La coordinación previa a nivel de disponibilidad y tiempo asignado para realizar el grupo focal. • La cobertura de los temas relevante según tiempo disponible. • La grabación del grupo focal teniendo en cuenta la aceptación de las mismas por parte de los participantes. • La promoción de la interacción de todos los participantes.
<p>5. Explica el objetivo de la actividad de control antes, durante y posterior al recojo de información.</p>	<p>El objetivo es asegurar la calidad de los datos en los tres momentos.</p> <p>Antes del recojo de información: se analizan la validez y confiabilidad de los instrumentos a partir de los resultados de la prueba piloto.</p> <p>Durante el recojo de información: se supervisa la calidad de datos.</p> <p>Después del recojo de información: Se evalúa la consistencia de las respuestas y/o evalúa el nivel de pérdida de datos por cada pregunta.</p>
<p>6. Describe las principales actividades del proceso de control de calidad en la recolección de datos.</p>	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <p>Las principales actividades son:</p> <ul style="list-style-type: none"> • <i>Control de desempeño.</i> Consiste en la revisión de la tarea del encuestador en cuanto a diligenciamiento oportuno de los formularios en cada uno de los encuestados y al logro de las entrevistas.

Pregunta	Respuesta
	<ul style="list-style-type: none"> • <i>Control de calidad.</i> El objetivo es orientar a la verificación de la calidad y consistencia de la información recolectada por los encuestadores. • <i>Control de visitas.</i> Tiene como finalidad la verificación de la concurrencia directa y personal del encuestador a los encuestados que debió visitar y encuestar. Además, que cumpla con las visitas concertadas para ubicar al informante indicado. • <i>Control de cobertura.</i> Se realiza conjuntamente con el control de visitas y tiene como propósito comprobar que las unidades encuestadas correspondan exactamente a las seleccionadas en la muestra. • <i>Evaluación de la entrevista.</i> Tiene como objetivo verificar el manejo del cuestionario por parte del encuestador. • <i>Revisión de cuestionarios.</i> Implica examinar en gabinete todos los formularios diligenciados para detectar y corregir en el mismo terreno, si fuera necesario, las inconsistencias, omisiones, legibilidad y grado de detalle de ciertas preguntas.

SECCIÓN 2: Unidad 3

Pregunta	Respuesta
1. Explica brevemente los pasos para el diseño de una base de datos.	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <ul style="list-style-type: none"> .1.1 Identificar y describir todas las variables y categorías de cada una de las variables de la encuesta a procesar. .1.2 Definir los parámetros de la matriz de dato en el programa (ítem, reactivo, categoría, subcategoría, tipo de variable, ancho de dígitos). .1.3 Codificar (si el caso lo requiera). .1.4 Captura de los datos mediante la digitación.
2. Explica los criterios para la revisión de la consistencia de datos.	<p>La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación:</p> <ul style="list-style-type: none"> - Cuando se observa una concentración de ausencia de respuestas en algunos segmentos poblacionales. - Cuando se observen datos faltantes o valores o códigos distintos a los registrados en el cuestionario.

	- Cuando se observen inconsistencias anormales en los datos.
3. ¿Cómo se determina la precisión estadística de los estimadores?	La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación: Al establecer el error relativo estimado o coeficiente de variación.
4. Describe los principales pasos durante la etapa de descubrimiento y codificación para el procesamiento de datos cualitativos.	La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación: <ul style="list-style-type: none"> • Descubrimiento: identificar temas y desarrollar conceptos y proposiciones. • Codificación: de los datos y el refinamiento de la comprensión del tema de estudio. • Uso de programas: específicos para el análisis de datos cualitativos • Relativización de los datos: el evaluador trata de comprender los datos en el contexto en que fueron recogidos.

SECCIÓN 2: Unidad 4

Pregunta	Respuesta
1. ¿Cuáles son los supuestos de partida del análisis paramétrico?	La respuesta será correcta si presenta las ideas clave que se muestran a continuación: <ul style="list-style-type: none"> ➤ La distribución poblacional de la variable dependiente es normal: el universo tiene una distribución normal. ➤ El nivel de medición de la variable es dependiente es por intervalos o razón. ➤ Cuando dos o más poblaciones son estudiadas, tienen una varianza homogénea: las poblaciones en cuestión poseen una dispersión similar en sus distribuciones.
2. Indica las principales características del análisis cualitativo.	La respuesta será correcta si presenta todas las ideas clave que se muestran a continuación: <p>El análisis cualitativo:</p> <ul style="list-style-type: none"> • Recolecta datos no estructurados. • Se analiza mediante el desarrollo de categorías a partir de los datos brutos dentro de un marco. • Los hallazgos son modelados por las suposiciones y las experiencias de los evaluadores.

