ACADEMIA HUMANITARIA · HUMANITAIRES JOBS

Mini MBA Programas y MEAL: diseñar, dirigir y demostrar el impacto

Manual completo

Programa de exactamente 3 meses (12 semanas, 144 horas) para diseñar, dirigir y evaluar programas humanitarios: evaluación de necesidades y severidad, cadena de resultados e indicadores, focalización y transferencias monetarias, planificación y presupuesto, plan MEAL y calidad de los datos, muestreo, análisis, rendición de cuentas, evaluación, medición del impacto, optimización de recursos y gestión adaptativa. Ejercicios numéricos corregidos, tres exámenes mensuales con estudios de caso y un diploma verificable de la Academia Humanitaria.

https://humanitairejobs.org/es/academia/mini-mba-programmes-meal/

Documento gratuito de la Academia Humanitaria. Puede imprimirse y compartirse libremente con fines de formación, sin modificaciones.

-

Índice

1. Semana 1 : Evaluar las necesidades: datos, severidad y prioridades
2. Semana 2 : Diseñar un programa: cadena de resultados, indicadores y metas
3. Semana 3 : Focalización y modalidades: a quién ayudar, cómo y con cuánto
4. Semana 4 : Planificar y presupuestar un programa: del plan de trabajo al ritmo de gasto
5. Semana 5 : El sistema de seguimiento: plan MEAL, calidad de los datos y seguimiento a distancia
6. Semana 6 : Encuestas y muestreo: obtener cifras correctas
7. Semana 7 : Analizar los datos: estadísticas útiles, trampas y visualización
8. Semana 8 : Rendición de cuentas a las poblaciones: participación, quejas y retroalimentación
9. Semana 9 : Evaluar: criterios, términos de referencia, matriz y utilización
10. Semana 10 : Medir el impacto: contrafactual, métodos y contribución
11. Semana 11 : Optimización de recursos: costo-eficiencia, costo-eficacia y arbitrajes
12. Semana 12 : Gestión adaptativa, aprendizaje e informes: dirigir con base en la evidencia

Semana 1 : Evaluar las necesidades: datos, severidad y prioridades

Objetivos

- Situar la evaluación de necesidades en el ciclo del programa humanitario y elegir el tipo de evaluación adecuado según la fase de la crisis.
- Combinar datos secundarios y primarios, detectar los sesgos y clasificar la severidad de una zona con una regla explícita.
- Pasar de la constatación a la decisión: análisis de la respuesta, viabilidad, valor añadido y elección argumentada de las prioridades.

La evaluación de necesidades en el ciclo del programa humanitario

El **ciclo del programa humanitario** (HPC) encadena cinco etapas: evaluación y análisis de necesidades, planificación estratégica, movilización de recursos, ejecución y seguimiento, revisión y evaluación. El primer producto colectivo es el **panorama de necesidades humanitarias** (HNO), que alimenta el plan de respuesta (desde 2024, la mayoría de los países publican un documento único, el HNRP). Una ONG que diseña un programa sin leer el HNO de su país se priva de la mejor síntesis disponible y corre el riesgo de proponer lo que otros ya cubren.

El tipo de evaluación depende de la fase de la crisis:

Fase	Herramienta típica	Plazo	Pregunta de decisión
Choque repentino	Evaluación rápida inicial multisectorial (MIRA o equivalente)	72 horas a 2 semanas	¿Dónde están las necesidades vitales y quién interviene dónde?
Crisis prolongada	Evaluación multisectorial de necesidades (MSNA), encuesta de hogares representativa	Anual	¿Qué severidad por zona y por grupo, cuántas personas con necesidades?
Programa en curso	Evaluación sectorial en profundidad (nutrición SMART, mercado, protección)	Según necesidad	¿Hay que ajustar la modalidad, la focalización, el valor de las transferencias?

Hay que distinguir tres nociones. La **necesidad** es la brecha entre la situación observada y una norma (por ejemplo, 15 litros de agua por persona y por día según Esfera). La **vulnerabilidad** es la probabilidad de sufrir un daño frente al choque, que depende de la exposición y de las capacidades. Las **capacidades** son los recursos de las personas, de las comunidades y de las instituciones. Un buen diagnóstico describe las tres; una lista de carencias sin capacidades conduce a respuestas que sustituyen lo que ya funcionaba.

> Regla de oro: cada pregunta de una evaluación debe servir a una decisión identificada de antemano. Una pregunta «interesante» que no cambia ninguna decisión le cuesta tiempo a las personas entrevistadas y diluye el análisis.

Los datos secundarios: empezar por lo que existe

La **revisión de datos secundarios** (SDR) reúne lo que ya se sabe antes de recopilar nada: HNO, informes de clúster, encuestas de demografía y salud (EDS), análisis CIF, evaluaciones de otras ONG, datos de los servicios técnicos del Estado, seguimiento de los desplazamientos (DTM de la OIM), precios en los mercados, informes de protección. Una SDR bien hecha reduce a la mitad el tamaño de una recopilación primaria, porque permite plantear solo las preguntas sin respuesta.

Cada fuente se califica en dos dimensiones. La **fiabilidad de la fuente**: método conocido, muestra descrita, autor sin conflicto de intereses. La **credibilidad de la información**: confirmada por otras fuentes, coherente con lo que se sabe, reciente. Una información de una fuente fiable pero con dos años de antigüedad en una zona de desplazamientos masivos es débil; una información de una fuente mediana confirmada por otras tres es sólida.

La SDR termina con un **análisis de las brechas de información**: para cada pregunta de decisión, ¿qué se sabe, con qué nivel de confianza y qué queda por recopilar? Es esta tabla, y no una lista de temas, la que justifica el cuestionario.

- **Triangulación de fuentes**: la misma información confirmada por fuentes independientes.
- **Triangulación de métodos**: una encuesta de hogares y unos grupos focales que convergen.
- **Triangulación de encuestadores**: equipos distintos que obtienen los mismos resultados.

Los datos primarios: elegir los métodos y controlar los sesgos

Cuatro métodos se complementan. La **entrevista con informantes clave** (KII) interroga a personas que tienen una visión de conjunto: jefe de aldea, enfermero, maestra, comerciante, responsable de una asociación de mujeres. Rápida y poco costosa, da tendencias pero no proporciones. El **grupo focal** (FGD), de 6 a 10 personas homogéneas (mismo sexo y mismo grupo de edad), explora las percepciones, las normas y las preferencias. La **observación directa** (letrinas, mercados, filas de espera) verifica lo declarado. La **encuesta de hogares** sobre una muestra representativa es la única que produce proporciones generalizables, siempre que el muestreo sea riguroso (semana 6).

Los sesgos más frecuentes:

- **Sesgo de selección**: entrevistar solo a las aldeas accesibles por carretera, o a los hogares presentes durante el día (se pierde a los jornaleros).
- **Sesgo de complacencia y de expectativa de ayuda**: las personas exageran o minimizan según lo que creen que el encuestador quiere oír o según la ayuda que esperan. Se reduce explicando claramente que no se promete ninguna ayuda y prefiriendo indicadores observables.

- **Sesgo del encuestador**: formulación sugestiva, presencia de un hombre para preguntar a mujeres sobre la violencia. Se forma, se prueba el cuestionario y se constituyen equipos mixtos.
- **Sesgo de supervivencia**: en una zona de alta mortalidad o de salidas, los hogares más afectados ya no están para responder.

> Ética mínima: consentimiento informado, posibilidad de negarse sin consecuencias, ninguna pregunta sensible sin protocolo de derivación (violencia de género, protección de la infancia) y datos personales reducidos a lo estrictamente necesario.

Medir la severidad: reglas explícitas y personas con necesidades

Comparar zonas exige una **escala de severidad** común. El marco conjunto de análisis intersectorial (JIAF) y la Clasificación Integrada de las Fases de la Seguridad Alimentaria (CIF) utilizan ambos cinco fases, de 1 (mínima) a 5 (catastrófica). La CIF clasifica una zona según una **regla del 20 %**: la zona se sitúa en la fase más severa que afecta, junto con las fases superiores, al menos al 20 % de su población.

Ejemplo. Población de 100 000 personas: fase 1: 35 %; fase 2: 30 %; fase 3: 22 %; fase 4: 10 %; fase 5: 3 %. Fases 4 y 5 juntas: 13 %, por debajo del umbral. Fases 3, 4 y 5 juntas: 35 %, por encima del 20 %. Por tanto, la zona se clasifica en **fase 3**. El número de personas que necesitan una acción urgente (fase 3 o superior) es 35 000.

El número de **personas con necesidades** (PIN) de un sector no es la población total de la zona: es la suma de las personas cuya situación alcanza el umbral de severidad definido (a menudo fase 3 o superior). Sumar los PIN de varios sectores cuenta varias veces a la misma persona. Por eso, para la cifra intersectorial, se retiene el máximo por zona y por grupo, o un análisis a nivel del hogar.

La severidad debe estar **desagregada** por sexo, edad y discapacidad. Para la discapacidad, el conjunto breve del Grupo de Washington plantea seis preguntas (ver, oír, caminar, recordar, cuidar de sí mismo, comunicarse); una persona se cuenta como persona con discapacidad si responde «mucha dificultad» o «no puede en absoluto» a al menos una pregunta.

Análisis de protección, de género y de mercado

Una evaluación que solo mira las carencias materiales pasa por alto los riesgos. El **análisis de protección** identifica las amenazas (quién hace qué a quién), las vulnerabilidades de las personas frente a esas amenazas y sus capacidades: riesgo = amenaza × vulnerabilidad / capacidad. Detecta los grupos excluidos de la ayuda (personas sin documentos, minorías, personas mayores aisladas) y los riesgos que la propia ayuda puede crear: distribuciones que exponen a robos, trayectos peligrosos, tensiones entre desplazados y población de acogida.

El **análisis de género** describe quién decide, quién controla los recursos, quién tiene acceso a los servicios y cómo la crisis modifica estos roles. Evita, por ejemplo, entregar la ayuda monetaria a un miembro del hogar que no controla su uso, o fijar horarios de distribución incompatibles con las tareas de las mujeres.

El **análisis de mercado** determina si las transferencias monetarias son viables: ¿funcionan los mercados, pueden los comerciantes responder a un aumento de la demanda sin inflación, es seguro el acceso físico y financiero? Las herramientas de referencia (evaluación rápida de mercado, cartografía de los sistemas de mercado) responden a tres preguntas: disponibilidad de los bienes, precios y su evolución, capacidad de reabastecimiento.

> No hacer daño: para cada opción de respuesta, se escribe al menos un riesgo que crea y la medida que lo reduce.

Del análisis a la decisión: el análisis de la respuesta

El **análisis de la respuesta** establece el vínculo entre lo que se sabe de las necesidades y lo que la organización va a hacer. Compara opciones (en especie, efectivo, cupones, servicios, incidencia) según criterios explícitos: pertinencia respecto a las necesidades prioritarias, preferencias de las personas, viabilidad (acceso, mercados, seguridad, capacidades), costo-eficacia, riesgos y lo que ya hacen los demás actores (matriz 3W o 4W: quién hace qué, dónde y cuándo).

Un método sencillo es la **matriz ponderada**: cada criterio recibe un peso fijado antes de calificar las opciones, para evitar elegir primero y justificar después. Por ejemplo, pertinencia 30 %, preferencias 20 %, viabilidad 20 %, costo-eficacia 20 %, riesgos 10 %. Una opción que obtiene 4/5 en pertinencia, 5 en preferencias, 3 en viabilidad, 4 en costo-eficacia y 3 en riesgos suma 0,3 × 4 + 0,2 × 5 + 0,2 × 3 + 0,2 × 4 + 0,1 × 3 = 3,9.

La decisión final también debe tener en cuenta el **valor añadido** de la organización: presencia, aceptación, experiencia, socios locales. Si un actor ya cubre correctamente una necesidad en una zona, la mejor decisión suele ser ir a otro lugar o complementar, y no duplicar.

El producto final es una **nota de análisis** de unas pocas páginas: necesidades prioritarias por zona y por grupo, severidad, brechas de cobertura, opciones comparadas, recomendación y nivel de confianza. Es esta nota la que leerá la dirección, no la base de datos.

Ejercicio: clasificar dos zonas y elegir dónde intervenir

Situación. Dos distritos afectados por la sequía. Distrito A (60 000 habitantes): fase 1: 30 %; fase 2: 32 %; fase 3: 25 %; fase 4: 11 %; fase 5: 2 %. Distrito B (80 000 habitantes): fase 1: 45 %; fase 2: 37 %; fase 3: 12 %; fase 4: 5 %; fase 5: 1 %. Otra ONG ya distribuye alimentos a 15 000 personas en fase 3 o superior en el distrito A. Los mercados del distrito B funcionan; los del distrito A están muy perturbados.

Trabajo solicitado. 1) Clasifique cada distrito según la regla del 20 %. 2) Calcule el número de personas en fase 3 o superior en cada distrito y el número no cubierto. 3) Recomiende una zona y una modalidad, y justifíquelo en tres frases.

Corrección. Distrito A: fases 4 y 5 = 13 %, fases 3 a 5 = 38 %: fase 3. Personas en fase 3 o superior: 0,38 × 60 000 = 22 800; no cubiertas: 22 800 − 15 000 = 7 800. Distrito B: fases 3 a 5 = 18 %, por debajo del umbral; fases 2 a 5 = 55 %: fase 2. Personas en fase 3 o superior: 0,18 × 80 000 = 14 400, todas no cubiertas. El distrito B tiene más personas no cubiertas (14 400 frente a 7 800) a pesar de una clasificación menos severa, y sus mercados permiten transferencias monetarias. Una respuesta defendible: transferencias monetarias focalizadas en los hogares en fase 3 o superior del distrito B, y un complemento en especie en el distrito A en coordinación con la otra ONG. La clasificación de una zona no basta para decidir: hay que mirar las cifras absolutas y la cobertura.

Para recordar

- Cada pregunta de evaluación debe servir a una decisión conocida de antemano.
- Primero la revisión de datos secundarios, con una calificación de fiabilidad y de credibilidad y un análisis de las brechas de información.
- Solo una encuesta representativa produce proporciones generalizables; las KII y los FGD dan tendencias y explicaciones.
- Regla del 20 %: una zona toma la fase más severa alcanzada, junto con las fases superiores, por al menos el 20 % de su población.
- El análisis de la respuesta compara opciones según criterios ponderados fijados de antemano, teniendo en cuenta la cobertura existente.

Cuestionario

1. Distribución de una zona: fase 1: 40 %; fase 2: 35 %; fase 3: 15 %; fase 4: 8 %; fase 5: 2 %. Según la regla del 20 %, la zona está en:
   A. Fase 2
   B. Fase 3
   C. Fase 4
   D. Fase 1

2. ¿Por qué no se pueden sumar las personas con necesidades de cada sector para obtener un total intersectorial?
   A. Porque los sectores utilizan monedas diferentes
   B. Porque la misma persona puede tener necesidades en varios sectores y se contaría varias veces
   C. Porque el total sería demasiado bajo
   D. Porque los donantes lo prohíben

3. ¿Qué método permite estimar una proporción representativa de hogares sin acceso a agua potable?
   A. Tres entrevistas con jefes de aldea
   B. Dos grupos focales
   C. Una encuesta de hogares sobre una muestra aleatoria correctamente dimensionada
   D. La observación de un punto de agua

4. Según el conjunto breve del Grupo de Washington, una persona se cuenta como persona con discapacidad si:
   A. Tiene un certificado médico
   B. Responde «alguna dificultad» a una pregunta
   C. Responde «mucha dificultad» o «no puede en absoluto» a al menos una de las seis preguntas
   D. Tiene más de 60 años

5. En una matriz ponderada, ¿por qué fijar los pesos antes de calificar las opciones?
   A. Para ir más rápido
   B. Para evitar elegir primero una opción y ajustar después los criterios para justificarla
   C. Porque el donante lo exige
   D. Para obtener puntuaciones más altas

Respuestas del cuestionario : 1. B · 2. B · 3. C · 4. C · 5. B

Semana 2 : Diseñar un programa: cadena de resultados, indicadores y metas

Objetivos

- Construir una cadena de resultados coherente, de la actividad al impacto, y una teoría del cambio cuyos supuestos sean comprobables.
- Elegir indicadores válidos y medibles, con definición, línea de base, meta y fuente, y utilizar los indicadores estándar del sector.
- Fijar metas realistas a partir de los datos de línea de base y de la capacidad real, e integrar género, edad y discapacidad desde el diseño.

La cadena de resultados: lo que depende de usted y lo que no

Una **cadena de resultados** une cinco niveles: **insumos** (dinero, personal, material), **actividades** (formar, distribuir, construir), **productos** (resultados directos: 400 letrinas construidas, 1 200 hogares formados), **efectos** (cambios de comportamiento o de situación: el 80 % de los hogares usa las letrinas), **impacto** (cambio duradero a largo plazo: disminución de las diarreas en los niños menores de cinco años).

La frontera decisiva está entre producto y efecto. La organización **controla** sus productos: si las letrinas no se construyen, es su responsabilidad. **Influye** en los efectos, que dependen también de las personas y del contexto. **Contribuye** al impacto, junto con muchos otros actores. Un informe que solo presenta productos dice lo que se hizo, no lo que cambió; un proyecto que promete un impacto sin efecto medible promete lo que no podrá demostrar.

> Prueba del «¿y qué?»: para cada producto, pregunte «¿y qué?» hasta llegar a un cambio que importe a las personas. Si la respuesta se detiene en «la gente recibió un kit», el programa no tiene un efecto definido.

Cada paso de un nivel a otro se basa en **supuestos**. Construir letrinas conduce a su uso solo si son culturalmente aceptables, seguras de noche para las mujeres, mantenidas y acompañadas de sensibilización. Estos supuestos son exactamente lo que el seguimiento deberá verificar.