SECCIÓN 2: Unidad 5

Afirmación	Respuesta
1. Las conclusiones derivan directamente de los datos	Falso
2. Las conclusiones deben hacer referencia a las preguntas de la evaluación	Verdadero
3. Las conclusiones se formulan tomando en cuenta las implicancias que tienen sobre los interesados, en especial los beneficiarios.	Verdadero
4. La elaboración de recomendaciones solo toma en cuenta los intereses de los financiadores y/o equipo de gestión del programa o proyecto.	Falso
1. Una lista de cotejo para evaluar el diseño es útil para los que practican la evaluación, así como para aquellos que la gestionan.	Verdadero
2. Se puede deducir del cuadro N° 21 que la elección de los métodos a elegir depende de las preguntas de evaluación.	Falso
3. Un informe de evaluación debe ser objetivo, es decir, evitar mostrar los valores e intereses de los involucrados en una evaluación.	Falso
4. El informe de evaluación debe ser riguroso pero entendible por los interesados en la evaluación.	Verdadero

SECCIÓN 3: Unidad 1

Pregunta	Respuesta
Los costos estimados de una evaluación se deben establecer sobre la base de costos de mercado.	Verdadero
El diseño de la evaluación es una etapa previa a la confección de TDR	Verdadero
Los TDR de una evaluación deben señalar los responsables de darles seguimiento.	Verdadero
Los TDR de una evaluación deben indicar los tiempos de revisión de los productos.	Verdadero

Anexo 2: Glosario de términos unificado de la OECD & Normatividad vigente del sector público del Perú

-A-

ACTIVIDADES. Acciones emprendidas o labor realizada mediante las cuales se movilizan los insumos, como son los fondos, la asistencia técnica y otros tipos de recursos, para generar productos determinados.

Término conexo: intervención para el desarrollo (OECD, 2002).

Es una acción sobre una lista específica y completa de insumos (bienes y servicios necesarios y suficientes), que en conjunto con otras actividades garantizan la provisión del producto. Se debe considerar que la actividad deberá ser relevante y presupuestable (MEF, 2014).

Las actividades son las tareas que se realizan en un programa o proyecto con el fin de obtener los productos previstos en el plan. En general, para obtener los productos previstos en el plan. En general, para obtener un producto es necesario realizar varias actividades. En el concepto actividad, la palabra clave es hacer (MIMP, 2011).

Conjunto de tareas que deben ejecutarse para el cumplimiento de los procesos internos que a su vez apuntan al logro de los objetivos institucionales (CEPLAN, 2010)

ANÁLISIS DE RIESGOS. Análisis o evaluación de factores (llamados « supuestos » en el marco lógico) que afectan o que pueden afectar el logro de los objetivos de una intervención. Examen detallado de las consecuencias indeseables y negativas que podría tener una intervención para el desarrollo en la vida humana, la salud, la propiedad o el medio ambiente; proceso sistemático que proporciona información sobre esas consecuencias indeseables; proceso de cuantificación de las probabilidades de que aparezcan los riesgos identificados y sus repercusiones previstas (OECD, 2002).

ASOCIADOS. Particulares y/u organismos que colaboran para lograr los objetivos mutuamente acordados. Nota: El concepto de asociación denota metas compartidas, responsabilidad común por los efectos directos, clara rendición de cuentas y obligaciones recíprocas. Los asociados pueden ser gobiernos, entidades de la sociedad civil, organizaciones no gubernamentales, universidades, asociaciones profesionales y empresariales, organismos multilaterales, empresas privadas, etc. (OECD, 2002).

ATRIBUCIÓN. Imputación de un vínculo causal entre cambios observados (o que se espera observar) y una intervención específica.

Nota: La atribución se refiere a la acción que explica los cambios observados o los resultados obtenidos. Representa la medida en que los efectos de desarrollo observados pueden atribuirse a una intervención específica o a la actuación de una o más partes teniendo en cuenta otras intervenciones, factores exógenos (previstos o imprevistos) o choques externos (OECD, 2002).

AUTOEVALUACIÓN. Evaluación por parte de los responsables del diseño y la implementación de una intervención para el desarrollo (OECD, 2002).

AUDITORÍA. Actividad de control independiente y objetivo que tiene como fin añadir valor y mejorar las operaciones de una organización. Ayuda a la organización a lograr sus objetivos aportando un enfoque sistemático y disciplinado para evaluar y mejorar la eficacia de la gestión de riesgos, el control y los procesos de dirección. Nota: Se debe distinguir entre la auditoría de regularidad (financiera), que se centra en el cumplimiento de los estatutos y regulaciones aplicables, y la auditoría de resultados, que se ocupa de la pertinencia, la economía, la eficiencia y la eficacia. La auditoría interna constituye una evaluación de los controles internos realizados por una unidad que rinde cuentas a la dirección, en tanto que la auditoría externa es realizada por una entidad independiente (OECD, 2002).

-B-

BENEFICIARIO. Individuos, grupos u organismos que se benefician, directa o indirectamente, de una intervención para el desarrollo, hayan sido o no los destinatarios de la intervención.

Términos conexos: destinatarios, grupo meta (OECD, 2002).

-C-

CADENA DE RESULTADOS. La secuencia causal de una intervención para el desarrollo que estipula la secuencia necesaria para lograr los objetivos deseados, comenzando con los insumos, pasando por las actividades y los productos, y culminando en el efecto directo, el impacto y la retroalimentación. En algunos organismos, los destinatarios (o el « alcance ») forman parte de la cadena de resultados.

Términos conexos: supuestos, marco de resultados (OECD, 2002).

CALIDAD. Mide la capacidad de la intervención para responder en forma rápida y directa a las necesidades de la población objetivo. Para esta medición se puede utilizar como referencia estándares de entrega de bienes y servicios. Son extensiones de la calidad: oportunidad, accesibilidad, precisión y continuidad en la entrega de los servicios, satisfacción del usuario. Cuando mide la capacidad de la intervención para responder a las necesidades de su población objetivo (MEF, 2014).

Mide los atributos de los bienes o servicios, en función de un estándar previamente establecido (MIDIS, 2012).

CALIDAD DE INFORMACIÓN. Se define como las características que debe tener la información, entre las que se encuentran: Exactitud, oportunidad, que sea completa, confiable, relevante, precisa y consistente (MIDIS, 2013).