De la teoría del cambio al marco lógico

La **teoría del cambio** explica por qué y cómo se producirá el cambio esperado, con sus supuestos causales, las evidencias que los sustentan y el papel de los demás actores. El **marco lógico** (logframe) es su traducción operativa: una matriz de cuatro columnas (lógica de intervención, indicadores, fuentes de verificación, supuestos) y cuatro filas (objetivo general, objetivo específico, resultados, actividades).

La **lógica vertical** se lee de abajo arriba: si se realizan las actividades y se cumplen los supuestos de ese nivel, entonces se logran los resultados; si se logran los resultados y se cumplen los supuestos, entonces se alcanza el objetivo específico. La **lógica horizontal** se lee

por fila: cada nivel tiene indicadores medibles y una fuente que permite efectivamente medirlos.

Errores frecuentes en la revisión de propuestas:

- Un objetivo específico que reformula los resultados («los hogares recibieron transferencias») en lugar de un cambio («los hogares focalizados mejoran su consumo de alimentos»).
- Varios objetivos específicos en un proyecto pequeño, que dispersan la medición.
- Supuestos que en realidad son riesgos que la organización debería gestionar («el personal se contrata a tiempo»). Un supuesto está fuera del control del proyecto; si es crítico y muy improbable, es un supuesto fatal y el proyecto debe rediseñarse.
- Fuentes de verificación que nadie producirá («estadísticas nacionales» que no existen a nivel de distrito).

Indicadores que miden de verdad

Un buen indicador es **válido** (mide lo que pretende medir), **fiable** (dos encuestadores obtienen el mismo resultado), **sensible** (se mueve si la situación cambia durante el proyecto), **sencillo** de recopilar y **útil** para una decisión. El criterio SMART (específico, medible, alcanzable, pertinente, con plazo) se aplica sobre todo a la meta asociada.

Cada indicador clave debe tener una **ficha de referencia** (PIRS): definición precisa, numerador y denominador, unidad, desagregación, método y herramienta de recopilación, frecuencia, responsable, línea de base y meta. Sin ficha, dos oficinas calculan el mismo indicador de dos maneras y las cifras no se pueden sumar.

Ejemplo de definición rigurosa. «Porcentaje de hogares focalizados con una puntuación de consumo de alimentos aceptable (superior a 35)». Numerador: hogares de la muestra con una puntuación superior a 35. Denominador: hogares focalizados de la muestra. Desagregación: sexo del jefe de hogar, presencia de una persona con discapacidad, estatus (desplazado, de acogida).

El sector dispone de **indicadores estándar** que es mejor utilizar que reinventar, porque permiten comparar y agregar: puntuación de consumo de alimentos (FCS: umbrales 21 y 35 en la mayoría de los contextos, 28 y 42 donde el consumo de aceite y azúcar es elevado), índice reducido de estrategias de afrontamiento (rCSI), escala del hambre en el hogar (HHS), perímetro braquial (MUAC: menos de 115 mm para la desnutrición aguda severa en niños de 6 a 59 meses), indicadores del programa conjunto de monitoreo WASH (JMP) y los registros de indicadores de los clústeres.

Línea de base y metas realistas

Una meta sin **línea de base** no tiene sentido: «el 70 % de los hogares con un consumo aceptable» es un enorme progreso si se parte del 30 %, y un retroceso si se parte del 80 %. La línea de base se mide antes del inicio de las actividades, con el mismo método que la medición final. Cuando es imposible (emergencia repentina), se utilizan datos secundarios recientes y se mide desde las primeras semanas, documentándolo.

Una meta realista se calcula, no se adivina. Combina: la línea de base, la **cobertura** real del proyecto (proporción de la población concernida que se alcanza), la **eficacia** esperada de la intervención según las evidencias disponibles y los factores externos (estación, precios).

Ejemplo. Línea de base: 40 % de hogares con un consumo aceptable. El proyecto alcanza al 80 % de los hogares focalizados. Proyectos comparables hicieron pasar al 50 % de los hogares alcanzados de un consumo insuficiente a aceptable. Entre el 60 % de hogares por debajo del umbral, 0,8 × 0,5 = 40 % pasan por encima: 0,6 × 0,4 = 24 puntos. Meta: 40 + 24 = 64 %, que se redondeará prudentemente al 60 % si el periodo de escasez es más duro de lo previsto.

> Una meta demasiado ambiciosa no motiva: obliga a explicar un fracaso que no lo es y empuja a embellecer los datos.

Género, edad, discapacidad e inclusión desde el diseño

La inclusión no se añade al final: se diseña. El **marcador de género con edad** del IASC (GAM) evalúa los proyectos sobre elementos concretos: análisis de las necesidades diferenciadas, actividades adaptadas, participación de los distintos grupos en el diseño y en la decisión, y mecanismos de retroalimentación accesibles. Un proyecto que no recopila datos desagregados no puede mostrar que las mujeres, las personas mayores o las personas con discapacidad se benefician de él.

En concreto: desagregar los indicadores por sexo, edad (al menos por tramos 0-5, 6-17, 18-59, 60 y más) y discapacidad; adaptar los sitios (rampas, horarios, distancia, filas prioritarias); prever el costo de la accesibilidad en el presupuesto (a menudo del 3 al 5 % de los costos de las actividades concernidas); consultar por separado a mujeres y hombres, jóvenes y mayores.

Las normas humanitarias de inclusión de las personas mayores y de las personas con discapacidad y las directrices del IASC sobre la inclusión de las personas con discapacidad (2019) proponen acciones mínimas: **identificar** (datos desagregados), **eliminar las barreras** (físicas, de comunicación, de actitud), **empoderar** (participación en las decisiones) y **fortalecer las capacidades** de las organizaciones de personas con discapacidad.

Verificar el diseño antes de presentarlo

Antes de presentar una propuesta, una breve revisión interna evita la mayoría de los rechazos y de las malas sorpresas durante el proyecto. Bastan cinco preguntas:

- **Coherencia**: ¿contribuye cada actividad a un resultado, y cada resultado al objetivo específico? Una actividad sin resultado debe suprimirse; un resultado sin actividad es irrealista.
- **Medibilidad**: ¿tiene cada indicador una definición, una línea de base, una meta, una fuente, un presupuesto y un responsable?
- **Proporcionalidad**: ¿cuesta el dispositivo de seguimiento entre el 3 y el 10 % del presupuesto (rango habitual), en relación con lo que piden el donante y la gestión?
- **Supuestos**: ¿se siguen los supuestos críticos con un indicador de contexto?

- **Personas**: ¿se ha consultado a las comunidades sobre la modalidad, la focalización y los mecanismos de quejas?

El marco lógico no es fijo. Se revisa cuando cambia el contexto, siempre que se documente el motivo, se consulte al donante para cualquier modificación de objetivo o de meta y se conserve el registro de las versiones. Un marco lógico que nunca se ha movido en un contexto volátil suele ser un marco lógico que nadie usa.

Ejercicio: corregir un marco lógico y fijar una meta

Situación. Un proyecto WASH de 12 meses propone: objetivo específico «400 letrinas construidas»; resultado 1 «los hogares adoptan buenas prácticas de higiene»; indicador del objetivo específico «número de kits de higiene distribuidos»; supuesto «el personal se contratará a tiempo». Línea de base: el 35 % de los hogares se lava las manos con jabón en los momentos clave. El proyecto alcanzará al 70 % de los hogares de la zona; en proyectos comparables, el 40 % de los hogares alcanzados que no tenían la práctica la adoptan.

Trabajo solicitado. 1) Identifique los cuatro errores. 2) Proponga un objetivo específico y un indicador correctos. 3) Calcule una meta realista para el lavado de manos.

Corrección. Errores: el objetivo específico es un producto; el resultado 1 es un efecto situado en el nivel de los resultados; el indicador mide una distribución y no un cambio; el supuesto es un riesgo interno que la organización debe gestionar. Objetivo específico posible: «reducir los riesgos de enfermedades relacionadas con el agua y el saneamiento en los hogares focalizados», indicador: «porcentaje de hogares focalizados que se lavan las manos con jabón en los momentos clave (observación y declaración)». Meta: el 65 % de los hogares no tiene la práctica; 0,7 × 0,4 = 28 % de ellos la adoptan; 0,65 × 0,28 = 18,2 puntos. Meta: 35 + 18 = aproximadamente 53 %.

Para recordar

- La organización controla sus productos, influye en los efectos y contribuye al impacto.
- El marco lógico se lee de abajo arriba (si… entonces…) y por fila (indicador, fuente, supuesto).
- Un indicador clave tiene una ficha de referencia: definición, numerador, denominador, desagregación, método, frecuencia, responsable.
- Meta = línea de base + cobertura × eficacia esperada aplicada a la proporción de la población concernida.
- Desagregar por sexo, edad y discapacidad desde el diseño, y presupuestar la accesibilidad.

Cuestionario

1. «1 500 hogares recibieron un kit de higiene» es:
   A. Un impacto
   B. Un efecto
   C. Un producto
   D. Un supuesto

2. Un supuesto crítico y muy improbable en un marco lógico es:
   A. Sin importancia
   B. Un supuesto fatal: el proyecto debe rediseñarse
   C. Un indicador
   D. Un riesgo que se puede ignorar

3. Línea de base: 50 % de hogares con un consumo aceptable. Cobertura: 60 %; eficacia: el 50 % de los hogares alcanzados por debajo del umbral pasan por encima. Meta calculada:
   A. 65 %
   B. 80 %
   C. 55 %
   D. 60 %

4. Según los umbrales más habituales, una puntuación de consumo de alimentos de 30 es:
   A. Aceptable
   B. Límite
   C. Pobre
   D. Imposible

5. ¿Qué elemento falta con más frecuencia en un indicador mal definido?
   A. Un título
   B. Un numerador y un denominador precisos
   C. Un color
   D. Una traducción

Respuestas del cuestionario : 1. C · 2. B · 3. A · 4. B · 5. B

Semana 3 : Focalización y modalidades: a quién ayudar, cómo y con cuánto

Objetivos

- Elegir un método de focalización adaptado al contexto y medir sus errores de inclusión y de exclusión.
- Comparar la ayuda en especie, los cupones y las transferencias monetarias según criterios explícitos, incluidos los riesgos.
- Calcular el valor de una transferencia a partir de una canasta mínima de gastos y ajustarlo a la inflación.

Por qué focalizar, y lo que cuesta una mala focalización

Con recursos inferiores a las necesidades, la **focalización** decide quién recibe la ayuda. Produce dos errores. El **error de inclusión**: hogares que no cumplen los criterios reciben la ayuda (despilfarro, pérdida de confianza). El **error de exclusión**: hogares elegibles no la reciben (perjuicio directo, a menudo para los más marginados). Reducir uno suele aumentar el otro: criterios muy estrictos excluyen a hogares vulnerables; criterios amplios incluyen a hogares que no la necesitan.

Cálculo. De 1 000 hogares evaluados, 400 son realmente elegibles. El proyecto selecciona 450, de los cuales 330 son elegibles. Error de inclusión: 120 / 450 = 26,7 % de los hogares asistidos no son elegibles. Error de exclusión: (400 − 330) / 400 = 17,5 % de los hogares elegibles quedan fuera.

> En el ámbito humanitario, el error de exclusión es el más grave cuando afecta a la supervivencia. Una política de focalización debe decir explícitamente cuál de los dos errores se acepta más, y por qué.

La focalización también tiene un **costo**: encuestas, verificación, quejas, tensiones sociales. Cuando más del 70 al 80 % de la población de una zona tiene necesidades, una distribución general (solo focalización geográfica) suele ser más justa y menos costosa que una focalización hogar por hogar.

Los métodos de focalización

Los métodos se combinan a menudo en varias etapas:

- **Geográfica**: elegir las zonas más severas (semana 1). Sencilla y poco cuestionable, pero imprecisa dentro de una zona.
- **Categórica**: criterios observables (hogar encabezado por una mujer sola, persona con discapacidad, niño menor de dos años, desplazado reciente). Transparente, pero estas categorías solo coinciden de forma imperfecta con la vulnerabilidad real.

- **Comunitaria**: un comité representativo identifica a los hogares según criterios explicados públicamente. Aprovecha el conocimiento local, pero corre el riesgo de captura por las élites y de exclusión de las minorías. Se controla mediante una verificación aleatoria y un mecanismo de quejas independiente.
- **Puntuación de vulnerabilidad** (proxy means test): un cuestionario de hogar da una puntuación calculada a partir de variables correlacionadas con la pobreza (bienes, vivienda, composición). Más objetiva, pero costosa, opaca para las personas y sensible a las declaraciones falsas.
- **Autofocalización**: la ayuda solo atrae a quienes la necesitan (trabajo a cambio de remuneración con un salario inferior al del mercado, productos poco apreciados). Barata, pero puede excluir a las personas que no pueden trabajar.

La elección depende de cuatro factores: la proporción de personas con necesidades, la disponibilidad de datos fiables, la cohesión social (riesgo de tensiones si la focalización parece injusta) y el tiempo disponible. En todos los casos, los **criterios deben ser públicos** y comprendidos por la comunidad, y un mecanismo de quejas debe permitir impugnar una exclusión.

En especie, cupones o transferencias monetarias

Las **transferencias monetarias** (efectivo) representan ya aproximadamente una quinta parte de la ayuda humanitaria internacional. Los datos comparativos muestran que, con resultados equivalentes, a menudo cuestan menos de entregar que la ayuda en especie, permiten que las personas elijan según sus prioridades y apoyan los mercados locales. No siempre son la respuesta correcta.

Criterio	En especie	Cupones	Efectivo sin restricción
Mercados no funcionales	Adaptado	Poco adaptado	Inadecuado
Necesidad de un producto concreto (alimento terapéutico, semillas certificadas)	Adaptado	Adaptado	Poco adaptado
Necesidades variadas según los hogares	Poco adaptado	Medio	Adaptado
Costo logístico	Elevado	Medio	Bajo
Dignidad y elección	Baja	Media	Alta

Los **riesgos** del efectivo se gestionan: inflación local (seguimiento de precios, topes de transferencia), inseguridad (pago electrónico, distribuciones fraccionadas), desvío por un miembro del hogar (elección del destinatario con análisis de género), sanciones y lucha antiterrorista (control de los proveedores financieros). La idea de que el efectivo se desviaría hacia el alcohol o el tabaco está contradicha por numerosos estudios: el gasto en «bienes de tentación» no aumenta de forma significativa.

El **efectivo multipropósito** (MPCA) cubre varias necesidades básicas a la vez; se coordina dentro del grupo de trabajo de efectivo, que fija un valor de transferencia común para evitar que las organizaciones compitan entre sí.

Calcular el valor de una transferencia

El valor de una transferencia parte de la **canasta mínima de gastos** (MEB): el costo mensual de los bienes y servicios que un hogar necesita para cubrir sus necesidades básicas (alimentación, agua, higiene, alquiler, energía, transporte, comunicación). Se calcula para un tamaño de hogar tipo, a partir de los precios registrados en los mercados.

El valor de la transferencia no es necesariamente la MEB completa: se resta lo que el hogar ya cubre por sí mismo (ingresos, otras ayudas) y se tiene en cuenta el objetivo del programa. **Fórmula básica: brecha = MEB − recursos medios del hogar focalizado.**

Ejemplo. MEB mensual para un hogar de 6 personas: 210 USD, de los cuales 130 para alimentación. Ingreso mensual medio de los hogares focalizados: 75 USD. Brecha: 135 USD. El grupo de trabajo de efectivo fijó un valor común del 80 % de la brecha para no crear dependencia y preservar los ingresos propios: 0,8 × 135 = 108 USD al mes.

Inflación. Si el precio de la canasta aumenta un 12 % en tres meses, el valor real de una transferencia fija de 108 USD cae a 108 / 1,12 = 96 USD en poder adquisitivo inicial. Por eso los programas prevén una regla de ajuste: por ejemplo, revisión en cuanto la MEB varíe más del 10 %, sobre la base del seguimiento mensual de precios.

> Tamaño de los hogares: un monto fijo por hogar perjudica a los hogares grandes. Los programas utilizan a menudo tramos (1-3, 4-6, 7 personas y más) con un tope, para que sean sencillos de explicar.

Los mecanismos de pago y la protección de datos

El **mecanismo de entrega** se elige después de la modalidad: entrega de efectivo en mano, agentes de pago, transferencia por teléfono móvil, tarjeta prepagada, cuenta bancaria. Los criterios: cobertura de la red, acceso de las mujeres y de las personas mayores a los teléfonos, distancia a los puntos de retiro, comisiones, liquidez de los agentes y requisitos de identificación («conozca a su cliente»).

El pago electrónico reduce los riesgos de robo y facilita la trazabilidad, pero crea **riesgos para los datos**: listas de beneficiarios compartidas con un operador, con autoridades o con otros actores, que pueden exponer a personas en peligro (refugiados, minorías). El Manual del CICR sobre protección de datos en la acción humanitaria recomienda: minimización de datos, finalidad definida, acuerdos escritos de intercambio, evaluación de impacto relativa a la protección de datos antes de cualquier nuevo sistema y plazo de conservación limitado.

La **conciliación** compara cada mes tres listas: hogares registrados, pagos enviados, pagos efectivamente retirados. Las discrepancias (retiros no efectuados, duplicados, números de teléfono compartidos por varios beneficiarios) son señales de error o de fraude que deben tratarse rápidamente.