CONCLUSIONES. Las conclusiones señalan los factores de éxito y fracaso de la intervención evaluada, prestando atención especial a los resultados y repercusiones intencionales o no y, de manera más general, a otras fortalezas y debilidades. Una conclusión se apoya en los datos recopilados y en los análisis realizados mediante una cadena transparente de enunciados (OECD, 2002).

CONTRAFACTUAL. Situación o condición que podría presentarse hipotéticamente para las personas, las organizaciones o los grupos si no hubiera una intervención para el desarrollo como la evaluada (OECD, 2002).

-D-

DESEMPEÑO. Medida en que una intervención para el desarrollo o una entidad que se ocupa de fomentar el desarrollo actúa conforme a criterios/normas/directrices específicos u obtiene resultados de conformidad con metas o planes establecidos (OECD, 2002).

DESTINATARIOS (O « ALCANCE »). Beneficiarios y otras partes interesadas de una intervención para el desarrollo. Término conexo: beneficiarios (OECD, 2002).

-E-

ECONOMÍA. Ausencia de desperdicio en la obtención de un resultado determinado. Nota: Se dice que una actividad es económica cuando los costos de los recursos escasos utilizados se aproximan al mínimo necesario para lograr los objetivos planeados (OECD, 2002).

EFFECTIVIDAD. Representa la contribución realizada por los resultados del PPP para el logro de los propósitos del PPP (OECD, 2002).

EFFECTO. Cambio intencional o no intencional debido directa o indirectamente a una intervención. Términos conexos: resultados, efecto directo (OECD, 2002).

EFFECTO DIRECTO. Representa el conjunto de resultados a corto y mediano plazo probables o logrados por los productos de una intervención. Términos conexos: resultados, producto, impacto, efecto (OECD, 2002).

EFICACIA. Medida en que se lograron o se espera lograr los objetivos de la intervención para el desarrollo, tomando en cuenta su importancia relativa. Nota: Se utiliza también como medida agregada o juicio sobre el mérito o el valor de una actividad, es decir, el grado al cual una intervención ha logrado, o se espera que logre, sus principales objetivos pertinentes de manera eficaz, en forma sostenible, y con un impacto institucional positivo en el desarrollo. Término conexo: eficiencia, efectividad (OECD, 2002).

Medida en que se lograron o se espera lograr los objetivos de la intervención para el desarrollo, tomando en cuenta su importancia relativa (MIDIS, 2012).

EFICIENCIA. Medida en que los recursos/insumos (fondos, tiempo, etc.) se han convertido económicamente en resultados (OECD, 2002).

ENSEÑANZAS APRENDIDAS. Generalizaciones basadas en las experiencias de evaluación de proyectos, programas o políticas en circunstancias específicas, que se aplican a situaciones más amplias. Con frecuencia, las enseñanzas destacan los puntos fuertes o débiles en la preparación, el diseño y la puesta en práctica que afectan al desempeño, los resultados y el impacto (OECD, 2002).

ESTUDIO DE BASE. Análisis que describe la situación previa a una intervención para el desarrollo, en relación con la cual puede medirse el avance o pueden efectuarse comparaciones (OECD, 2002).

ESTUDIO DEFINITIVO. Estudio que permite definir a detalle la alternativa seleccionada en el nivel de pre inversión y calificada como viable. Para su elaboración se deben realizar estudios especializados que permitan definir: el dimensionamiento a detalle del proyecto, los costos unitarios por componentes, especificaciones técnicas para la ejecución de obras o equipamiento, medidas de mitigación de impactos ambientales negativos, necesidades de operación y mantenimiento, el plan de implementación, entre otros requerimientos considerados como necesarios de acuerdo a la tipología del proyecto. En proyectos de infraestructura, a los estudios especializados se les denomina de ingeniería de detalle (topografía, estudios de suelos, etc.) Los contenidos de los Estudios Definitivos varían con el tipo de proyecto y son establecidos de acuerdo

con la reglamentación sectorial vigente y los requisitos señalados por la Unidad Formuladora y/o Unidad Ejecutora del Proyecto (MEF, 2011).

EVALUABILIDAD. Medida en que puede evaluarse una actividad o un programa de manera fiable y creíble. Nota: Para determinar la evaluabilidad es necesario un examen anticipado de una evaluación propuesta para determinar si sus objetivos están definidos adecuadamente y si sus resultados pueden ser objeto de verificación (OECD, 2002).

EVALUACIÓN. Apreciación sistemática y objetiva de un proyecto, programa o política en curso o concluido, de su diseño, su puesta en práctica y sus resultados. El objetivo es determinar la pertinencia y el logro de los objetivos, así como la eficiencia, la eficacia, el impacto y la sostenibilidad para el desarrollo. Una evaluación deberá proporcionar información creíble y útil, que permita incorporar las enseñanzas aprendidas en el proceso de toma de decisiones de beneficiarios y donantes. La evaluación también se refiere al proceso de determinar el valor o la significación de una actividad, política o programa. Se trata de una apreciación, tan sistemática y objetiva como sea posible, de una intervención para el desarrollo planeada, en curso o concluida. Nota: En algunos casos, la evaluación entraña la definición de normas adecuadas, el examen de los resultados en función de esas normas, una apreciación de los resultados efectivos y previstos y la identificación de lecciones pertinentes.

Término conexo: revisión (OECD, 2002).