El monitoreo posdistribución

El **monitoreo posdistribución** (PDM) verifica, entre dos y seis semanas después de una distribución o una transferencia, que la ayuda ha llegado, que se utiliza y que produce los efectos esperados. Se realiza sobre una muestra aleatoria de beneficiarios y consta de cuatro bloques:

- **Proceso**: monto o cantidades recibidos, tiempo de espera, distancia, costo del trayecto, seguridad, respeto a las personas.
- **Uso**: gastos realizados con la transferencia, reventa de bienes en especie, reparto.
- **Efectos**: puntuación de consumo de alimentos, estrategias de afrontamiento, deuda.
- **Rendición de cuentas**: conocimiento de los criterios y del mecanismo de quejas, satisfacción, problemas señalados (exigencias de pago, acoso).

Un PDM solo tiene valor si desencadena decisiones. Por ejemplo: si más del 10 % de los beneficiarios declara haber pagado algo para ser inscrito, se abre una investigación y se cambia el proceso de registro; si el 40 % revende el aceite recibido, se ajusta la composición de la canasta o se pasa al efectivo.

Ejercicio: evaluar una focalización y fijar una transferencia

Situación. De 2 000 hogares evaluados, una verificación independiente muestra que 900 son realmente elegibles. El comité comunitario seleccionó 1 000, de los cuales 700 son elegibles. MEB mensual: 180 USD; ingreso medio de los hogares focalizados: 60 USD; regla común: 75 % de la brecha. Los precios aumentan un 15 % en el tercer mes.

Trabajo solicitado. 1) Calcule los errores de inclusión y de exclusión. 2) Calcule el valor de la transferencia. 3) Calcule su valor real tras el aumento y proponga una decisión.

Corrección. Inclusión: 300 / 1 000 = 30 %. Exclusión: (900 − 700) / 900 = 22,2 %. Ambos errores son elevados: el comité probablemente favoreció a ciertos hogares; se necesita una verificación aleatoria, un mecanismo de quejas independiente y una revisión de la lista. Transferencia: brecha 120 USD; 0,75 × 120 = 90 USD. Tras el aumento: 90 / 1,15 = 78,3 USD en poder adquisitivo inicial, es decir, una pérdida del 13 %. El aumento supera el umbral habitual del 10 %: se propone al grupo de trabajo de efectivo una revisión a 90 × 1,15 = 103,5 USD, verificando primero que las propias transferencias no son la causa de la inflación local.

Para recordar

- Error de inclusión = no elegibles asistidos / total asistido; error de exclusión = elegibles no asistidos / total de elegibles.
- La focalización comunitaria debe controlarse mediante una verificación aleatoria y un mecanismo de quejas independiente.
- La elección entre especie, cupones o efectivo depende ante todo del funcionamiento de los mercados y de la necesidad.
- Valor de la transferencia = proporción acordada de (MEB − recursos del hogar), ajustada a la inflación según una regla escrita.
- El monitoreo posdistribución debe desencadenar decisiones, no solo producir un informe.

Cuestionario

1. 500 hogares asistidos, de los cuales 400 son elegibles; 600 hogares elegibles en total. El error de exclusión es:
   A. 20 %
   B. 33,3 %
   C. 16,7 %
   D. 25 %

2. En una zona donde el 85 % de la población tiene necesidades, el método más justo y menos costoso suele ser:
   A. Una puntuación de vulnerabilidad hogar por hogar
   B. Una distribución general con focalización geográfica
   C. Un sorteo
   D. La focalización por parte de las autoridades

3. Una transferencia de 100 USD tras un aumento de precios del 25 % vale, en poder adquisitivo inicial:
   A. 75 USD
   B. 80 USD
   C. 125 USD
   D. 100 USD

4. ¿Qué condición hace inadecuadas las transferencias monetarias?
   A. Hogares con necesidades variadas
   B. Mercados que no funcionan o que no pueden responder a la demanda
   C. Una fuerte presencia del teléfono móvil
   D. Personas que prefieren el efectivo

5. La conciliación mensual de los pagos compara:
   A. El presupuesto y los gastos de la sede
   B. Los hogares registrados, los pagos enviados y los pagos retirados
   C. Los precios de dos mercados
   D. Los informes de dos donantes

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 4 : Planificar y presupuestar un programa: del plan de trabajo al ritmo de gasto

Objetivos

- Construir un plan de trabajo realista: desglose de tareas, dependencias, ruta crítica y holguras.
- Vincular el presupuesto con las actividades mediante costos unitarios, y calcular el costo por persona y el costo por resultado.
- Seguir la ejecución con el ritmo de gasto y detectar a tiempo las subejecuciones y los sobrecostos.

Del marco lógico al plan de trabajo

El marco lógico dice qué; el **plan de trabajo** dice quién hace qué, cuándo y con qué recursos. Se parte de las actividades del marco lógico y se desglosan en tareas de 1 a 4 semanas (estructura de desglose del trabajo): para «distribuir transferencias a 3 000 hogares», hay que registrar, verificar, contratar un proveedor de pagos, formar a los equipos, informar a las comunidades, pagar y hacer el monitoreo posdistribución.

Cada tarea recibe una **duración**, un **responsable** único y sus **dependencias**: la mayoría son del tipo «fin-inicio» (el pago solo puede comenzar después de la firma del contrato). El diagrama de Gantt muestra estas tareas en el tiempo, pero no basta: hay que identificar la ruta crítica.

El plan de trabajo integra también las **limitaciones del contexto**: temporada de lluvias que corta las carreteras, periodo de escasez, fiestas religiosas, elecciones, calendario agrícola y plazos administrativos (autorización de las autoridades, despacho de aduanas, contratación). En muchos contextos, la contratación de un puesto clave lleva de 8 a 12 semanas y la importación de bienes de 3 a 6 meses; un plan que lo ignora es erróneo desde el primer día.

> El primer trimestre de un proyecto casi siempre va con retraso. Planificar los gastos de forma lineal (una doceava parte por mes) garantiza una brecha importante desde el tercer mes.

La ruta crítica y las holguras

La **ruta crítica** es la secuencia más larga de tareas dependientes; su duración es la duración mínima del proyecto. Cualquier tarea de la ruta crítica que se retrase retrasa todo el proyecto. Las demás tareas tienen una **holgura**: pueden desplazarse sin consecuencias, dentro del límite de esa holgura.

Ejemplo. A: evaluación (3 semanas). B: licitación para el proveedor de pagos (6 semanas, después de A). C: registro de los hogares (4 semanas, después de A). D: verificación de las listas (2 semanas, después de C). E: primer pago (1 semana, después de B y D).

- Ruta A-B-E: 3 + 6 + 1 = 10 semanas.
- Ruta A-C-D-E: 3 + 4 + 2 + 1 = 10 semanas.

Las dos rutas duran 10 semanas: ambas son críticas y ninguna tarea tiene holgura. Un retraso de una semana en el registro o en la licitación aplaza el primer pago. Si la licitación durara 4 semanas, la ruta A-B-E sería de 8 semanas y la tarea B tendría 2 semanas de holgura.

Para **acortar** un proyecto, solo se actúa sobre la ruta crítica: acelerar una tarea que tiene holgura no cambia nada. Dos técnicas: hacer en paralelo lo que se hacía en serie (iniciar el registro durante el final de la evaluación, con un riesgo de tener que rehacerlo) y añadir recursos a una tarea crítica (dos equipos de registro en lugar de uno).

El presupuesto del programa por costos unitarios

Un presupuesto sólido se construye **a partir de las actividades**, con costos unitarios verificables: cantidad × costo unitario × duración o frecuencia. Cada línea tiene una justificación (cotización, precio de mercado, escala salarial, histórico). Los donantes rechazan sobre todo los presupuestos con montos redondos, sin base de cálculo o sin vínculo con el marco lógico.

El presupuesto distingue habitualmente: **costos directos de las actividades** (transferencias, insumos, formaciones), **personal** (con la parte del tiempo imputada al proyecto), **costos de funcionamiento de la oficina** compartidos entre proyectos según una clave de reparto documentada, **seguimiento y evaluación**, y **costos indirectos** (porcentaje fijado por el donante, a menudo del 7 al 10 %).

Costos compartidos. Una oficina cuesta 20 000 USD al mes y sirve a tres proyectos. La clave más defendible suele ser el tiempo del personal o la proporción del presupuesto: si el proyecto representa el 40 % de los gastos directos de la oficina, asume 8 000 USD al mes. La clave debe estar escrita, aplicarse de forma constante y ser verificable por un auditor.

> Un presupuesto no es un techo que hay que alcanzar: es una previsión. Gastar al final del proyecto para «consumir» el presupuesto es una falta de gestión que se detecta en las auditorías.

Costo por persona y costo por resultado

Dos indicadores de costo iluminan las decisiones, siempre que se definan bien. El **costo por persona alcanzada** divide el costo total por el número de personas asistidas. Es sencillo pero engañoso: favorece los programas que dan poco a mucha gente y no dice nada del efecto.

El **costo por resultado** divide el costo por el efecto obtenido: costo por niño curado de desnutrición aguda severa, costo por hogar que pasa de un consumo insuficiente a aceptable, costo por punto de agua funcional después de un año. Permite comparar enfoques.

Ejemplo. Proyecto A: 600 000 USD, 4 000 niños admitidos, tasa de curación del 70 %: costo por niño curado = 600 000 / 2 800 = 214 USD. Proyecto B: 450 000 USD, 3 500 niños

admitidos, tasa de curación del 55 %: 450 000 / 1 925 = 234 USD. El proyecto B parece más barato por niño admitido (129 frente a 150 USD), pero es más caro por niño curado.

La **proporción de los costos que llega directamente a las personas** (ratio de transferencia, o «cost-transfer ratio» para el efectivo) mide la eficiencia de la entrega: para un programa de efectivo, monto transferido / costo total. Un ratio de 0,80 significa que 80 centavos de cada dólar llegan a los hogares. Se compara entre programas similares, nunca entre modalidades o contextos muy diferentes.

Seguir la ejecución: el ritmo de gasto

El **ritmo de gasto** (burn rate) compara los gastos reales acumulados con los gastos previstos en el plan de tesorería, y no con una línea recta. Se sigue cada mes, línea por línea, con el responsable financiero y el jefe de proyecto juntos.

Ejemplo. Presupuesto de 12 meses: 900 000 USD. Final del mes 6: gastos previstos acumulados 420 000 USD (el plan tenía en cuenta un arranque lento), gastos reales 290 000 USD. Tasa de ejecución: 290 000 / 420 000 = 69 %. Al ritmo medio actual (290 000 / 6 = 48 333 USD al mes), quedarían 610 000 USD por gastar en 6 meses, es decir, 101 667 USD al mes: más del doble del ritmo actual. Hay que actuar ahora.

Las causas de una subejecución son casi siempre operativas: retrasos de acceso, de contratación o de compras, actividades aplazadas. La respuesta sigue este orden: entender las causas línea por línea; acelerar lo que se pueda sin pérdida de calidad; proponer al donante una revisión presupuestaria o una extensión sin costo **antes** de los últimos meses (la mayoría de los donantes exigen una solicitud al menos 30 a 60 días antes del final); no hacer nunca compras innecesarias de fin de proyecto.

Por el contrario, un sobrecosto en una línea se compensa dentro del límite de **flexibilidad** previsto en el contrato (a menudo del 10 al 20 % entre líneas o rubros sin acuerdo previo). Más allá, el acuerdo escrito del donante es obligatorio; de lo contrario, el gasto corre el riesgo de ser declarado no elegible.

Recursos humanos del programa y organización

El plan de dotación de personal se deriva del plan de trabajo: ¿cuántos encuestadores para registrar 3 000 hogares en 4 semanas? Si un encuestador realiza 12 entrevistas al día durante 20 días laborables al mes, un equipo de 13 encuestadores hace 3 120 entrevistas en 4 semanas. El cálculo evita tanto los equipos demasiado pequeños (retrasos) como los demasiado grandes (costos, calidad).

El **organigrama del programa** aclara la responsabilidad de cada resultado y la separación de funciones: la persona que selecciona a los beneficiarios no debe ser la que valida los pagos ni la que trata las quejas. La matriz **RACI** (realiza, aprueba, consultado, informado) por proceso clave (registro, pago, seguimiento, quejas) hace visibles y controlables estas separaciones.

Por último, el jefe de programa organiza un **ritmo de pilotaje** estable: reunión semanal del equipo sobre el plan de trabajo; revisión mensual de los indicadores, del presupuesto y de los

riesgos; revisión trimestral con la dirección de país, que puede decidir reasignaciones. Un programa sin ritmo de pilotaje descubre sus problemas en el informe final.

Ejercicio: ruta crítica y ritmo de gasto

Situación. Tareas: A, evaluación (2 semanas); B, contratación de los equipos (5 semanas, después de A); C, compra de los kits (8 semanas, después de A); D, formación de los equipos (1 semana, después de B); E, distribución (2 semanas, después de C y D). Presupuesto total: 600 000 USD en 10 meses. Al final del mes 5, gastos previstos acumulados: 260 000 USD; gastos reales: 182 000 USD.

Trabajo solicitado. 1) Calcule la duración mínima y la ruta crítica. 2) ¿Cuál es la holgura de la contratación? 3) Calcule la tasa de ejecución y el ritmo mensual necesario para consumir el presupuesto restante.

Corrección. A-C-E: 2 + 8 + 2 = 12 semanas. A-B-D-E: 2 + 5 + 1 + 2 = 10 semanas. Ruta crítica: A-C-E, 12 semanas; la contratación tiene 2 semanas de holgura. Para adelantar la distribución, hay que actuar sobre la compra de los kits (lanzar la licitación durante la evaluación o utilizar un acuerdo marco). Ejecución: 182 000 / 260 000 = 70 %. Restante: 418 000 USD en 5 meses, es decir, 83 600 USD al mes frente a 36 400 de media hasta ahora. Hay que identificar las líneas retrasadas y hablar desde ahora con el donante sobre una revisión o una extensión sin costo.

Para recordar

- Solo la ruta crítica determina la duración del proyecto; acelerar una tarea que tiene holgura no sirve de nada.
- Un presupuesto se construye por actividades: cantidad × costo unitario × duración, con una justificación por línea.
- El costo por resultado compara mejor los enfoques que el costo por persona alcanzada.
- El ritmo de gasto se compara con el plan de tesorería, no con una línea recta, y se sigue cada mes.
- Revisión presupuestaria o extensión sin costo: se solicitan pronto, nunca mediante compras de fin de proyecto.

Cuestionario

1. A (2 sem.), luego B (4 sem.) y C (3 sem.) en paralelo después de A, luego D (1 sem.) después de B y C. Duración mínima del proyecto:
   A. 10 semanas
   B. 7 semanas
   C. 6 semanas
   D. 9 semanas
2. Proyecto de 300 000 USD, 1 500 niños admitidos, tasa de curación del 80 %. Costo por niño curado:
   A. 200 USD
   B. 250 USD
   C. 240 USD
   D. 160 USD
3. Para acortar un proyecto, se actúa primero sobre:
   A. Las tareas que tienen más holgura
   B. Las tareas de la ruta crítica
   C. El presupuesto de los costos indirectos
   D. El informe final
4. Un programa de efectivo transfiere 720 000 USD con un costo total de 900 000 USD. El ratio de transferencia es:
   A. 0,80
   B. 1,25
   C. 0,72
   D. 0,20
5. A cuatro meses del final, el 40 % del presupuesto no se ha gastado debido a retrasos de acceso. La mejor respuesta es:
   A. Comprar material para consumir el presupuesto
   B. Analizar las causas línea por línea y solicitar pronto al donante una revisión o una extensión sin costo
   C. No decir nada al donante
   D. Devolver el dinero sin explicación

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. A · 5. B

Semana 5 : El sistema de seguimiento: plan MEAL, calidad de los datos y seguimiento a distancia

Objetivos

- Construir un plan MEAL completo y proporcionado, desde el formulario de recopilación hasta la decisión de gestión.
- Evaluar y mejorar la calidad de los datos según cinco dimensiones, y organizar controles de calidad periódicos.
- Organizar el seguimiento en las zonas inaccesibles: seguimiento a distancia, monitoreo por terceros y triangulación.

El plan MEAL: la columna vertebral del seguimiento

El **plan MEAL** describe cómo se medirá cada indicador, por quién, cuándo y para qué decisión. Se presenta en forma de tabla: indicador, definición, línea de base, meta, fuente y método, frecuencia, desagregación, responsable de la recopilación, responsable del análisis, uso (revisión mensual, informe al donante, ajuste de la focalización).

Abarca cuatro tipos de seguimiento:

- **Seguimiento de las actividades y de los productos**: lo que se ha realizado, en comparación con el plan (registros de distribución, listas de asistencia, fichas de obra).
- **Seguimiento de los efectos**: cambios en las personas (encuestas de línea de base, intermedias y finales; monitoreo posdistribución).
- **Seguimiento del contexto**: precios, acceso, seguridad, movimientos de población e indicadores de los supuestos críticos del marco lógico.
- **Seguimiento de la rendición de cuentas**: quejas y retroalimentación recibidas, plazo de respuesta, satisfacción, participación (semana 8).

El costo del plan MEAL debe figurar en el presupuesto: personal, herramientas, transporte, encuestas, evaluaciones. Un rango del 3 al 10 % del presupuesto del proyecto es habitual; un proyecto piloto cuyo efecto se quiere demostrar puede justificar más.

> La mejor prueba de un plan MEAL: para cada indicador, ¿puede nombrar la reunión en la que se discutirá y la decisión que puede desencadenar? Si no, pregúntese si hay que recopilarlo.

De las herramientas de recopilación al tablero de control

La cadena de datos tiene cinco eslabones: **recopilación**, **entrada y almacenamiento**, **limpieza**, **análisis**, **difusión**. La recopilación móvil (KoboToolbox, ODK, CommCare) ha sustituido al papel en la mayoría de los contextos: controles automáticos en la entrada (rangos

de valores, campos obligatorios, saltos lógicos), geolocalización y marca de tiempo, envío directo al servidor.

Un buen formulario se diseña a partir del plan de análisis: primero se escriben las tablas que se quieren producir, luego las preguntas necesarias y solo esas. Se prueba en el terreno con encuestados reales antes de desplegarlo y luego se congela una versión; cualquier modificación posterior se documenta, porque hace que los datos de antes y de después sean incomparables.

La **limpieza** sigue reglas escritas: duplicados, valores atípicos (un hogar de 45 personas, un gasto en alimentos de 2 000 USD por semana), incoherencias (un niño de 3 años casado), entrevistas demasiado cortas (menos de 10 minutos para un cuestionario de 40 minutos). Nunca se corrige un valor sin fuente; se señala, se verifica con el encuestador o se retira, y se lleva un registro de limpieza.

El **tablero de control** muestra pocos indicadores, con la meta, la tendencia y un código de colores según umbrales fijados de antemano. Se actualiza al ritmo de las decisiones: semanal para las distribuciones, mensual para los indicadores de efecto intermedio.

La calidad de los datos: cinco dimensiones

Una **evaluación de la calidad de los datos** (DQA) verifica cinco dimensiones:

Dimensión	Pregunta	Ejemplo de problema
Validez	¿Miden los datos lo que prevé el indicador?	Se cuentan las personas formadas en lugar de las que aplican la práctica.
Fiabilidad	¿Son estables los métodos entre lugares, personas y periodos?	Dos oficinas calculan la tasa de curación de forma diferente.
Precisión	¿Es el margen de error lo bastante bajo para la decisión?	Una muestra de 50 hogares para seguir una variación de 5 puntos.
Puntualidad	¿Llegan los datos a tiempo para decidir?	El informe mensual llega seis semanas después del final del mes.
Integridad	¿Están los datos protegidos contra los errores y las manipulaciones?	Un mismo agente introduce, valida y corrige las listas de pago.

El método más concreto es la **verificación por recuento**: se elige una muestra de sitios, se recuenta a partir de las fuentes primarias (registros, fichas) y se compara con las cifras reportadas. Un **ratio de verificación** de 1,00 es perfecto; se fija una tolerancia, por ejemplo entre 0,95 y 1,05. Por encima de 1,05, las cifras reportadas están infladas; por debajo de 0,95, hay actividades que no se reportan.

Ejemplo. Un centro de salud reporta 1 240 consultas; el recuento de los registros encuentra 1 085. Ratio: 1 240 / 1 085 = 1,14. El centro sobredeclara un 14 %: hay que entender la causa (doble conteo de las visitas de seguimiento, presión sobre las metas) antes de cualquier sanción.

Los controles de calidad de la recopilación

La calidad se juega durante la recopilación, no después. Los controles estándar:

- **Formación y prueba** de los encuestadores, con un examen escrito y entrevistas simuladas; los encuestadores que no lo superan no recopilan datos.
- **Supervisión directa** de una parte de las entrevistas (por ejemplo, el 5 %).
- **Verificaciones posteriores** (back-checks): el 10 % de los hogares se vuelve a entrevistar sobre algunas preguntas estables (tamaño del hogar, tipo de refugio). Una tasa de concordancia baja para un encuestador indica errores o entrevistas inventadas.
- **Controles diarios a distancia**: duración de las entrevistas, coordenadas GPS (entrevistas todas en el mismo lugar), hora (entrevistas a las 3 de la madrugada), distribución de las respuestas por encuestador (un encuestador cuyos hogares tienen todos exactamente 5 miembros).

Los datos inventados («curbstoning») son más frecuentes de lo que se cree cuando los encuestadores cobran por entrevista, están sometidos a presión por cuotas o se les envía a zonas peligrosas. Prevenirlos pasa por objetivos realistas, un pago por jornada en lugar de por unidad y controles anunciados de antemano.

> Un responsable de MEAL que nunca encuentra errores no busca lo suficiente. Publicar internamente los errores encontrados y corregidos refuerza la confianza en las cifras.

El seguimiento a distancia y el monitoreo por terceros

En las zonas inaccesibles para el personal (inseguridad, restricciones), el **seguimiento a distancia** combina varias fuentes para compensar la ausencia de visitas:

- **Socios locales** o personal local que tienen acceso, con herramientas estandarizadas y pruebas (fotos geolocalizadas, listas firmadas).
- **Monitoreo por terceros** (TPM): una empresa u organización independiente, a menudo exigida por los donantes, que verifica una muestra de actividades.
- **Llamadas telefónicas** a una muestra aleatoria de beneficiarios (monitoreo posdistribución por teléfono).
- **Datos remotos**: imágenes satelitales (campamentos, cultivos, destrucción), datos de transacciones de los operadores móviles, datos de precios.
- **Retroalimentación de la comunidad** por línea gratuita y mensajería.

El seguimiento a distancia tiene **límites**: muestras telefónicas sesgadas hacia los hogares con teléfono y a menudo masculinas, riesgos de presión sobre los socios, transferencia de los riesgos de seguridad al personal local. Hay que decirlo en los informes, pagar correctamente a los socios por este trabajo y no pedirles que asuman riesgos que la organización rechazaría para su propio personal.

La **triangulación** es la regla: una distribución se considera verificada cuando al menos dos fuentes independientes concuerdan (por ejemplo, lista firmada y llamadas a una muestra de beneficiarios).

Hacer útil el seguimiento: del dato a la decisión

El seguimiento rara vez fracasa por falta de datos; fracasa porque nadie los utiliza. Tres mecanismos hacen útil el seguimiento:

- **Umbrales de alerta** fijados de antemano: si el tiempo de espera en las distribuciones supera las 3 horas, si la tasa de quejas no resueltas supera el 20 %, si el precio de la canasta aumenta más del 10 %, se desencadena una acción predefinida.
- **Revisiones periódicas** en las que el equipo de programa, MEAL, logística y finanzas miran juntos las mismas cifras y deciden. El acta registra las decisiones, su responsable y su plazo.
- **Una devolución a las personas y a los equipos**: presentar a las comunidades lo que se ha aprendido y cambiado; mostrar a los encuestadores para qué han servido sus datos.

La independencia del MEAL importa. Si el responsable de MEAL rinde cuentas únicamente al jefe de proyecto cuyos resultados mide, la tentación de suavizar las malas cifras es fuerte. Una doble línea jerárquica (operativa con el proyecto, técnica con un responsable de MEAL de país o de sede) protege la calidad de la información.

Ejercicio: verificar la calidad de los datos de tres sitios

Situación. Un proyecto de nutrición reporta, para un mes, las admisiones de tres sitios. Sitio 1: reportado 320, recontado 312. Sitio 2: reportado 410, recontado 342. Sitio 3: reportado 198, recontado 205. Tolerancia: ratio entre 0,95 y 1,05. Las verificaciones posteriores de la encuesta de cobertura muestran una tasa de concordancia del 96 % para todos los encuestadores salvo uno (61 %), cuyas entrevistas duran de media 9 minutos frente a 35 para los demás.

Trabajo solicitado. 1) Calcule los ratios de verificación e interprételos. 2) ¿Qué hace con los datos del encuestador sospechoso? 3) ¿Qué medidas de prevención propone?

Corrección. Sitio 1: 320 / 312 = 1,03, dentro de la tolerancia. Sitio 2: 410 / 342 = 1,20: sobredeclaración del 20 %, que hay que investigar (doble conteo, presión sobre las metas, fraude). Sitio 3: 198 / 205 = 0,97, dentro de la tolerancia (ligera subdeclaración). Total reportado 928 frente a 859 recontados: hay que corregir el informe mensual e informar al donante. Los datos del encuestador sospechoso se apartan y una parte de sus hogares se vuelve a entrevistar; si se confirma la falsificación, sus entrevistas se rehacen o se retiran, y se trata la causa (cuotas, pago por entrevista). Prevención: pago por jornada, controles diarios de duración y de GPS, verificaciones posteriores anunciadas, objetivos realistas.

Para recordar

- El plan MEAL vincula cada indicador con un método, un responsable, una frecuencia y una decisión.
- Cinco dimensiones de la calidad de los datos: validez, fiabilidad, precisión, puntualidad, integridad.
- Ratio de verificación = cifra reportada / cifra recontada; fuera de la tolerancia, se investiga antes de sancionar.
- Las verificaciones posteriores, la duración de las entrevistas y el GPS detectan los datos inventados.
- En el seguimiento a distancia, una actividad está verificada cuando dos fuentes independientes concuerdan.

Cuestionario

1. Un sitio reporta 500 distribuciones; el recuento encuentra 400. El ratio de verificación es:
   A. 0,80
   B. 1,25
   C. 1,20
   D. 0,90

2. Dos oficinas calculan de forma diferente el mismo indicador. ¿Qué dimensión de la calidad está en juego?
   A. Puntualidad
   B. Fiabilidad
   C. Integridad
   D. Precisión

3. Un encuestador realiza entrevistas de 8 minutos para un cuestionario de 35 minutos, todas en el mismo punto GPS. Lo más probable:
   A. Es muy eficiente
   B. Entrevistas inventadas o hechas de prisa
   C. Una avería del servidor
   D. Una muestra demasiado grande

4. ¿Por qué evitar pagar a los encuestadores por entrevista?
   A. Es más caro
   B. Incita a hacer las entrevistas de prisa o a inventarlas
   C. Los donantes siempre lo prohíben
   D. Reduce el tamaño de la muestra

5. ¿Cuál es el principal límite del monitoreo posdistribución por teléfono?
   A. Es demasiado caro
   B. La muestra está sesgada hacia los hogares que tienen teléfono, a menudo representados por hombres
   C. Está prohibido
   D. No permite ninguna pregunta

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 6 : Encuestas y muestreo: obtener cifras correctas

Objetivos

- Elegir un diseño muestral (aleatorio simple, sistemático, estratificado, por conglomerados) adaptado a la pregunta y al terreno.
- Calcular un tamaño de muestra con margen de error, corrección por población finita, efecto de diseño y no respuestas.
- Diseñar un cuestionario fiable e interpretar correctamente una proporción y su intervalo de confianza.

Población, muestra y representatividad

Una encuesta busca conocer una **población** (todos los hogares desplazados de un distrito) a partir de una **muestra**. Los resultados solo son generalizables si cada unidad de la población tiene una **probabilidad conocida y no nula** de ser seleccionada: es el muestreo probabilístico. Una muestra «de conveniencia» (los hogares cercanos a la carretera, los presentes en la distribución) puede ser grande, pero sigue estando sesgada.

Primero se necesita un **marco muestral**: lista de los hogares, de los sitios, de las aldeas, con su población. Sin una lista fiable, se trabaja por etapas: se seleccionan aldeas o bloques de un campamento y luego hogares dentro de ellos, enumerándolos sobre el terreno.

Los principales diseños:

- **Aleatorio simple**: sorteo en la lista completa. El más sencillo estadísticamente, a menudo impracticable en el terreno.
- **Sistemático**: se calcula un intervalo (población / tamaño de la muestra), se sortea un punto de partida y luego se toma un hogar cada «intervalo». Para 2 400 refugios y 240 hogares que entrevistar, intervalo = 10; punto de partida sorteado entre 1 y 10, por ejemplo 7: refugios 7, 17, 27…
- **Estratificado**: la población se divide en grupos (desplazados y de acogida, zonas urbanas y rurales) y se selecciona una muestra en cada grupo. Indispensable si se quieren resultados fiables para cada grupo.
- **Por conglomerados**: primero se seleccionan conglomerados (aldeas) y luego hogares en cada conglomerado. Menos costoso en desplazamientos, pero menos preciso: se necesita una muestra más grande.

Calcular el tamaño de la muestra

Para estimar una proporción, la fórmula básica es: $\mathbf{n = z^2 \times p \times (1 - p) / e^2}$, donde z vale 1,96 para un nivel de confianza del 95 %, p es la proporción esperada (0,5 cuando no se sabe, porque es el caso más exigente) y e el margen de error.

- Margen del 5 %: $1{,}96^2 \times 0{,}25 / 0{,}0025 = 384$ hogares.
- Margen del 7 %: 196 hogares.
- Margen del 10 %: 96 hogares.

Dividir el margen por dos multiplica la muestra por cuatro. La elección del margen depende de la decisión: para seguir una variación de 10 puntos entre dos encuestas, un margen del 10 % no basta.

Corrección por población finita. Cuando la población es pequeña, la muestra necesaria disminuye: $n = n_0 / (1 + (n_0 - 1) / N)$. Para N = 2 000 hogares: 384 / (1 + 383 / 2 000) = 384 / 1,19 = 323 hogares. Para N = 100 000, la corrección es insignificante.

Efecto de diseño. En el muestreo por conglomerados, los hogares de una misma aldea se parecen, lo que reduce la información que aporta cada uno. Se multiplica la muestra por el efecto de diseño: **DEFF = 1 + (m − 1) × ICC**, donde m es el número de hogares por conglomerado e ICC el coeficiente de correlación intraconglomerado. Con 10 hogares por conglomerado y un ICC de 0,05: DEFF = 1 + 9 × 0,05 = 1,45. A falta de datos, se suele utilizar de 1,5 a 2.

No respuestas. Se divide por la tasa de respuesta esperada: con un 10 % de no respuestas, n / 0,9.

> Cálculo completo: margen del 5 %, p = 0,5, DEFF 1,5, no respuestas del 10 %: 384 × 1,5 = 576; 576 / 0,9 = 640 hogares, por ejemplo 64 conglomerados de 10 hogares.

Estratificación y comparaciones

Si se quieren resultados para cada estrato con la misma precisión, se necesita el tamaño de muestra **en cada estrato**. Para tres distritos con un margen del 5 % cada uno, la muestra es tres veces mayor que para una estimación global. Esta suele ser la principal causa de las encuestas demasiado caras: se prometen resultados por distrito, por sexo del jefe de hogar y por estatus, sin calcular la muestra necesaria.

Cuando los estratos tienen tamaños diferentes y se quiere un resultado global, se **pondera**: cada estrato cuenta según su peso en la población y no según su peso en la muestra. **Ejemplo.** Desplazados: 20 000 hogares, muestra de 400, 60 % en inseguridad alimentaria. Comunidad de acogida: 80 000 hogares, muestra de 400, 30 % en inseguridad alimentaria. Media no ponderada: 45 %. Media ponderada: 0,2 × 60 + 0,8 × 30 = 36 %. Presentar el 45 % sobrestimaría mucho la situación.

Para **comparar dos encuestas** (línea de base y final) o dos grupos, se necesita una muestra capaz de detectar la diferencia esperada: detectar un paso del 40 % al 50 % con un 95 % de confianza y un 80 % de potencia requiere unos 385 hogares por grupo en muestreo aleatorio simple, más en conglomerados. Un cambio pequeño exige una muestra grande.

Diseñar un cuestionario fiable

Un cuestionario se construye a partir de los indicadores y del plan de análisis. Reglas esenciales:

- **Una idea por pregunta**: «¿Tiene acceso a agua y a jabón?» son dos preguntas.
- **Periodo de recuerdo definido**: «En los últimos 7 días» para la alimentación, «en los últimos 30 días» para los gastos. Un periodo demasiado largo produce errores de memoria.
- **Módulos estándar** para los indicadores estándar (puntuación de consumo de alimentos, rCSI, Grupo de Washington), sin modificarlos, para seguir siendo comparables.
- **Formulación neutra**: «¿Qué opina de la distribución?» en lugar de «¿Está satisfecho con la distribución?».
- **Orden**: preguntas sencillas al principio, sensibles al final, y nunca una pregunta sensible sin formación de los encuestadores y sin posibilidad de derivación.
- **Traducción y retrotraducción** a las lenguas locales, y luego prueba en el terreno.
- **Duración**: más allá de 45 minutos, la calidad de las respuestas disminuye; suprimir lo que no sea indispensable.

> El consentimiento informado se lee a cada persona encuestada: objeto de la encuesta, duración, carácter voluntario, ausencia de vínculo con la ayuda, confidencialidad. Una negativa se respeta y se sustituye según la regla prevista, nunca por el vecino más práctico.

Interpretar una proporción y su intervalo de confianza

Un resultado de encuesta es una **estimación**, rodeada de una incertidumbre. El intervalo de confianza al 95 % de una proporción es aproximadamente: $p \pm 1{,}96 \times \sqrt{(p \times (1 - p) / n)}$, multiplicado por √DEFF en conglomerados.

Ejemplo. 384 hogares en muestreo aleatorio simple, 45 % sin acceso suficiente al agua. Error estándar: $\sqrt{(0{,}45 \times 0{,}55 / 384)} = 0{,}0254$. Margen: $1{,}96 \times 0{,}0254 = 5{,}0$ puntos. Intervalo: 40 % a 50 %.

Consecuencias prácticas:

- Si la línea de base era del 45 % (intervalo 40-50 %) y la final del 49 % (intervalo 44-54 %), no se puede afirmar que la situación haya empeorado.
- Los resultados por subgrupo tienen intervalos más amplios, porque los subgrupos son más pequeños: un 45 % sobre 80 hogares encabezados por una mujer tiene un margen de unos 11 puntos.
- Siempre se presenta el tamaño de la muestra, y el intervalo para los indicadores clave.

La **significación estadística** no significa importancia práctica: con una muestra muy grande, una diferencia de 1 punto puede ser significativa sin cambiar ninguna decisión. A la inversa, una diferencia importante puede no ser significativa con una muestra pequeña: se dice entonces que los datos no permiten concluir, no que no haya efecto.