EVALUACIÓN CONJUNTA. Evaluación en la que participan distintas entidades donantes y/o asociados. Nota: Existen varios grados de « asociación », dependiendo de la medida en que cada parte coopera en el proceso de evaluación, y combina sus recursos de evaluación y sus mecanismos de declaración de información. Las evaluaciones conjuntas pueden ayudar a superar problemas de atribución al evaluar la eficacia de los programas y las estrategias, el carácter complementario de los esfuerzos respaldados por diferentes asociados, la calidad de la coordinación de la asistencia, etc. (OECD, 2002).

EVALUACIÓN DE CONJUNTO. Evaluación de un conjunto de actividades, proyectos y/o programas relacionados entre sí (OECD, 2002).

EVALUACIÓN DE PROCESOS. Evaluación de la dinámica interna de los organismos de ejecución, sus instrumentos de política, sus mecanismos de prestación de servicios, sus procedimientos de gestión y los vínculos que existen entre todos estos componentes.

Término conexo: evaluación formativa (OECD, 2002).

EVALUACIÓN DE PROGRAMAS. Evaluación de un conjunto de intervenciones estructuradas para alcanzar objetivos específicos de desarrollo a nivel global, regional, de país o sectorial. Nota: Un programa de desarrollo es una intervención de duración limitada que comprende múltiples actividades que pueden abarcar varios sectores, temas y/o zonas geográficas.

Términos conexos: evaluación del programa/ estrategia del país (OECD, 2002).

EVALUACIÓN DE PROYECTOS. Evaluación de una intervención individual para el desarrollo que tiene por objeto alcanzar objetivos específicos con recursos especificados y dentro de calendarios de ejecución determinados, a menudo enmarcada en un programa más amplio. Nota: El análisis de costos y beneficios es uno de los principales instrumentos de evaluación de proyectos con beneficios mensurables. Cuando no puedan cuantificarse los beneficios, el enfoque adecuado sería el de eficacia en función de los costos (OECD, 2002).

EVALUACIÓN DEL PROGRAMA DE PAÍS/ EVALUACIÓN DE LA ASISTENCIA AL PAÍS. Evaluación de la cartera de intervenciones para el desarrollo de uno o más donantes o instituciones, y de la estrategia de asistencia que la sustenta, en un país determinado (OECD, 2002).

EVALUACIÓN DE PROGRAMAS SECTORIALES. Evaluación de un conjunto de intervenciones para el desarrollo dentro de un país o un grupo de países en un sector, todas las cuales contribuyen al logro de un objetivo específico de desarrollo. Nota: Un sector incluye actividades de desarrollo que se agrupan comúnmente para fines de acción pública, tales como la salud, la educación, la explotación agropecuaria, el transporte, etc. (OECD, 2002).

EVALUACIÓN DE TÉRMINO MEDIO. Evaluación que se efectúa a mediados del periodo de implementación de la intervención.

Término conexo: evaluación formativa (OECD, 2002).

EVALUACIÓN EX ANTE. Determinación general de la pertinencia, la factibilidad y la sostenibilidad potencial de una intervención para el desarrollo antes de tomar la decisión de otorgar financiamiento. Nota: En los organismos, bancos, etc. de desarrollo, el propósito de la evaluación ex ante es presentar información durante el proceso de toma de decisiones para determinar si la actividad representa el uso adecuado de los recursos de la institución.

Término conexo: pre-evaluación (OECD, 2002).

EVALUACIÓN EX POST. Evaluación de una intervención para el desarrollo que se efectúa una vez que ha concluido. Nota: Puede emprenderse inmediatamente o mucho después de su conclusión. La intención es identificar los factores de éxito o fracaso, evaluar la sostenibilidad de los resultados y las repercusiones, y extraer conclusiones que puedan ser útiles para otras intervenciones (OECD, 2002).

EVALUACIÓN EXTERNA. Evaluación de una intervención para el desarrollo realizada por entidades y/o individuos ajenos a los organismos donantes y de ejecución (OECD, 2002).

EVALUACIÓN INDEPENDIENTE. Evaluación realizada por entidades y personas que no están bajo el control de los responsables del diseño y la implementación de la intervención para el desarrollo. Nota: La credibilidad de una evaluación depende en parte del grado de independencia con la que se llevó a cabo. La independencia significa que la evaluación no está sujeta a ninguna influencia política ni a presiones de ninguna organización. Se cuenta con pleno acceso a la información y total autonomía para llevar a cabo las investigaciones y notificar lo que se haya determinado (OECD, 2002).

EVALUACIÓN INTERNA. Evaluación de una intervención para el desarrollo realizada por una unidad y/o personas que dependen de la gerencia o dirección de un organismo donante, asociado o de ejecución.

Término conexo: autoevaluación (OECD, 2002).

EVALUACIÓN FORMATIVA. Evaluación cuyo objeto es mejorar el desempeño y que generalmente se realiza durante la fase de implementación de un proyecto o programa. Nota: También pueden realizarse evaluaciones formativas por otras razones como la verificación del cumplimiento, por exigencias legales o como parte de una iniciativa de evaluación de mayor envergadura.

Término conexo: evaluación de proceso (OECD, 2002).

EVALUACIÓN PARTICIPATIVA. Método de evaluación en la que los representantes de entidades y partes interesadas (incluidos los beneficiarios) trabajan conjuntamente en el diseño, implementación e interpretación de una evaluación (OECD, 2002).

EVALUACIÓN SUMATIVA. Estudio que se realiza al final de una intervención (o de una fase de esa intervención) para determinar en qué medida se produjeron los resultados previstos. El objetivo de la evaluación sumativa es proporcionar información sobre el valor del programa.