Organizar una encuesta en el terreno

Una encuesta de hogares se prepara en semanas, no en días. El calendario tipo:

- Semana 1: objetivos, indicadores, plan de análisis, diseño muestral, autorización de las autoridades.
- Semana 2: cuestionario, traducción, programación en tableta, pruebas.
- Semana 3: contratación y formación de los encuestadores (3 a 5 días, con examen y práctica), prueba piloto, correcciones.
- Semanas 4 y 5: recopilación, con controles de calidad diarios y reuniones de balance.
- Semana 6: limpieza, ponderación, análisis, validación con el equipo y los socios.
- Semana 7: informe y devolución, incluida la devolución a las comunidades.

Logística. Número de equipos = tamaño de la muestra / (entrevistas por equipo y por día × días de recopilación). Para 640 hogares, 4 entrevistas por encuestador y por día, 10 días: 640 / 40 = 16 encuestadores, es decir, 4 equipos de 4 con un supervisor cada uno. Se prevé entre un 10 y un 20 % de días de reserva para los imprevistos (lluvia, seguridad, acceso).

La **sustitución** de los hogares ausentes o que se niegan sigue una regla escrita (volver una vez y luego sustituir por el hogar siguiente en la lista seleccionada); de lo contrario, los encuestadores sustituyen los hogares alejados por hogares cercanos, lo que sesga la muestra.

Ejercicio: dimensionar una encuesta de línea de base

Situación. Un proyecto quiere conocer, con un margen del 5 % y un 95 % de confianza, la proporción de hogares en inseguridad alimentaria, por separado para los desplazados (3 000 hogares) y para la comunidad de acogida (40 000 hogares). Muestreo por conglomerados de 12 hogares, ICC estimado en 0,06. No respuestas esperadas: 8 %.

Trabajo solicitado. 1) Calcule la muestra básica con la corrección por población finita para cada estrato. 2) Calcule el efecto de diseño y la muestra final por estrato. 3) ¿Cuántos conglomerados se necesitan en total?

Corrección. Base: 384. Desplazados: 384 / (1 + 383 / 3 000) = 384 / 1,128 = 340. Acogida: 384 / (1 + 383 / 40 000) = 384 / 1,0096 = 380. Efecto de diseño: 1 + 11 × 0,06 = 1,66. Desplazados: 340 × 1,66 = 565; 565 / 0,92 = 614. Acogida: 380 × 1,66 = 631; 631 / 0,92 = 686. Total: 1 300 hogares. Conglomerados: 614 / 12 = 51,2, es decir, 52 conglomerados; 686 / 12 = 57,2, es decir, 58 conglomerados; total 110 conglomerados. Para un resultado global, los dos estratos se ponderarán según su peso en la población (7 % y 93 %).

Para recordar

- $n = z^2 \times p \times (1 - p) / e^2$: 384 para un margen del 5 % con un 95 % de confianza y p = 0,5.
- Corrección para poblaciones pequeñas: $n = n_0 / (1 + (n_0 - 1) / N)$.
- Efecto de diseño por conglomerados: $1 + (m - 1) \times ICC$; luego se divide por la tasa de respuesta.
- Un resultado por estrato exige la muestra en cada estrato; un resultado global exige una ponderación.
- Presentar siempre el tamaño de la muestra y el intervalo de confianza; no significativo no quiere decir ausencia de efecto.

Cuestionario

1. Con p = 0,5, un 95 % de confianza y un margen del 10 %, el tamaño de una muestra aleatoria simple es de aproximadamente:
   A. 38
   B. 96
   C. 196
   D. 384

2. Conglomerados de 20 hogares, ICC de 0,04. El efecto de diseño es:
   A. 1,04
   B. 1,76
   C. 1,80
   D. 0,80

3. Estrato A: 10 000 hogares, 50 % afectados. Estrato B: 30 000 hogares, 20 % afectados. La proporción global ponderada es:
   A. 35 %
   B. 27,5 %
   C. 30 %
   D. 25 %

4. Línea de base: 52 % (intervalo 47-57 %). Final: 55 % (intervalo 50-60 %). Conclusión más correcta:
   A. La situación ha mejorado claramente
   B. Los datos no permiten concluir que haya habido un cambio
   C. La situación ha empeorado
   D. La encuesta es errónea

5. Para 2 000 refugios y 200 hogares que entrevistar en muestreo sistemático, el intervalo es:
   A. 5
   B. 10
   C. 20
   D. 200

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 7 : Analizar los datos: estadísticas útiles, trampas y visualización

Objetivos

- Elegir las estadísticas descriptivas adecuadas y desagregar para revelar las desigualdades ocultas por las medias.
- Evitar las trampas clásicas: correlación y causalidad, paradoja de Simpson, denominadores engañosos.
- Analizar datos cualitativos con rigor y presentar los resultados para que se entiendan y se utilicen.

Estadísticas descriptivas: media, mediana, dispersión

La **media** es sensible a los valores extremos; la **mediana** (valor que divide la muestra en dos mitades) no lo es. Para los ingresos, los gastos, las deudas o los tiempos de espera, distribuciones muy asimétricas, la mediana describe mejor el hogar típico.

Ejemplo. Deudas de diez hogares (USD): 0, 0, 20, 30, 40, 50, 60, 80, 100, 1 620. Media: 2 000 / 10 = 200 USD. Mediana: (40 + 50) / 2 = 45 USD. Decir «los hogares tienen de media 200 USD de deudas» describe un solo hogar, no la situación.

La **dispersión** cuenta tanto como el centro: desviación estándar, o rango intercuartílico (entre el percentil 25 y el 75) para las distribuciones asimétricas. Dos programas con un tiempo de espera medio de 2 horas son muy diferentes si uno tiene esperas entre 1 h 30 y 2 h 30, y el otro entre 10 minutos y 7 horas.

Para los indicadores expresados en proporción, se presentan el numerador y el denominador: «62 % (248 de 400)». Los porcentajes calculados sobre efectivos muy pequeños («el 50 % de las personas mayores», es decir, 2 de 4) deben evitarse o señalarse.

> Variación en puntos y en porcentaje: pasar del 20 % al 30 % es un aumento de 10 puntos porcentuales, es decir, un aumento relativo del 50 %. Confundir ambos es el error más frecuente de los informes.

Desagregar para ver las desigualdades

Una media global satisfactoria puede ocultar un grupo que se queda atrás. La **desagregación** (por sexo y edad, discapacidad, estatus, zona, tipo de hogar) es el medio para verificar que el programa llega a todo el mundo.

Ejemplo. Tasa de satisfacción global: 82 %. Por grupo: hogares sin ninguna persona con discapacidad 86 %; hogares con una persona con discapacidad 54 %. El programa no funciona para un grupo entero, y la cifra global lo ocultaba.

La **tabla cruzada** es la herramienta básica: un indicador en fila, un grupo en columna. Para saber si la diferencia observada entre dos grupos supera lo que el azar del muestreo puede producir, se utiliza una prueba (prueba de chi cuadrado para proporciones, prueba t o prueba no paramétrica para medias) o se comparan los intervalos de confianza. Un programa informático lo calcula; el papel del responsable es plantear la pregunta correcta e interpretar.

Atención a las **comparaciones múltiples**: si se prueban 20 diferencias con un umbral del 5 %, cabe esperar encontrar una «significativa» por puro azar. Por eso se definen de antemano las comparaciones principales del plan de análisis, y las demás se presentan como exploratorias.

Correlación, causalidad y paradoja de Simpson

Dos variables que evolucionan juntas no prueban que una cause la otra. Siempre existen tres explicaciones alternativas: el **azar**, un **factor común** (variable de confusión) y la **causalidad inversa**.

Ejemplo. Los hogares que participaron en las formaciones agrícolas tienen mejores rendimientos. Pero también son los que tienen más tierras y herramientas, y los que estaban más motivados para inscribirse: la formación quizá tenga un efecto, pero la comparación simple no permite separarlo de esas diferencias iniciales. Medir un efecto exige un contrafactual (semana 10).

La **paradoja de Simpson** se produce cuando una tendencia presente en cada grupo se invierte al agrupar los datos, debido a una distribución desigual entre grupos.

	Centro A	**Centro B**
Casos severos	60 curados de 100 (60 %)	110 curados de 200 (55 %)
Casos moderados	270 curados de 300 (90 %)	170 curados de 200 (85 %)
Total	330 de 400 (82,5 %)	280 de 400 (70 %)

Aquí, A es mejor en cada grupo y en el total: no hay paradoja. Pero si B recibiera sobre todo casos moderados y A sobre todo casos severos, B podría parecer mejor en el total siendo peor en cada grupo. Regla: comparar desempeños solo con una **composición comparable**, o desagregando por el factor que cambia la dificultad.

Analizar los datos cualitativos

Las entrevistas y los grupos focales explican el «porqué» detrás de las cifras. Su análisis sigue un método, sin el cual solo se retienen las citas que confirman lo que ya se pensaba.

- **Transcripción o notas detalladas** el mismo día, con las palabras de las personas.
- **Codificación**: se asignan códigos (temas) a los fragmentos. Se parte de códigos vinculados a las preguntas de evaluación y se añaden los que emergen.
- **Matriz**: temas en fila, grupos (mujeres, hombres, jóvenes, desplazados, comunidad de acogida) en columna. Así se ve qué temas aparecen en qué grupos.

- **Búsqueda activa de contraejemplos**: los casos que contradicen la tendencia general suelen ser los más instructivos.
- **Saturación**: se detiene la recopilación cuando las nuevas entrevistas ya no aportan temas nuevos.

Los resultados cualitativos no se presentan en porcentajes, pero se puede indicar la frecuencia: «este tema apareció en 9 de los 12 grupos focales, incluidos todos los grupos de mujeres». Una cita ilustra un resultado, no lo prueba.

> Métodos mixtos: la cifra dice «el 30 % de los hogares no retira su transferencia»; lo cualitativo dice «el punto de retiro está a 3 horas a pie y los agentes a menudo no tienen liquidez». Ambos juntos llevan a una decisión.

Visualizar para hacer comprender

Un gráfico debe responder a una pregunta en pocos segundos. Reglas prácticas:

- **Barras horizontales** para comparar categorías, ordenadas por valor; **líneas** para las evoluciones en el tiempo; **mapas** para las diferencias entre zonas; evitar los gráficos circulares de más de tres categorías y los efectos 3D.
- El eje de las barras empieza en **cero**; si no, una diferencia de 2 puntos parece enorme.
- Un **título que diga el mensaje** («La cobertura ha bajado en las zonas norte desde marzo») en lugar de un título neutro («Cobertura por zona»).
- Pocos colores, un color de acento para lo que importa y una referencia para la meta.
- La fuente, la fecha y el tamaño de la muestra al pie del gráfico.

Para un tablero de control de dirección, se limita a 6 a 10 indicadores clave, con para cada uno el valor actual, la meta, la tendencia y un código de colores cuyos umbrales están definidos. Un tablero de 40 indicadores no lo lee nadie.

En un mapa, se representan **tasas** y no cifras absolutas cuando las zonas tienen poblaciones diferentes: una gran ciudad siempre tendrá más casos que una aldea.

Del resultado a la recomendación

Un análisis útil sigue una progresión en tres tiempos: **constatación** (lo que dicen los datos), **interpretación** (por qué, con los datos cualitativos y el contexto) y **recomendación** (lo que hay que hacer, quién y cuándo). Las recomendaciones son pocas, precisas y realistas, y cada una se deriva de una constatación.

También hay que indicar el **nivel de confianza**: alto (varias fuentes concordantes, muestra suficiente), medio, bajo (una sola fuente, muestra pequeña). Una dirección puede decidir sobre una evidencia media en una emergencia, pero debe saber que lo hace.

Por último, la **devolución** de los resultados se hace antes del informe definitivo: una sesión con el equipo y los socios permite corregir los errores de interpretación y construir la adhesión

a las recomendaciones. Los resultados también se comparten con las comunidades, en una forma adaptada (reunión, cartel, mensaje de voz).

Ejercicio: analizar un monitoreo posdistribución

Situación. Monitoreo posdistribución de 400 hogares. Tiempo de espera: media 2 h 40, mediana 1 h 20. Satisfacción global: 78 %. Hogares encabezados por una mujer (120): 60 % satisfechos; encabezados por un hombre (280): 85,7 % satisfechos. La puntuación media de consumo de alimentos pasó de 31 a 38; la proporción de hogares con un consumo aceptable pasó del 30 % al 45 %. En los grupos focales, 8 de 10 grupos de mujeres mencionan la hora de distribución (después de las 15 h) y el trayecto de vuelta.

Trabajo solicitado. 1) ¿Qué dice la diferencia entre la media y la mediana del tiempo de espera? 2) Exprese la variación del consumo aceptable en puntos y en porcentaje. 3) ¿Qué conclusión y qué recomendación extrae sobre la satisfacción?

Corrección. Una media muy superior a la mediana indica que la mayoría espera poco, pero que un grupo espera muchísimo: hay que identificar qué sitios o qué días. Consumo aceptable: +15 puntos porcentuales, es decir, +50 % en términos relativos (no se puede atribuir este cambio únicamente al programa sin contrafactual). Satisfacción: verificación, 0,3 × 60 + 0,7 × 85,7 = 78 %. La diferencia de 26 puntos entre hogares encabezados por una mujer y por un hombre es importante y la explican los grupos focales: horarios tardíos y trayecto de vuelta peligroso. Recomendación: distribuciones por la mañana, fila prioritaria, puntos de distribución más cercanos y verificación en el siguiente monitoreo.

Para recordar

- Mediana en lugar de media para las distribuciones asimétricas (ingresos, deudas, tiempos de espera).
- Una variación del 20 % al 30 % es +10 puntos, es decir, +50 % en términos relativos.
- Desagregar revela los grupos que se quedan atrás y que la media oculta.
- Correlación no es causalidad; comparar desempeños con una composición comparable.
- Análisis cualitativo: codificación, matriz por grupo, búsqueda de contraejemplos; una cita ilustra, no prueba.

Cuestionario

1. Gastos de 5 hogares: 10, 15, 20, 25, 430 USD. Media y mediana:
   A. Media 100, mediana 20
   B. Media 20, mediana 100
   C. Media 25, mediana 20
   D. Media 100, mediana 25

2. Una cobertura pasa del 40 % al 50 %. Es:
   A. +10 % en términos relativos
   B. +10 puntos porcentuales, es decir, +25 % en términos relativos
   C. +50 puntos
   D. +25 puntos

3. Se prueban 40 diferencias entre grupos con un umbral del 5 %. ¿Cuántos resultados «significativos» cabe esperar solo por azar?
   A. Ninguno
   B. Unos 2
   C. Unos 20
   D. Exactamente 5

4. En un mapa que compara distritos con poblaciones muy diferentes, hay que representar:
   A. Las cifras absolutas de casos
   B. Tasas (por ejemplo, por cada 1 000 habitantes)
   C. Solo los nombres de los distritos
   D. Los presupuestos de las ONG

5. Un tema aparece en 9 de los 12 grupos focales. La presentación correcta es:
   A. El 75 % de la población piensa eso
   B. Este tema apareció en 9 de los 12 grupos, precisando cuáles
   C. Todo el mundo piensa eso
   D. Este resultado es estadísticamente significativo

Respuestas del cuestionario : 1. A · 2. B · 3. B · 4. B · 5. B

Semana 8 : Rendición de cuentas a las poblaciones: participación, quejas y retroalimentación

Objetivos

- Traducir la Norma Humanitaria Esencial (CHS) en prácticas concretas y medibles en un programa.
- Diseñar y dirigir un mecanismo de quejas y retroalimentación accesible, seguro y que cierre el ciclo.
- Organizar la participación real de las personas y una información transparente, incluso sobre las decisiones difíciles.

Lo que significa rendir cuentas a las personas

La **rendición de cuentas a las poblaciones afectadas** (AAP) es el compromiso de usar el propio poder de forma responsable: tener en cuenta la opinión de las personas, rendirles cuentas y responder ante ellas de los propios actos. No sustituye a la rendición de cuentas ante los donantes; la complementa, y ambas entran a menudo en tensión (un donante quiere cifras, las personas quieren ser escuchadas).

La **Norma Humanitaria Esencial en materia de Calidad y Rendición de Cuentas** (CHS), revisada en 2024, fija nueve compromisos con las personas y las comunidades. En esencia: una ayuda que respeta sus derechos y su participación, accesible a tiempo, pertinente, que refuerza sus capacidades y no causa daño, con quejas tratadas de forma segura, una ayuda coordinada, una organización que aprende, un personal competente y bien gestionado, y recursos utilizados de forma responsable.

Para cada compromiso, una organización puede plantearse tres preguntas: ¿qué hacemos concretamente?, ¿cómo lo sabemos?, ¿qué dicen las propias personas? Los indicadores más elocuentes proceden de las personas: «¿Cree que la ayuda responde a sus necesidades más importantes?», «¿Sabe cómo presentar una queja?», «¿Se siente tratado con respeto?».

> La rendición de cuentas se juzga por lo que cambia. Un programa que consulta sin modificar nunca nada enseña a las personas que su opinión no cuenta, y la siguiente participación será menor.

Diseñar un mecanismo de quejas y retroalimentación

Un **mecanismo de quejas y retroalimentación** (CFM o FCRM) recibe, registra, trata y responde a todas las comunicaciones de las personas: preguntas, sugerencias, quejas simples, quejas sensibles. Su diseño parte de la consulta: ¿qué canales utilizan las personas y en quién confían? Las mujeres, los niños, las personas mayores y las personas con discapacidad no dan las mismas respuestas.

Varios canales se complementan: línea telefónica gratuita, mensajería, buzones de sugerencias (para el anonimato, si los vacía una persona independiente del equipo de distribución), puntos de atención y mostradores de acogida durante las distribuciones, visitas a domicilio para las personas que no pueden desplazarse, comités comunitarios y retroalimentación recogida durante el seguimiento.