Término conexo: evaluación del impacto (OECD, 2002).

EVALUACIÓN TEMÁTICA. Evaluación de un conjunto determinado de intervenciones para el desarrollo, todas las cuales se refieren a una prioridad específica de desarrollo que abarca distintos países, regiones y sectores (OECD, 2002).

EVIDENCIA. Documentación científica que muestre la validez y consistencia en las relaciones de causalidad y eficacia identificadas (MEF, 2014).

-F-

FIABILIDAD. Es la coherencia de los datos y los juicios de evaluación, y la medida en que se puede confiar en ellos, con referencia a la calidad de los instrumentos, procedimientos y análisis empleados para recopilar e interpretar

datos de evaluación. Nota: La información de una evaluación es fiable si observaciones repetidas utilizando instrumentos similares, en condiciones similares dan lugar a resultados similares (OECD, 2002).

-G-

GARANTÍA DE CALIDAD. La garantía de calidad abarca toda actividad que tenga por objeto evaluar y mejorar el mérito o el valor de una intervención para el desarrollo o su cumplimiento con normas establecidas. Nota: Cabe citar como ejemplos de actividades de garantía de calidad la evaluación *ex ante*, la gestión basada en resultados, los exámenes durante la ejecución, las evaluaciones, etc. La garantía de calidad también puede referirse a la evaluación de la calidad de una cartera y de su eficacia en términos del desarrollo (OECD, 2002).

GESTIÓN BASADA EN RESULTADOS. Estrategia de gestión que se centra en el desempeño y el logro de productos, efectos e impacto.
Término conexo: marco lógico (OECD, 2002).

GESTIÓN DEL CONOCIMIENTO. Es una disciplina emergente que tiene como objetivo generar, compartir y utilizar el conocimiento tácito (know-how) y explícito (formal) existente en un determinado colectivo u organización, para dar respuestas a las necesidades de los individuos y de las comunidades en su desarrollo. El objetivo es administrar conocimiento y los aprendizajes organizacionales para mejorar el funcionamiento de las entidades, tomando en cuenta buenas prácticas propias o de terceros para retroalimentar el diseño e implementación de sus estrategias de acción y asegurar así resultados positivos y relevantes. Cuando la gestión del conocimiento se implementa formalmente en el sector público, los sistemas se vuelven cada vez más interconectados, los procesos se hacen más visibles y dinámicos, se pueden optimizar los recursos y mejora la transparencia en el manejo de los asuntos públicos (PCM, 2013).

GESTIÓN DE LA EVIDENCIA. Conjunto de procesos que promueven que las evidencias y las recomendaciones de las evaluaciones y reportes de seguimiento

sean debidamente utilizadas y oportunamente implementadas, y que sus resultados sean vinculados a actividades futuras (MIDIS, 2012).

GRUPO META. Individuos u organismos específicos en cuyo beneficio se lleva a cabo la intervención para el desarrollo (OECD, 2002).

-H-

HALLAZGO (CONSTATACIÓN). Un hallazgo (o una constatación) utiliza evidencias obtenidas de una o más evaluaciones para realizar afirmaciones basadas en hechos (OECD, 2002).

HERRAMIENTAS DE RECOLECCIÓN DE DATOS: Metodologías empleadas para identificar fuentes de información y recopilar datos durante una evaluación. (OECD, 2002).

Nota: Cabe citar como ejemplos las encuestas informales y formales, la observación directa y participativa, las entrevistas en comunidades, los grupos de control, la opinión de expertos, los estudios de casos y la investigación bibliográfica (OECD, 2002).

-I-

INDICADOR. Variable o factor cuantitativo o cualitativo que proporciona un medio sencillo y fiable para medir logros, reflejar los cambios vinculados con una intervención o ayudar a evaluar los resultados de un organismo de desarrollo. (OECD, 2002).

Es un enunciado que permite medir el estado de cumplimiento de un objetivo, facilitando su seguimiento (CEPLAN, 2014).

Enunciado que define una medida cuantitativa sobre el nivel de logro de resultados, productos y/o actividades, vinculados a las políticas, planes programas o proyectos (MIDIS, 2012).

INDICADOR DE DESEMPEÑO. Enunciado que define una medida sobre el logro de resultados, la entrega de productos o la satisfacción de los atributos del producto en las dimensiones de eficacia, eficiencias, calidad y/o economía (MIDIS, 2012).

INSUMOS. Recursos financieros, humanos y materiales empleados en una intervención para el desarrollo (OECD, 2002).

IMPACTOS. Efectos de largo plazo positivo y negativo, primario y secundario, producidos directa o indirectamente por una intervención para el desarrollo, intencionalmente o no (OECD, 2002).

IMPACTO EN EL DESARROLLO INSTITUCIONAL. Medida en que una intervención mejora o debilita la capacidad de un país o región de hacer uso más eficiente, equitativo y sostenible de sus recursos humanos, financieros y naturales, por ejemplo, a través de: (a) mecanismos institucionales mejor definidos, más estables, transparentes y aplicados de manera eficaz y previsible y/o (b) mejor ajuste de la misión y la capacidad de una organización con su mandato, que se deriva de estos mecanismos institucionales. Estos impactos pueden incluir los efectos intencionales o no intencionales de una acción (OECD, 2002).