El mecanismo define **categorías** y **plazos** por escrito. Por ejemplo: solicitud de información, respuesta en 3 días; queja sobre el programa (no inscrito, monto erróneo), respuesta en 7 días; queja sensible (fraude, abuso, explotación sexual, seguridad), transmisión en 24 horas a la persona designada, por un circuito confidencial separado.

Los principios no negociables: **confidencialidad**, **seguridad** de las personas que se quejan (ninguna represalia), **accesibilidad** (idiomas, analfabetismo, discapacidad), **gratuidad** y **separación** entre quienes tratan las quejas y quienes son objeto de ellas.

Cerrar el ciclo y medir el desempeño

«Cerrar el ciclo» significa que cada persona recibe una respuesta, incluso cuando la respuesta es no, y que las tendencias de la retroalimentación alimentan las decisiones. La mayoría de los mecanismos recopilan, pero muchos no responden.

Indicadores de desempeño útiles:

- Número de comentarios recibidos al mes y por cada 1 000 personas asistidas, por canal, por sexo y edad.
- Proporción de comentarios cerrados dentro del plazo previsto.
- Proporción de personas satisfechas con la respuesta (verificada mediante una llamada a una muestra).
- Proporción de personas que conocen al menos un canal (encuesta de seguimiento).
- Número de cambios en el programa derivados de la retroalimentación durante el trimestre.

Ejemplo. 340 comentarios en un mes para 17 000 personas: 20 por cada 1 000. 289 cerrados dentro del plazo: 85 %. Comparado con un mes anterior de 2 comentarios por cada 1 000, no es un deterioro: a menudo es la señal de que el mecanismo por fin es conocido y de que las personas confían en él.

> Un mecanismo que recibe muy pocas quejas rara vez es señal de un programa perfecto: casi siempre es señal de un mecanismo desconocido, inaccesible o que no inspira confianza.

Quejas sensibles y protección contra la explotación y el abuso sexuales

Las quejas de **explotación y abuso sexuales** (EAS) por parte del personal humanitario siguen un circuito estricto, basado en los seis principios fundamentales del IASC: la EAS constituye una falta grave que justifica el despido; toda actividad sexual con un niño (menor de 18 años) está prohibida sea cual sea la edad de consentimiento local; el intercambio de

dinero, empleo, bienes o servicios por relaciones sexuales está prohibido; las relaciones sexuales con beneficiarios se desaconsejan firmemente, porque se basan en una dinámica de poder desigual; el personal debe señalar cualquier preocupación; y los responsables deben crear un entorno que prevenga la EAS.

En el mecanismo, esto significa: toda persona que recibe una denuncia la transmite **inmediatamente** al punto focal designado, sin investigar por su cuenta; la **persona sobreviviente** es derivada hacia una asistencia (atención médica, apoyo psicosocial, protección, asistencia jurídica) según sus deseos; la persona acusada se aparta de los contactos con los beneficiarios durante la investigación; la información se comparte estrictamente según la necesidad de conocerla; el donante es informado según el contrato.

Un enfoque **centrado en la persona sobreviviente** respeta cuatro principios: seguridad, confidencialidad, respeto de sus decisiones, no discriminación. Nunca se confronta a la persona con el presunto autor y no se le pide un testimonio repetido.

Participación e información

La **participación** va de la información a la decisión compartida. En la escala clásica se distinguen: informar, consultar, implicar, colaborar y transferir la decisión. La mayoría de los programas humanitarios se quedan en el nivel de la consulta; apuntar más alto en ciertas decisiones (elección de los sitios, criterios de focalización, composición de los kits, calendario) mejora la pertinencia y la aceptación.

Los comités comunitarios son útiles pero frágiles: a menudo reproducen las relaciones de poder existentes. Se cuida su composición (mujeres, jóvenes, personas con discapacidad, minorías), se fija un mandato escrito y no se les confía en solitario la selección de los beneficiarios sin control.

La **información** es una forma de ayuda: las personas tienen derecho a saber quién es la organización, qué hace, para quién, según qué criterios, durante cuánto tiempo, cómo quejarse y que la ayuda es gratuita. Se utilizan los idiomas y los formatos adecuados (radio, mensajes de voz, carteles ilustrados, reuniones separadas para las mujeres).

Las **decisiones difíciles** (reducción de las raciones, fin de un programa, cambio de focalización) deben anunciarse pronto, explicarse con honestidad y acompañarse: derivación hacia otros servicios, periodo de transición. Anunciar un corte el día de la distribución es una falta de rendición de cuentas que además crea riesgos de seguridad para el personal.

Protección de datos e integración en el ciclo del programa

Los comentarios y las quejas contienen **datos personales**, a veces muy sensibles. Reglas mínimas: acceso limitado a las personas que tratan los casos; separación de los datos de identificación y del contenido para los análisis; anonimización en los informes; plazo de conservación definido; ningún intercambio con autoridades o terceros sin una base clara y sin evaluación de los riesgos para las personas.

La rendición de cuentas debe integrarse en cada etapa del ciclo del programa, y presupuestarse:

- **Evaluación de necesidades**: preguntar a las personas por sus prioridades y sus canales preferidos.
- **Diseño**: criterios de focalización explicados, mecanismo de quejas previsto, línea presupuestaria (personal, línea telefónica, material informativo).
- **Ejecución**: información, comentarios tratados, ajustes.
- **Seguimiento**: indicadores de rendición de cuentas en el tablero de control.
- **Evaluación y aprendizaje**: las propias personas juzgan la pertinencia y la calidad de la ayuda.

A nivel colectivo, los mecanismos comunes (líneas de información interagenciales) reducen la confusión para las personas, que a menudo no conocen el nombre de las organizaciones. Exigen acuerdos claros de intercambio y de derivación de casos.

Ejercicio: corregir un mecanismo de quejas

Situación. Programa de 24 000 personas. En seis meses se recibieron 31 comentarios, todos a través del buzón de sugerencias instalado en la oficina de la organización y vaciado por el jefe de sitio; 12 recibieron respuesta. Una encuesta de seguimiento muestra que el 22 % de las personas conoce algún medio para quejarse (13 % de las mujeres, 31 % de los hombres). Un encuestador informa de un rumor según el cual un distribuidor pediría favores sexuales a mujeres a cambio de una inscripción.

Trabajo solicitado. 1) Calcule la tasa de comentarios por cada 1 000 personas y la tasa de respuesta. 2) Identifique los cuatro problemas principales. 3) ¿Qué hace en las 24 horas siguientes con respecto al rumor? 4) Proponga tres indicadores de seguimiento.

Corrección. 31 comentarios para 24 000 personas: 1,3 por cada 1 000 en seis meses, extremadamente bajo. Tasa de respuesta: 12 / 31 = 39 %. Problemas: canal único, poco accesible y situado en la oficina; buzón vaciado por una persona que puede ser objeto de las quejas (falta de independencia); escaso conocimiento, sobre todo entre las mujeres; ciclo no cerrado. En las 24 horas: transmitir el rumor al punto focal de PSEA sin investigar por cuenta propia, organizar la separación preventiva del distribuidor de las actividades de inscripción, prever la asistencia para cualquier sobreviviente identificada, informar a la dirección y al donante según los procedimientos. Indicadores: comentarios por cada 1 000 personas por sexo y canal; proporción cerrada dentro del plazo; proporción de personas, por sexo, que conocen al menos un canal.

Para recordar

- La CHS fija nueve compromisos; los mejores indicadores proceden de lo que dicen las personas.
- Un mecanismo de retroalimentación tiene varios canales elegidos con las personas, plazos escritos y un circuito sensible separado.
- Cerrar el ciclo: cada comentario recibe una respuesta y las tendencias cambian el programa.
- Pocas quejas suelen indicar un mecanismo desconocido o que no inspira confianza, no un programa perfecto.
- Denuncia de EAS: transmisión inmediata al punto focal, asistencia centrada en la persona sobreviviente, ninguna investigación improvisada.

Cuestionario

1. Un mecanismo recibe 3 quejas en seis meses para 20 000 personas asistidas. La interpretación más probable:
   A. El programa es excelente
   B. El mecanismo es desconocido, inaccesible o no inspira confianza
   C. Las personas no tienen nada que decir
   D. Hay que suprimir el mecanismo

2. Un agente recibe una denuncia de explotación sexual contra un colega. Debe:
   A. Interrogar él mismo al colega
   B. Transmitirla inmediatamente al punto focal designado, sin investigar por su cuenta
   C. Esperar a tener pruebas
   D. Hablar de ello en una reunión de equipo

3. Según los principios del IASC, toda actividad sexual con una persona menor de 18 años está prohibida:
   A. Solo si la edad de consentimiento local es de 18 años
   B. Sea cual sea la edad de consentimiento local
   C. Solo para el personal internacional
   D. Solo durante las horas de trabajo

4. ¿Quién no debe vaciar un buzón de quejas?
   A. Un responsable de rendición de cuentas independiente
   B. El jefe del equipo de distribución cuyas actividades son objeto de las quejas
   C. Un binomio que incluya a una mujer
   D. Un miembro del equipo de MEAL

5. Se decide una reducción de las raciones. La práctica conforme a la rendición de cuentas es:
   A. Anunciarla el día de la distribución
   B. Anunciarla pronto, explicarla con honestidad y derivar hacia otros servicios durante una transición
   C. No anunciarla
   D. Dejar que las autoridades la anuncien

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 9 : Evaluar: criterios, términos de referencia, matriz y utilización

Objetivos

- Elegir el tipo de evaluación y los criterios adaptados a la decisión que debe orientar.
- Redactar términos de referencia precisos y construir una matriz de evaluación rigurosa.
- Gestionar una evaluación de forma independiente y transformar sus recomendaciones en decisiones con seguimiento.

Por qué y cuándo evaluar

Una **evaluación** es una apreciación sistemática y objetiva de una intervención, de su diseño, de su ejecución y de sus resultados. Se diferencia del seguimiento: el seguimiento es continuo e interno, y pregunta «¿hacemos lo que habíamos previsto?»; la evaluación es puntual, a menudo externa, y pregunta «¿hacemos lo correcto, de la manera correcta, y qué cambia con ello?».

Tiene dos funciones, a veces en tensión: la **rendición de cuentas** (rendir cuentas a las personas, a los donantes, a la junta directiva) y el **aprendizaje** (mejorar este programa y los siguientes). Una evaluación concebida únicamente para la rendición de cuentas produce calificaciones; una evaluación concebida para el aprendizaje produce decisiones.

Tipo	Momento	Pregunta principal
Evaluación en tiempo real	Primeras semanas de una emergencia	¿Hay que corregir la respuesta ahora?
Evaluación de medio término	Mitad del proyecto	¿Qué hay que ajustar para la segunda mitad?
Evaluación final	Final del proyecto	¿Qué resultados, qué lecciones para lo que sigue?
Evaluación de impacto	Durante y después	¿Qué cambio es atribuible a la intervención?
Evaluación ex post	De 1 a 3 años después del final	¿Han perdurado los efectos?

Antes de encargar una evaluación, escriba quién la utilizará y para qué decisión. Si nadie puede responder, la evaluación será un informe más en una estantería.

Los criterios del CAD de la OCDE

Los criterios de evaluación del Comité de Ayuda al Desarrollo de la OCDE, revisados en 2019, son seis:

- **Pertinencia**: ¿responde la intervención a las necesidades y prioridades de las personas, y sigue siendo pertinente si cambia el contexto?
- **Coherencia** (añadida en 2019): ¿es compatible con las demás intervenciones en el mismo contexto, internas y externas a la organización?
- **Eficacia**: ¿alcanza sus objetivos, también para los distintos grupos?
- **Eficiencia**: ¿produce sus resultados de forma económica y a tiempo?
- **Impacto**: ¿qué efectos a más largo plazo, positivos o negativos, buscados o no?
- **Sostenibilidad**: ¿continuarán los beneficios tras el final de la financiación?

En el ámbito humanitario también se utilizan los criterios adaptados por ALNAP: **cobertura** (¿se ha llegado a todas las personas con necesidades, sin exclusión?), **conectividad** (vínculo entre la emergencia y el largo plazo) y **coordinación**.

No todos los criterios se aplican con el mismo peso. Una evaluación en tiempo real de una respuesta de tres meses se centra en la pertinencia, la cobertura y la eficacia; hablar de sostenibilidad en esa fase tiene poco sentido. Elegir tres o cuatro criterios y profundizar en ellos es mejor que sobrevolar seis.

Términos de referencia que consiguen una buena evaluación

Los **términos de referencia** (TdR) son el contrato intelectual de la evaluación. Unos TdR vagos producen evaluaciones vagas. Contienen:

- **Contexto** y descripción de la intervención (con la teoría del cambio y el marco lógico en anexo).
- **Objeto y usuarios**: por qué ahora, para quién, para qué decisiones.
- **Alcance**: periodo, zonas, componentes, lo que queda excluido.
- **Preguntas de evaluación**: de cinco a ocho preguntas principales, no treinta. Cada pregunta es evaluable con los medios y el tiempo disponibles.
- **Enfoque y métodos** esperados, dejando que el equipo proponga, y los datos existentes disponibles.
- **Ética**: consentimiento, protección de datos, protocolos para los temas sensibles, participación de las personas.
- **Productos y calendario**: informe inicial (inception), devolución, borrador de informe, informe final, resumen.
- **Perfil del equipo**, presupuesto, gobernanza (grupo de referencia) y reglas de independencia.

Orden de magnitud. Una evaluación final externa de calidad de un proyecto multisectorial cuesta a menudo entre 25 000 y 80 000 USD según el terreno y los métodos. Un presupuesto

demasiado bajo conduce a unos pocos días de terreno y a un informe sin evidencias; es mejor hacer menos evaluaciones, mejor hechas.

La matriz de evaluación y el informe inicial

La **matriz de evaluación** es la herramienta central del informe inicial: para cada pregunta, precisa las subpreguntas, los indicadores o criterios de juicio, las fuentes de datos, los métodos de recopilación y de análisis, y los límites.

Ejemplo. Pregunta: «¿En qué medida las transferencias monetarias han mejorado la seguridad alimentaria de los hogares más vulnerables?». Criterios de juicio: evolución de la puntuación de consumo y del rCSI entre la línea de base y el final, comparada con una zona de comparación; diferencias según el sexo del jefe de hogar y la presencia de una persona con discapacidad. Fuentes: encuestas de línea de base y final, monitoreo posdistribución, grupos focales. Métodos: diferencia en diferencias, análisis cualitativo por grupo. Límites: zona de comparación parcialmente asistida por otros actores a partir del mes 6.

El **informe inicial** es el momento de corregir el rumbo: el equipo confirma lo que es factible, señala las preguntas imposibles de tratar y propone el muestreo. El gestor lo valida con el grupo de referencia antes del trabajo de campo. Es la mejor inversión de tiempo de toda la evaluación.

La **calidad** de una evaluación se juzga por la claridad del vínculo entre **hallazgos** (respaldados por datos), **conclusiones** (juicios basados en los hallazgos) y **recomendaciones** (acciones que se derivan de las conclusiones). Una recomendación sin hallazgo correspondiente es una opinión del evaluador.

Independencia, ética y gestión de la evaluación

La **independencia** protege la credibilidad: el equipo de evaluación no participó en el diseño ni en la ejecución; el gestor de la evaluación no es el jefe del proyecto evaluado; el equipo tiene libre acceso a los datos y a las personas; el programa puede comentar los hechos del borrador de informe, pero no imponer que se cambien las conclusiones.

Las normas y estándares de evaluación del Grupo de Evaluación de las Naciones Unidas (UNEG) y los principios de ALNAP insisten en la **ética**: no causar daño a las personas entrevistadas, consentimiento informado, confidencialidad, atención a las personas en situación de angustia y a los riesgos de represalias, devolución a las comunidades.

El **gestor de la evaluación** tiene un papel activo: prepara los datos, organiza el acceso al terreno, reúne al grupo de referencia, controla la calidad en las etapas clave, protege la independencia del equipo y organiza la devolución. Los principales riesgos son conocidos: retrasos de arranque que reducen el trabajo de campo, datos de línea de base inexistentes, acceso limitado, presión para suavizar los hallazgos.

> Una evaluación que solo encuentra aspectos positivos no es creíble. Un programa que acepta hallazgos negativos bien fundamentados demuestra que es serio.

La respuesta de la dirección y la utilización

La **respuesta de la dirección** (management response) es el documento que transforma una evaluación en decisiones. Para cada recomendación: aceptada, parcialmente aceptada o rechazada, con la justificación; acciones previstas; responsable; plazo. Se redacta en el mes siguiente al informe final y se publica con él.

El **seguimiento** de las acciones se hace en las revisiones trimestrales de dirección: ¿en qué punto está cada acción? Las organizaciones maduras publican cada año una tasa de aplicación de las recomendaciones aceptadas.

Una recomendación puede rechazarse legítimamente: demasiado costosa, fuera del mandato, ya en curso. Pero la justificación debe constar por escrito. Lo que no es aceptable es la ausencia de respuesta.

Por último, la **difusión** adapta el formato al público: resumen de dos páginas para la dirección y los donantes, presentación para los equipos, versión sencilla y oral para las comunidades, y publicación del informe completo (por ejemplo, en la biblioteca de ALNAP) para el sector. La publicación de las evaluaciones, incluidas las críticas, es un compromiso del Grand Bargain en materia de transparencia.

Ejercicio: redactar las preguntas y la matriz de una evaluación

Situación. Un proyecto de 18 meses (agua, higiene y protección) en tres campamentos de desplazados llega a su mitad. La dirección de país debe decidir dentro de cuatro meses si hay que extender el proyecto a un cuarto campamento y cómo. El donante ha pedido una evaluación pero no ha fijado preguntas. Presupuesto: 30 000 USD. Existen datos de línea de base para el agua, no para la protección.