INDICADOR DEL DESEMPEÑO (O DE LOS RESULTADOS). Variable que permite verificar cambios debidos a la intervención para el desarrollo o que muestra resultados en relación con lo que se ha planeado

Términos conexos: seguimiento del desempeño, medición de los resultados (OECD, 2002)

INTERVENCIÓN PARA EL DESARROLLO. Instrumento de asistencia utilizado por un organismo (donante o no donante), con el objeto de fomentar el desarrollo. Nota: Cabe citar como ejemplos el asesoramiento en materia de políticas, los proyectos y los programas (OECD, 2002).

-M-

MARCO LÓGICO. Herramienta que se utiliza para mejorar el diseño de las intervenciones, más frecuentemente a nivel del proyecto. Comprende la

identificación de elementos estratégicos (insumos, productos, efectos, impacto) y sus relaciones causales, indicadores y los supuestos o riesgos que pueden influir en el éxito o el fracaso. De esa manera facilita la planeación, la ejecución y la evaluación de una intervención para el desarrollo. Término conexo: gestión basada en resultados (OECD, 2002).

MARCO DE RESULTADOS. Lógica del programa que explica la forma en que deberá lograrse el objetivo de desarrollo, que incluye las relaciones causales y los supuestos subyacentes.

Términos conexos: cadena de resultados, marco lógico. (OECD, 2002).

MEDICIÓN DE LOS RESULTADOS. Sistema para evaluar los resultados de las intervenciones para el desarrollo en función de los objetivos declarados. Términos conexos: seguimiento del desempeño, indicador (OECD, 2002).

MEDIOS DE VERIFICACIÓN. Son los mecanismos o herramientas utilizados para la medición de los indicadores y están constituidos por las fuentes de los datos y los procedimientos de cálculo. Estos deben cumplir con las siguientes características: confiabilidad, oportunidad, transparencia y replicabilidad, costo – eficiencia, pertinencia y validez.

META. Objetivo global hacia el cual se prevé que contribuya una intervención para el desarrollo.

Término conexo: objetivo de desarrollo (OECD, 2002).

Es el valor proyectado del indicador para hacer seguimiento al logro de los objetivos estratégicos (CEPLAN, 2014).

META-EVALUACIÓN. Este término se utiliza para evaluaciones cuyo objeto es sintetizar constataciones de un conjunto de evaluaciones. También puede utilizarse para indicar la evaluación de otra evaluación a fin de juzgar su calidad y/o medir el desempeño de los evaluadores (OECD, 2002).

-N-

NORMA DE REFERENCIA. Base de comparación o norma (standard) en relación con la cual pueden medirse los resultados o los logros obtenidos.

Nota: Una norma se refiere a los resultados logrados en el pasado reciente por otros organismos similares, o a lo que razonablemente puede inferirse que se ha logrado en circunstancias similares (OECD, 2002).

-O-

OBJETIVO DE DESARROLLO. Impacto intencional esperado que contribuye a generar beneficios físicos, financieros, institucionales, sociales, ambientales o de otra índole a una sociedad, una comunidad o un grupo de personas mediante una o más intervenciones para el desarrollo (OECD, 2002).

OBJETIVO DEL PROYECTO O PROGRAMA. Se refiere a los resultados físicos, financieros, institucionales, sociales, ambientales o de otra índole que se espera que el proyecto o programa contribuya a lograr (OECD, 2002).

-P-

PARTES INTERESADAS. Entidades, organizaciones, grupos o particulares que tienen un interés directo o indirecto en la intervención para el desarrollo o en su evaluación (OECD, 2002).

PERTINENCIA. Medida en que los objetivos de una intervención para el desarrollo son congruentes con los requisitos de los beneficiarios, las necesidades del país, las prioridades globales y las políticas de los asociados y donantes.

Nota: En retrospectiva, la cuestión de la pertinencia suele referirse a si los objetivos o el diseño de una intervención siguen siendo adecuados aun cuando hayan cambiado las circunstancias (OECD, 2002).

PROCESO. Conjunto de actividades interrelacionadas que producen un valor agregado al producto o servicio destinado a un cliente determinado, sea este interno o extremo. (CEPLAN, 2010)

PRODUCTO. Comprende los productos, los bienes de capital y los servicios que resultan de una intervención para el desarrollo; puede incluir también los cambios resultantes de la intervención que son pertinentes para el logro de los efectos directos (OECD, 2002).

PROGRAMA PRESUPUESTAL. Es la categoría que constituye un instrumento del Presupuesto por Resultados, que es una unidad de programación de las acciones de las entidades públicas, las que integradas y articuladas se orientan a proveer productos, para lograr un Resultado Específico en la población y así contribuir al logro (MEF, 2014).

PROGRAMAS SOCIALES. Toda intervención pública, previamente planificada y articulada de acciones, prestaciones y beneficios temporales, estructurada a través de objetivos, estrategias, instrumentos y metas, que brinden bienes y/o servicios destinados a lograr un propósito específico sobre personas o poblaciones en situación de pobreza, vulnerabilidad o riesgo social, con el fin de atender una necesidad urgente o revertir un problema que les afecte. Para efectos de la presente directiva, toda mención a programas sociales estará referida a los programas sociales del MIDIS (MIDIS, 2012).

PROPÓSITO. Objetivos declarados públicamente del programa o proyecto de desarrollo (OECD, 2002).

PROYECTOS Y PROGRAMAS DE INVERSIÓN PÚBLICA. Son el conjunto de intervenciones limitadas en el tiempo, de las cuales resulta un producto final que concurre a la expansión de la acción del Gobierno. (MEF, 2014).