Trabajo solicitado. 1) ¿Qué tipo de evaluación y qué criterios principales? 2) Redacte tres preguntas de evaluación. 3) Para una de ellas, complete una fila de la matriz (criterios de juicio, fuentes, métodos, límites).

Corrección. Evaluación de medio término, formativa, cuyos usuarios son la dirección de país y el donante; criterios principales: pertinencia, eficacia, cobertura, eficiencia (la sostenibilidad es secundaria en esta fase). Preguntas posibles: 1) ¿En qué medida los servicios de agua e higiene alcanzan las normas previstas para todos los grupos, incluidas las personas con discapacidad? 2) ¿Responden las actividades de protección a los riesgos prioritarios identificados por las mujeres y los niños? 3) ¿Qué condiciones hay que reunir para extender el proyecto a un cuarto campamento con la misma calidad y un costo por persona comparable? Fila de la matriz para 2): criterios: correspondencia entre los riesgos prioritarios declarados y las actividades, uso de los servicios, percepción de seguridad; fuentes: grupos focales separados por sexo y edad, entrevistas con los actores de protección, datos de los servicios (anonimizados); métodos: análisis cualitativo codificado, cartografía participativa de los riesgos; límites: ausencia de línea de base en protección, de ahí un análisis de contribución en lugar de una medición de la evolución.

Para recordar

- El seguimiento pregunta «¿hacemos lo que habíamos previsto?»; la evaluación pregunta «¿es lo correcto, bien hecho, y qué cambia?».
- Seis criterios del CAD (2019): pertinencia, coherencia, eficacia, eficiencia, impacto, sostenibilidad; más cobertura, conectividad y coordinación en el ámbito humanitario.
- Unos TdR precisos: usuarios, decisiones, de cinco a ocho preguntas evaluables, ética, productos.
- La matriz de evaluación vincula cada pregunta con criterios, fuentes, métodos y límites.
- Respuesta de la dirección: cada recomendación aceptada o no, con acción, responsable y plazo.

Cuestionario

1. ¿Qué criterio añadió el CAD de la OCDE en 2019?
   A. Pertinencia
   B. Coherencia
   C. Eficiencia
   D. Impacto

2. Una evaluación que pregunta si se ha llegado a todas las personas con necesidades, sin exclusión, trata del criterio:
   A. Sostenibilidad
   B. Cobertura
   C. Eficiencia
   D. Conectividad

3. El momento más útil para corregir una evaluación mal orientada es:
   A. La lectura del informe final
   B. La validación del informe inicial y de la matriz de evaluación
   C. La publicación
   D. La respuesta de la dirección

4. Una recomendación que no se apoya en ningún hallazgo del informe es:
   A. Una buena práctica
   B. Una opinión del evaluador, que hay que cuestionar
   C. Obligatoria
   D. Una conclusión

5. ¿Quién no debería gestionar la evaluación de un proyecto?
   A. Un responsable de MEAL de país
   B. El jefe del proyecto evaluado
   C. Un asesor de la sede
   D. Un comité de pilotaje

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 10 : Medir el impacto: contrafactual, métodos y contribución

Objetivos

- Comprender el problema del contrafactual y por qué una comparación antes-después no mide un efecto.
- Elegir y aplicar un método cuasiexperimental o experimental adaptado al contexto humanitario, con sus condiciones de validez.
- Utilizar los enfoques de contribución cuando la atribución es imposible, y decidir qué es razonable demostrar.

El problema fundamental: ¿qué habría ocurrido sin el programa?

El **impacto** de un programa es la diferencia entre lo que ocurrió con el programa y lo que habría ocurrido sin él: el **contrafactual**. Nunca se puede observar directamente el mismo hogar con y sin programa; todos los métodos de evaluación de impacto son maneras de construir una estimación creíble de ese contrafactual.

Dos comparaciones ingenuas son engañosas:

- **Antes-después** en los beneficiarios: el cambio observado mezcla el efecto del programa y todo lo que cambió al mismo tiempo (cosecha, precios, seguridad, otras ayudas). Si la cosecha fue buena, la seguridad alimentaria mejora incluso sin programa; si fue mala, un programa eficaz puede parecer sin efecto.
- **Beneficiarios frente a no beneficiarios** después del programa: los dos grupos eran diferentes al principio, precisamente porque la focalización eligió a los más vulnerables. La comparación subestima entonces el efecto, o lo sobrestima si fueron los más motivados los que se inscribieron (sesgo de selección).

> Pregunta que hay que plantear ante cualquier afirmación de impacto: «¿en comparación con qué?». Si la respuesta es «en comparación con antes», no es una medición de impacto.

El ensayo controlado aleatorizado y su uso en el ámbito humanitario

En un **ensayo controlado aleatorizado** (ECA), se sortea, entre las personas elegibles, quiénes reciben el programa (grupo de tratamiento) y quiénes no lo reciben o lo reciben más tarde (grupo de control). El sorteo hace que los dos grupos sean comparables en promedio en todas las características, observadas o no. El efecto es simplemente la diferencia de los resultados medios al final.

En el ámbito humanitario, negar una ayuda vital a personas elegibles para constituir un grupo de control es **contrario a la ética**. Pero los ECA siguen siendo posibles y han producido conocimientos importantes (sobre el efectivo, la nutrición, la salud mental) en situaciones concretas:

- **Recursos insuficientes** de todos modos para todo el mundo: el sorteo es entonces una forma equitativa de repartir, y el grupo no atendido sirve de control.
- **Implantación progresiva**: todo el mundo recibirá el programa, pero el orden se sortea (diseño escalonado, «stepped wedge»).
- **Comparación de dos variantes** de un programa, ambas consideradas aceptables: efectivo mensual o en un solo pago, con o sin acompañamiento.

Los ECA son caros, llevan tiempo y responden a preguntas concretas; se justifican cuando la respuesta servirá a numerosos programas. Para la mayoría de los proyectos, bastan los métodos cuasiexperimentales.

La diferencia en diferencias

La **diferencia en diferencias** compara la evolución de un grupo asistido con la evolución de un grupo de comparación no asistido: efecto = (después − antes) en el grupo asistido − (después − antes) en el grupo de comparación. Elimina las diferencias iniciales estables entre los grupos y las tendencias comunes (cosecha, precios).

Ejemplo. Proporción de hogares con un consumo aceptable. Grupo asistido: 28 % antes, 52 % después (+24 puntos). Grupo de comparación: 35 % antes, 44 % después (+9 puntos). Efecto estimado: 24 − 9 = 15 puntos. Una comparación antes-después habría atribuido 24 puntos al programa; una comparación de los resultados finales por sí solos le habría atribuido 8.

Condición de validez: el supuesto de **tendencias paralelas**. Sin el programa, los dos grupos habrían evolucionado de la misma manera. Se verifica, si es posible, con varias mediciones antes del programa, y se comprueba que el grupo de comparación no recibe otra ayuda importante durante el periodo (contaminación) ni sufre un choque específico.

En el ámbito humanitario, el grupo de comparación suele ser una zona vecina no cubierta por falta de financiación, o hogares elegibles inscritos en lista de espera. Nunca se crea un grupo de comparación negando la ayuda para las necesidades del estudio.

Emparejamiento y regresión discontinua

El **emparejamiento** construye un grupo de comparación asociando a cada hogar asistido uno o varios hogares no asistidos que se le parecen en las características medidas antes del programa (tamaño, bienes, ingresos, estatus, zona). El método del **puntaje de propensión** resume estas características en una probabilidad de ser asistido, y se comparan hogares con puntajes cercanos. Su límite: solo corrige las diferencias observadas. Si la motivación o las redes sociales influyeron en la inscripción y no se miden, el sesgo persiste.

La **regresión discontinua** se aplica cuando la elegibilidad depende de un umbral en una puntuación: por ejemplo, se asiste a los hogares con una puntuación de vulnerabilidad de 60 o más, y no a los que están por debajo. Los hogares justo por encima y justo por debajo del

umbral (57-59 y 60-62) son casi idénticos; la diferencia de resultados entre ellos estima el efecto del programa en torno al umbral. Condiciones: el umbral se aplica estrictamente y los hogares no pueden manipular su puntuación para superarlo.

Ejemplo. Hogares con puntuación 57-59: 38 % con un consumo aceptable. Hogares con puntuación 60-62 (asistidos): 51 %. Efecto estimado en torno al umbral: unos 13 puntos. Este resultado vale para los hogares cercanos al umbral, no necesariamente para los más vulnerables.

La **serie temporal interrumpida** compara la tendencia de un indicador antes y después del inicio de un programa, sobre numerosos puntos de medición (por ejemplo, las admisiones semanales en nutrición). Es útil cuando no existe ningún grupo de comparación, siempre que se disponga de suficientes puntos antes de la intervención.

Enfoques de contribución: cuando la atribución es imposible

Muchos efectos importantes no se prestan a los métodos estadísticos: una labor de incidencia que cambia una ley, un fortalecimiento de capacidades de organizaciones locales, una coordinación mejorada. No se pueden comparar con un contrafactual cuantificado. Se busca entonces establecer de forma creíble que el programa **contribuyó** al cambio.

- **Análisis de contribución** (Mayne): se expone la teoría del cambio, se reúnen las evidencias para cada vínculo, se examinan las explicaciones alternativas y se construye un «relato de contribución» que se somete a la crítica de las partes interesadas.
- **Rastreo de procesos** (process tracing): se ponen a prueba hipótesis causales buscando evidencias cuya presencia o ausencia permite confirmar o descartar una explicación (por ejemplo, la reproducción palabra por palabra de una nota de incidencia en un decreto).
- **Cosecha de alcances** (outcome harvesting): primero se identifican los cambios observados en los actores y luego se rastrea hacia atrás para determinar si el programa contribuyó a ellos y cómo.
- **Cambio más significativo**: se recopilan relatos de cambio entre las personas, y unos paneles seleccionan los más significativos explicando por qué. Útil para descubrir efectos no previstos.

> Rigor no quiere decir cuantitativo. Un análisis de contribución bien realizado, que examina seriamente las explicaciones alternativas, es más creíble que una comparación antes-después cuantificada.

Elegir el método y decidir qué hay que demostrar

La elección de un método depende de cuatro preguntas:

- **¿Qué decisión?** Extender un modelo piloto a escala nacional justifica una evidencia sólida; ajustar un programa en curso se conforma con una evidencia suficiente.
- **¿Existen ya las evidencias?** El efecto del efectivo sobre el consumo de alimentos está muy documentado: es inútil volver a demostrarlo con un ECA en cada proyecto. A menudo

basta con verificar que el programa es fiel al modelo y que los efectos intermedios están presentes.

- **¿Qué es factible y ético?** ¿Existe un grupo de comparación sin negar la ayuda? ¿Un umbral de elegibilidad nítido? ¿Una implantación progresiva?
- **¿Qué presupuesto y qué plazo?** Una evaluación de impacto seria exige una línea de base antes del inicio: se diseña con el programa, no al final.

En cuanto al **tamaño de la muestra**, una evaluación de impacto debe tener una potencia suficiente para detectar el efecto esperado. Detectar un efecto de 15 puntos requiere muchos menos hogares que un efecto de 5 puntos. Si el efecto esperado es pequeño y la muestra reducida, el estudio concluirá casi con seguridad «ningún efecto significativo», lo que se interpretará erróneamente como «ningún efecto».

Último punto: presentar con honestidad el **nivel de evidencia**. «Los hogares asistidos mejoraron su consumo en 24 puntos; en comparación con una zona de comparación, el efecto atribuible se estima en 15 puntos, bajo el supuesto de tendencias paralelas que respaldan los datos de dos campañas anteriores» es una frase que inspira confianza.

Ejercicio: estimar un efecto y criticar el método

Situación. Un programa de medios de vida apoyó a 1 200 hogares. Ingreso mensual mediano: hogares asistidos 42 USD antes, 71 USD después. Zona vecina no cubierta (por falta de financiación): 47 USD antes, 58 USD después. Durante el programa, otra ONG distribuyó semillas en la mitad de las aldeas de la zona vecina. El informe final anuncia: «el programa aumentó los ingresos un 69 %».

Trabajo solicitado. 1) ¿De dónde viene la cifra del 69 % y por qué es engañosa? 2) Calcule el efecto por diferencia en diferencias. 3) ¿Cuál es el efecto probable de la distribución de semillas en su estimación, y qué propone?

Corrección. 69 %: (71 − 42) / 42, comparación antes-después que atribuye al programa todo el aumento, incluido el debido al contexto. Diferencia en diferencias: (71 − 42) − (58 − 47) = 29 − 11 = 18 USD al mes, es decir, aproximadamente el 43 % del ingreso inicial de los hogares asistidos. Las semillas probablemente aumentaron los ingresos de la zona de comparación, lo que infla su tendencia y subestima el efecto del programa (la contaminación viola el supuesto de tendencias paralelas). Propuesta: analizar por separado las aldeas de comparación sin semillas, o controlar esta variable; presentar la estimación como un límite inferior e indicar claramente la limitación.

Para recordar

- El impacto se mide en comparación con un contrafactual; antes-después no es una medición de impacto.
- Diferencia en diferencias = evolución del grupo asistido − evolución del grupo de comparación, bajo el supuesto de tendencias paralelas.
- Un ECA en el ámbito humanitario solo es ético si no priva a nadie de una ayuda disponible: recursos limitados, implantación progresiva o comparación de variantes.
- Regresión discontinua: comparar los hogares justo por encima y justo por debajo de un umbral estricto de elegibilidad.
- Cuando la atribución es imposible: análisis de contribución, rastreo de procesos, cosecha de alcances, cambio más significativo.

Cuestionario

1. Grupo asistido: 30 % antes, 50 % después. Comparación: 32 % antes, 40 % después. Efecto por diferencia en diferencias:
   A. 20 puntos
   B. 12 puntos
   C. 10 puntos
   D. 8 puntos

2. ¿Qué supuesto es esencial para la diferencia en diferencias?
   A. Los grupos son idénticos al principio
   B. Sin el programa, los dos grupos habrían evolucionado de la misma manera
   C. La muestra supera los 1 000 hogares
   D. El programa duró más de un año

3. ¿En qué situación es éticamente aceptable un sorteo en el ámbito humanitario?
   A. Para negar una ayuda vital disponible a la mitad de los elegibles
   B. Cuando los recursos no alcanzan de todos modos para todos los elegibles, o para el orden de una implantación progresiva
   C. Nunca
   D. Solo si el donante lo impone

4. El principal límite del emparejamiento por puntaje de propensión es:
   A. Solo funciona con ECA
   B. Solo corrige las diferencias medidas, no las diferencias no observadas como la motivación
   C. Exige un umbral de elegibilidad
   D. Está prohibido por la OCDE

5. Para demostrar que una labor de incidencia contribuyó a la adopción de una ley, el método más adecuado es:
   A. Un ensayo aleatorizado
   B. Una diferencia en diferencias
   C. Un análisis de contribución o un rastreo de procesos
   D. Una encuesta de hogares

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. C

Semana 11 : Optimización de recursos: costo-eficiencia, costo-eficacia y arbitrajes

Objetivos

- Aplicar el marco de las 4E (economía, eficiencia, eficacia, equidad) para juzgar la optimización de los recursos de un programa.
- Realizar un análisis riguroso de costo-eficiencia y de costo-eficacia, con asignación de los costos compartidos y análisis de sensibilidad.
- Utilizar estos análisis para arbitrar entre opciones, sin sacrificar la equidad ni la calidad.

La optimización de recursos: el marco de las 4E

La **optimización de recursos** (value for money, VfM) no significa «lo más barato». Es el mejor uso de los recursos para obtener los resultados buscados. El marco más utilizado, procedente de la cooperación británica, tiene cuatro dimensiones:

- **Economía**: comprar los insumos con la calidad requerida al mejor precio (precio de los kits, salarios, alquileres comparados con el mercado).
- **Eficiencia**: transformar los insumos en productos al menor costo (costo por kit distribuido, por hogar registrado, ratio de transferencia).
- **Eficacia**: transformar los productos en efectos (proporción de hogares que mejoran su consumo, tasa de curación).
- **Equidad**: llegar a los más vulnerables y a los más difíciles de alcanzar, lo que a menudo cuesta más por persona.

A veces se añade la **costo-eficacia** como síntesis: costo por efecto obtenido. Las cuatro dimensiones van juntas: un programa muy económico que compra kits de mala calidad, o muy eficiente que evita las zonas remotas, no optimiza los recursos.

> Llegar a las personas más aisladas o a las personas con discapacidad cuesta más por persona. Un informe de VfM honesto lo explica en lugar de ocultarlo, y muestra que ese sobrecosto es una elección de equidad.

Análisis de costo-eficiencia: el método

Un **análisis de costo-eficiencia** calcula el costo de producción de un producto (por ejemplo, el costo por hogar que recibe una transferencia durante seis meses). Exige atribuir a la actividad **todos** los costos que consume, no solo los costos directos: personal del programa, parte del personal de apoyo (logística, finanzas, RR. HH.), oficinas, vehículos y parte de los costos de la sede.

Las etapas, inspiradas en el enfoque de análisis sistemático de costos desarrollado por ONG y retomado por la herramienta Dioptra:

- **Definir** el producto y el periodo.
- **Reunir** los gastos reales (no el presupuesto) de todas las líneas concernidas.
- **Asignar** cada línea a la actividad según una clave: 100 % para un costo vinculado exclusivamente; proporción del tiempo para el personal compartido; proporción de los gastos directos para los costos de apoyo.
- **Dividir** el costo total asignado por el número de productos.
- **Comparar** con otros programas u otros periodos, teniendo en cuenta el contexto.