-R-

RECOMENDACIONES. Propuestas que tienen por objeto mejorar la eficacia, la calidad o la eficiencia de una intervención para el desarrollo, rediseñar los

objetivos y/o reasignar los recursos. Las recomendaciones deberán estar vinculadas a las conclusiones (OECD, 2002).

RENDICIÓN DE CUENTAS. Obligación de demostrar que se ha efectuado el trabajo cumpliendo con las reglas y normas acordadas o de declarar de manera precisa e imparcial los resultados obtenidos en comparación con las funciones y/o planes encomendados. Esto puede exigir una demostración cuidadosa, que pueda defenderse incluso judicialmente, de que la labor realizada es congruente con los términos contractuales.

Nota: La rendición de cuentas en materia de desarrollo puede referirse a las obligaciones de las partes de actuar de conformidad con responsabilidades, funciones y expectativas de resultados claramente definidas, a menudo con respecto al uso prudente de los recursos. Para los evaluadores, denota la responsabilidad de proporcionar informes de seguimiento y evaluaciones de resultados precisos, imparciales y creíbles. Los administradores y ejecutivos del sector público, deben rendir cuentas ante los contribuyentes y/o los ciudadanos (OECD, 2002).

RIESGO. Es un evento desfavorable que no ha ocurrido aún, y de ocurrir puede afectar negativamente el cumplimiento de los objetivos estratégicos, siempre que no se tomen decisiones para disminuir o anular su efecto (CEPLAN, 2014).

RESULTADOS. Producto, efecto o impacto (intencional o no, positivo y/o negativo) de una intervención para el desarrollo.

Términos conexos: efecto directo, repercusión, impacto (OECD, 2002).

RETROALIMENTACIÓN. Transmisión de las constataciones (hallazgos) a las que se arribó en el proceso de evaluación a aquellas partes para quienes son pertinentes y útiles con objeto de facilitar el aprendizaje. Puede abarcar la recopilación y divulgación de observaciones, conclusiones, recomendaciones y lecciones de la experiencia (OECD, 2002).

REVISIÓN. Evaluación de los resultados de una intervención, periódicamente o de manera ad hoc. Nota: Con frecuencia, se realiza una « evaluación » para tener

una determinación más completa y/o profunda que la que produce una revisión. Las revisiones tienden a hacer hincapié en los aspectos operativos. En ocasiones, los términos « revisión » y « evaluación » se utilizan como sinónimos.

Término conexo: evaluación (OECD, 2002).

-S-

SEGUIMIENTO. Función continua que utiliza una recopilación sistemática de datos sobre indicadores especificados para proporcionar a los administradores y a las partes interesadas principales de una intervención para el desarrollo, indicaciones sobre el avance y el logro de los objetivos así como de la utilización de los fondos asignados.

Términos conexos: seguimiento del desempeño, indicador (OECD, 2002).

(o monitoreo) es un proceso organizado para verificar que una actividad o una secuencia de actividades transcurre como se había previsto dentro de un determinado periodo de tiempo. Reporta las fallas en el diseño y la implementación de los programas y proyectos, permite comprobar si se está manteniendo la ruta hacia el objetivo establecido. Estima la probabilidad de alcanzar los objetivos planeados, identifica las debilidades que deben ser atendidas y oportunamente recomienda cambios y propone soluciones (PCM, 2013).

Es el proceso de revisión regular de indicadores específicos, que tiene como objetivo generar información periódica, confiable, relevante y oportuna que verifica el cumplimiento de los resultados esperados y permite tomar decisiones para mejorar la eficacia, calidad, eficiencia, transparencia y rendición de cuentas de las políticas, planes, proyectos y programas (MIDIS, 2012).

SEGUIMIENTO (Fase). Fase del proceso de planeamiento estratégico mediante la cual se realiza el seguimiento continuo a las metas con el fin de retroalimentar el proceso de planeamiento estratégico para el logro de los objetivos estratégicos. Contribuye al logro de los objetivos estratégicos a través del control y evaluación

continuos de las metas con el fin de retroalimentar el proceso de planeamiento estratégico (CEPLAN, 2014).

SEGUIMIENTO DEL DESEMPEÑO. Proceso continuo de recolección y análisis de datos para comparar en qué medida se está ejecutando un proyecto, programa o política en función de los resultados previstos (OECD, 2002).

SISTEMA DE INFORMACIÓN. El sistema de información es un proceso para recoger, organizar y analizar datos, con el objetivo de convertirlos e información útil para la toma de decisiones. El sistema de información para el seguimiento, monitoreo y evaluación debe diseñar los procesos de recojo, sistematización y análisis de la información, desde la etapa inicial de diseño de los indicadores, hasta las evaluaciones de resultados e impacto (PCM, 2013).

Es un conjunto ordenado de elementos (recursos, herramientas y procedimientos) que interactúan entre sí con el propósito de satisfacer las demandas de información de una organización, programa o proyecto, y elevar el nivel de conocimientos que permitan un mejor apoyo a la toma de decisiones y el desarrollo de actividades, para lo cual debe enfatizarse en la recopilación, administración y manipulación de todo aquel dato o información que sea necesaria (MIDIS, 2012).

SISTEMA DE SEGUIMIENTO: Sistema observacional para verificar el cumplimiento de las actividades en función de la programación y metas establecidas, así como el adecuado y eficiente uso de los recursos. El sistema de seguimiento debe proveer el flujo continuo de información durante la operación de la organización, programa o proyecto para hacer posible una toma de decisiones informada, a partir del sistema de información (MIDIS, 2012).

SOSTENIBILIDAD. Continuación de los beneficios de una intervención para el desarrollo después de concluida. Probabilidad de que continúen los beneficios en el largo plazo. Situación en la que las ventajas netas son susceptibles de resistir los riesgos con el correr del tiempo (OECD, 2002).

SUPERVISIÓN TÉCNICA. Proceso dinámico permanente de orientación, guía, educación y apoyo técnico al personal para que se desarrolle, supere y realice su trabajo en forma eficaz y eficiente. Es inherente a la acción directriz o ejecutiva de todo jefe y se sustenta en mecanismos de enseñanza - aprendizaje que permite el intercambio de experiencias entre supervisor y supervisado (MIMP, 2011).

SUPUESTOS. Hipótesis sobre los factores o riesgos que pueden afectar el avance o éxito de una intervención para el desarrollo. Nota: Este término puede también aplicarse a las condiciones hipotéticas que inciden en la validez de la evaluación misma, por ejemplo, las características de la población de la cual se extrae una muestra para una encuesta. Los supuestos se plantean de manera explícita en evaluaciones basadas en la teoría, cuando las mismas son verificadas sistemáticamente mediante una comparación con la cadena de resultados prevista (OECD, 2002).

-T-

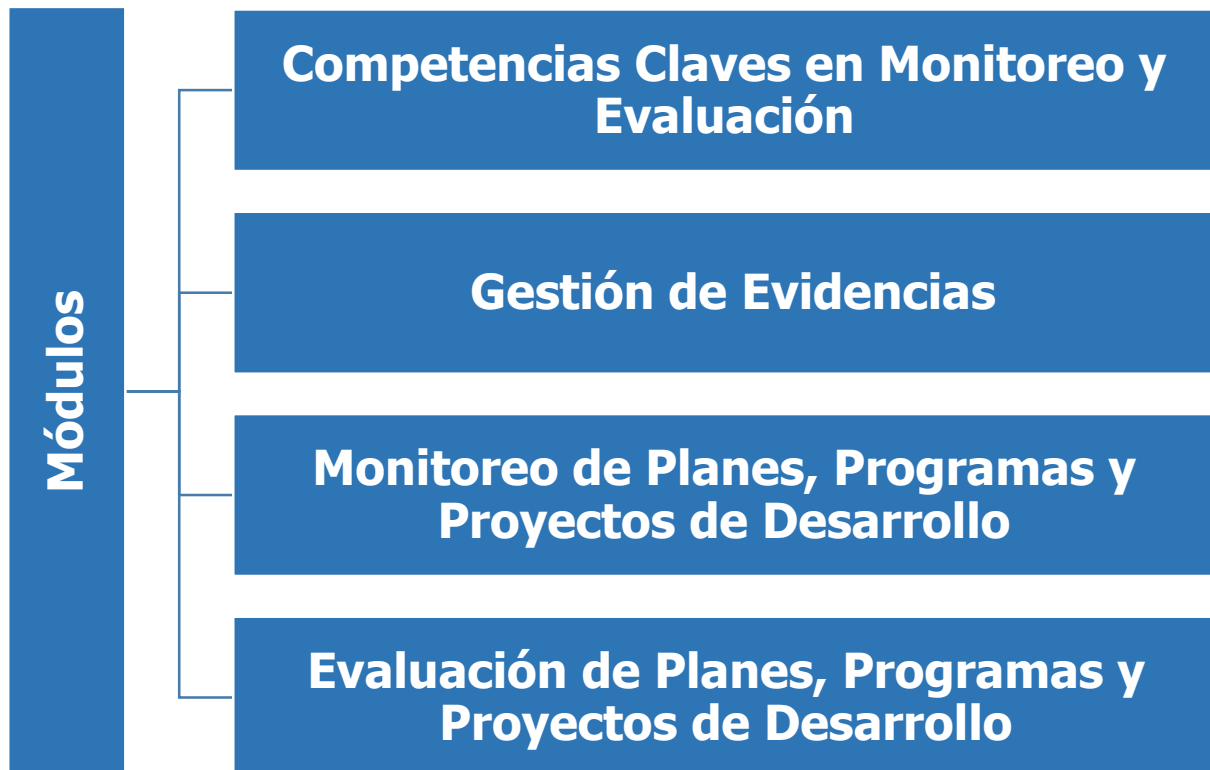
TÉRMINOS DE REFERENCIA. Documento escrito que presenta el propósito y el alcance de la evaluación, los métodos que se han de utilizar, la norma con la que se evaluarán los resultados o los análisis que se han de realizar, los recursos y el tiempo asignado, y los requisitos de presentación de informes. A veces se utilizan también las expresiones «alcance del trabajo » o « mandato de la evaluación » (OECD, 2002).

TRIANGULACIÓN. Uso de tres o más teorías, fuentes o tipos de información, o tipos de análisis para verificar y sustentar una evaluación. Nota: Al combinar múltiples fuentes de datos, métodos, análisis o teorías, los evaluadores procuran eliminar el sesgo que surge cuando se recurre a una sola fuente de información, a un solo método, a un único observador o a una teoría única (OECD, 2002).

-V-

VALIDEZ. Medida en que las estrategias y los instrumentos de recopilación de datos miden lo que tienen por objeto medir (OECD, 2002)

Serie Modular para el desarrollo de capacidades en Monitoreo y Evaluación



Agencia de los Estados Unidos para el Desarrollo Internacional
Av. La Encalada s/n, Santiago de Surco
Lima, Perú