Ejemplo. Programa de efectivo de 6 meses para 2 000 hogares: transferencias 1 080 000 USD; personal del programa 96 000; comisiones del proveedor de pagos (2 %) 21 600; parte del apoyo de país 54 000; parte de la sede 42 400. Costo total: 1 294 000 USD. Costo por hogar: 647 USD. Ratio de transferencia: 1 080 000 / 1 294 000 = 0,83.

Análisis de costo-eficacia

El **análisis de costo-eficacia** relaciona el costo con un efecto: costo por niño curado, por hogar que sale de la inseguridad alimentaria, por punto de puntuación ganado, por caso de cólera evitado. Permite comparar enfoques que persiguen el mismo efecto.

Ejemplo de comparación. Dos opciones para mejorar el consumo de alimentos de 1 000 hogares durante seis meses. Opción A, alimentos: costo total 520 000 USD, 380 hogares pasan a un consumo aceptable (en comparación con el contrafactual). Opción B, efectivo: costo total 450 000 USD, 410 hogares. Costo por hogar que ha progresado: A = 1 368 USD; B = 1 098 USD. B es más costo-eficaz en este contexto.

Tres reglas de interpretación:

- El efecto debe ser **atribuible** (semana 10): dividir el costo por una mejora antes-después sobrestima o subestima groseramente la costo-eficacia.
- Las comparaciones solo valen entre opciones que persiguen el **mismo efecto**, en contextos comparables.
- Los efectos **múltiples** complican la comparación: el efectivo también puede reducir la deuda y apoyar los mercados. En salud se utilizan medidas sintéticas (años de vida ajustados por discapacidad, AVAD), rara vez disponibles en los demás sectores.

El **retorno social de la inversión** (SROI), que convierte todos los efectos en dinero, es atractivo pero frágil: los valores monetarios atribuidos a los efectos sociales son muy discutibles. Hay que leer estos ratios con prudencia y examinar siempre los supuestos.

Economías de escala, costos fijos y análisis de sensibilidad

El costo unitario depende mucho de la **escala**. Los costos fijos (equipo de coordinación, sistema de pago, oficina) se reparten entre más beneficiarios cuando el programa crece.

Ejemplo. Costos fijos: 150 000 USD; costo variable por hogar: 400 USD. Para 500 hogares: (150 000 + 200 000) / 500 = 700 USD por hogar. Para 2 000 hogares: (150 000 + 800 000) / 2 000 = 475 USD. Comparar el costo unitario de un proyecto piloto con el de un programa a gran escala es, por tanto, injusto; se compara a una escala comparable, o se modeliza el costo a la escala prevista.

El **análisis de sensibilidad** pone a prueba la solidez de las conclusiones cuando cambian los supuestos: ¿qué pasa con la comparación si el efecto de la opción B es un 20 % más bajo de lo previsto?, ¿y si las comisiones de pago se duplican? Si la conclusión cambia ante una variación realista, hay que decirlo: la decisión se basa entonces en un supuesto que hay que verificar.

El **punto de equilibrio** de una inversión es otra herramienta útil: un sistema de pago móvil que cuesta 60 000 USD de instalación y ahorra 12 USD por hogar y por mes se amortiza a partir de 5 000 hogares-mes (60 000 / 12), por ejemplo 1 000 hogares durante 5 meses.

Presentar la optimización de recursos a los donantes

Los donantes piden cada vez más una sección de VfM en las propuestas y los informes. Una sección sólida:

- Presenta **indicadores para cada una de las 4E**, con su valor y una comparación (periodo anterior, otra zona, referencia externa).
- Muestra las **medidas concretas** adoptadas: compras agrupadas, acuerdos marco, oficinas y vehículos compartidos con otros proyectos, pago electrónico, reducción de pérdidas.
- Explica los **sobrecostos justificados**: zonas difíciles, accesibilidad, seguridad, calidad, localización (apoyo a los socios).
- Reconoce los **límites**: efectos no atribuibles, datos de costos incompletos.

Las trampas que hay que evitar: presentar ahorros a corto plazo que degradan la calidad (menos seguimiento, materiales menos duraderos); comparar costos unitarios de contextos muy diferentes; mostrar un ratio de transferencia elevado obtenido suprimiendo el seguimiento y la rendición de cuentas.

> La optimización de recursos es también un argumento a favor de los costos de apoyo: un programa sin logística, finanzas y MEAL sólidos no es eficiente, es arriesgado.

Arbitrar: eficiencia, equidad y calidad

Los análisis de costos iluminan las decisiones, no las toman. Los arbitrajes habituales:

- **Cobertura frente a profundidad**: ¿ayudar a 10 000 hogares con una transferencia que cubre el 50 % de la canasta, o a 6 000 con una transferencia que cubre el 85 %? Si una transferencia parcial no permite alcanzar el efecto buscado (por ejemplo, evitar las estrategias de afrontamiento negativas), ayudar a más gente peor no es más eficaz.
- **Zonas fáciles frente a zonas difíciles**: el costo por persona puede duplicarse en las zonas remotas, donde las necesidades suelen ser las más severas.

- **Corto plazo frente a sostenibilidad**: reparar un punto de agua cuesta menos que formar un comité de gestión e implantar un sistema de recuperación de costos, pero el punto de agua corre el riesgo de volver a averiarse.

Un método para decidir es combinar el análisis de costo-eficacia con **criterios explícitos**: umbral mínimo de calidad (normas Esfera), prioridad a los más severos, techo de costo por persona por encima del cual se busca otro enfoque. La decisión y sus razones se ponen por escrito, para poder explicarlas a las personas, a los equipos y a los donantes.

Ejercicio: comparar dos opciones de suministro de agua

Situación. Para suministrar 15 litros de agua por persona y por día a 12 000 personas durante 12 meses, hay dos opciones. Opción A, camiones cisterna: costos fijos 30 000 USD, costo variable 0,012 USD por litro. Opción B, rehabilitación de dos pozos y red: inversión 210 000 USD, funcionamiento 2 500 USD al mes. La vida útil de los pozos rehabilitados es de al menos cinco años.

Trabajo solicitado. 1) Calcule el costo de cada opción en 12 meses. 2) Calcule el costo por persona y por año. 3) ¿A partir de cuántos meses resulta más barata la opción B? 4) ¿Qué recomendación, y con qué condiciones?

Corrección. Volumen: 12 000 × 15 × 365 = 65 700 000 litros al año. Opción A: 30 000 + 0,012 × 65 700 000 = 30 000 + 788 400 = 818 400 USD, es decir, 68,2 USD por persona y por año. Opción B: 210 000 + 12 × 2 500 = 240 000 USD, es decir, 20 USD por persona el primer año. Umbral: costo mensual del transporte en camión ≈ 788 400 / 12 = 65 700 USD; B es más barata en cuanto 30 000 + 65 700 × m > 210 000 + 2 500 × m, es decir, m > 180 000 / 63 200 = 2,85 meses. Recomendación: B, con suministro en camión durante las obras (2 a 3 meses), siempre que la hidrogeología confirme el caudal y la calidad del agua, y que se prevean un comité de gestión y una financiación del funcionamiento más allá del proyecto.

Para recordar

- 4E: economía, eficiencia, eficacia, equidad; lo más barato no es lo óptimo.
- Costo-eficiencia = costo total asignado (incluidos apoyo y sede) / productos.
- Costo-eficacia = costo / efecto atribuible; comparar solo opciones que persiguen el mismo efecto.
- Los costos fijos hacen que los programas pequeños sean más caros por persona: comparar a una escala comparable.
- Análisis de sensibilidad: si la conclusión cambia ante una variación realista, la decisión se basa en un supuesto que hay que verificar.

Cuestionario

1. Llegar a las personas con discapacidad y a las zonas remotas aumenta el costo por persona. En el marco de las 4E, este sobrecosto corresponde a:
   A. La economía
   B. La equidad
   C. La eficiencia
   D. Un error de gestión
2. Costos fijos 100 000 USD, costo variable 300 USD por hogar, 1 000 hogares. Costo por hogar:
   A. 300 USD
   B. 400 USD
   C. 1 300 USD
   D. 100 USD
3. Opción A: 600 000 USD para 400 niños curados. Opción B: 500 000 USD para 250 niños curados. ¿Cuál es más costo-eficaz?
   A. B, porque cuesta menos en total
   B. A: 1 500 USD por niño curado frente a 2 000 para B
   C. Son equivalentes
   D. No se pueden comparar
4. ¿Por qué hay que incluir una parte de los costos de apoyo y de la sede en un análisis de costo-eficiencia?
   A. Para inflar las cifras
   B. Porque la actividad consume realmente esos recursos; ignorarlos subestima su costo
   C. Porque el donante lo paga todo
   D. No es necesario
5. Una inversión de 40 000 USD ahorra 8 USD por hogar y por mes. Se amortiza a partir de:
   A. 500 hogares-mes
   B. 5 000 hogares-mes
   C. 320 000 hogares-mes
   D. 50 000 hogares-mes

Respuestas del cuestionario : 1. B · 2. B · 3. B · 4. B · 5. B

Semana 12 : Gestión adaptativa, aprendizaje e informes: dirigir con base en la evidencia

Objetivos

- Implantar una gestión adaptativa: revisiones periódicas, reglas de decisión y revisión documentada del programa.
- Construir un sistema de aprendizaje organizacional que capitalice y reutilice las lecciones, más allá de los informes.
- Redactar informes a los donantes honestos y convincentes, y presentar un expediente MEAL completo como proyecto final.

La gestión adaptativa

Los contextos humanitarios cambian más rápido que los marcos lógicos. La **gestión adaptativa** consiste en ajustar de forma deliberada y documentada las actividades, los métodos o las metas en función de los datos y del contexto, manteniendo el objetivo. No es improvisación: se basa en supuestos explícitos, un seguimiento que los pone a prueba y momentos de decisión previstos.

Sus elementos:

- **Hipótesis de trabajo** claramente enunciadas en la teoría del cambio («las mujeres podrán desplazarse a los puntos de retiro»).
- **Indicadores de contexto y de alerta** que las ponen a prueba (tiempo de trayecto declarado, incidentes de seguridad en las carreteras).
- **Reglas de decisión** escritas de antemano: «si más del 20 % de las transferencias no se retiran en 10 días, se abren puntos de retiro móviles».
- **Revisiones periódicas** (trimestrales, o mensuales en crisis aguda) con autoridad para decidir.
- **Un acuerdo con el donante** sobre lo que puede cambiar sin solicitud formal (actividades, métodos) y lo que exige una enmienda (objetivos, presupuesto global, zonas).

Los donantes suelen estar más abiertos de lo que se cree a una adaptación bien argumentada. Lo que rechazan son los cambios no anunciados que se descubren en el informe final.

> Un proyecto que no ha cambiado nada en 18 meses en un contexto volátil probablemente no ha mirado sus datos, o no se ha atrevido a hablar de ellos.

Las revisiones del programa y la revisión posterior a la acción

La **revisión trimestral** de un programa reúne al equipo de programa, MEAL, logística, finanzas y, si es posible, a representantes de los socios y de las comunidades. El orden del día tipo:

- Avance de los indicadores clave respecto a las metas y a las tendencias.
- Ejecución presupuestaria y ritmo de gasto.
- Retroalimentación de las personas y de los socios: temas principales, respuestas dadas.
- Contexto y supuestos: lo que ha cambiado.
- Decisiones: qué, quién, cuándo. Y seguimiento de las decisiones de la revisión anterior.

La **revisión posterior a la acción** (after action review), procedente de los ejércitos y luego ampliamente adoptada, es una herramienta corta (una a dos horas) que se utiliza después de un acontecimiento importante: una distribución, una evacuación, una primera oleada de respuesta. Cuatro preguntas: ¿qué habíamos previsto?, ¿qué ocurrió realmente?, ¿por qué hay una diferencia?, ¿qué haremos de otra manera la próxima vez? Funciona si se habla de hechos, sin buscar culpables, y si la última pregunta produce acciones con un responsable.

Las lecciones deben ser **específicas** y **accionables**: «hay que mejorar la comunicación» no es una lección; «anunciar las fechas de distribución por radio y por mensaje de voz al menos 5 días antes, porque el 40 % de los hogares no había recibido la información» sí lo es.

El aprendizaje organizacional

Las organizaciones humanitarias producen muchos informes y aprenden poco: los mismos errores se repiten de una crisis a otra, a menudo porque el personal rota rápido y las lecciones se quedan en documentos que nadie lee. Un sistema de aprendizaje eficaz:

- **Define preguntas de aprendizaje** prioritarias para la organización (por ejemplo: «¿cómo mantener la calidad del seguimiento cuando el acceso es intermitente?») y orienta el MEAL de varios proyectos para responderlas.
- **Capitaliza** en formatos utilizables: fichas de dos páginas, listas de verificación, modelos de documentos, en lugar de informes largos.
- **Hace circular** las lecciones en el momento adecuado: al iniciar un nuevo proyecto, se leen las lecciones de los proyectos similares (revisión de lecciones en la fase de diseño).
- **Valora** que se informe de los fracasos, por ejemplo mediante sesiones periódicas en las que los equipos presentan lo que no funcionó y lo que aprendieron de ello.
- **Comparte** con el sector: evaluaciones publicadas, contribuciones a los grupos de trabajo de los clústeres, estudios de caso.

El papel del responsable de MEAL cambia entonces: de productor de datos para los donantes, pasa a ser un **facilitador de la reflexión** que ayuda a los equipos a plantear buenas preguntas, a leer los datos y a decidir.

Redactar informes a los donantes

Un **informe narrativo** de calidad sigue la estructura del donante y responde a tres exigencias: **exactitud** (cifras coherentes con el informe financiero y con los informes anteriores), **análisis** (no solo lo que se hizo, sino lo que cambió y por qué) y **honestidad** (desviaciones y dificultades explicadas).

Para cada resultado, funciona una estructura sencilla: **meta, logrado, desviación, explicación, consecuencia, medida**. Por ejemplo: «Meta: 3 000 hogares. Logrado: 2 450 (82 %). Desviación debida al cierre de la carretera de X durante 7 semanas. Consecuencia: retraso de dos ciclos de transferencias en 4 aldeas. Medida: pago móvil implantado en mayo; recuperación prevista de aquí a agosto, sin costo adicional».

Errores que hay que evitar:

- Cifras de productos presentadas como efectos («12 000 personas sensibilizadas mejoraron sus prácticas»).
- Dobles conteos entre actividades (la misma persona contada en cada actividad); hay que distinguir beneficiarios únicos y participaciones.
- Éxitos individuales presentados como representativos; una historia ilustra, no prueba.
- La ausencia de mención de las quejas recibidas y de las respuestas dadas.

> El donante se entera tarde o temprano de los problemas. Es mejor que se entere por usted, con el análisis y las medidas adoptadas, que por una evaluación o una auditoría.

El cierre del programa y la sostenibilidad

El **cierre** se prepara desde el diseño (estrategia de salida) y se pone en marcha al menos tres meses antes del final. Comprende:

- **Información** a las comunidades y a las autoridades sobre la fecha y las modalidades de finalización, y derivación hacia otros servicios.
- **Transferencia** de los activos (puntos de agua, equipos) según las normas del donante, con acuerdos escritos con los receptores de esos activos.
- **Traspaso** a los socios, a los servicios técnicos o a los comités, con un fortalecimiento de sus capacidades realmente planificado y presupuestado.
- **Encuesta final** y evaluación, y luego devolución a las comunidades.
- **Archivo** de los datos y documentos, y destrucción segura de los datos personales según la política de conservación.

La **sostenibilidad** se juzga por las condiciones reunidas: una institución o una comunidad que asume la responsabilidad, recursos para el funcionamiento (presupuesto público, cuotas, recuperación de costos), competencias y una demanda real de los usuarios. Sin estas condiciones, es mejor reconocerlo y buscar un relevo (otra financiación, otro actor) en lugar de anunciar una sostenibilidad que no existirá.

Proyecto final: el expediente MEAL completo

El proyecto final de este Mini MBA consiste en presentar un **expediente MEAL completo** para un programa real o realista de su elección. Reúne todo lo que ha aprendido en doce semanas:

- Una **nota de análisis de necesidades** de dos páginas: fuentes, severidad, grupos prioritarios, brechas de cobertura, opciones de respuesta comparadas.

- Una **teoría del cambio** y un **marco lógico** con indicadores, líneas de base, metas calculadas y supuestos.
- Una **estrategia de focalización** y, en su caso, el cálculo del valor de las transferencias.
- Un **plan MEAL** con dos fichas de referencia de indicadores, el diseño muestral y el cálculo del tamaño de la muestra de la encuesta de línea de base.
- Un **mecanismo de quejas y retroalimentación**: canales, plazos, circuito sensible, indicadores.
- Un **plan de evaluación**: tipo, preguntas, método de estimación del efecto y sus condiciones de validez.
- Un **análisis de VfM**: indicadores de las 4E y una comparación de opciones.
- Un **calendario de revisiones** y tres reglas de decisión adaptativas.

Criterios de calidad: coherencia entre las partes, cálculos correctos y verificables, realismo (presupuesto MEAL del 3 al 10 %), lugar real otorgado a las personas y honestidad sobre los límites. Los estudios de caso del examen del diploma retoman exactamente estas competencias.

Ejercicio: redactar una regla de decisión y un párrafo de informe

Situación. Programa de efectivo para 4 000 hogares, pago móvil. En el tercer mes: el 18 % de las transferencias no se retiran después de 10 días (6 % el primer mes); en 3 de 20 aldeas, esta tasa supera el 50 %. La retroalimentación señala agentes sin liquidez y una carretera cortada por grupos armados. Meta del trimestre: 12 000 transferencias retiradas; logrado: 9 840.

Trabajo solicitado. 1) Redacte una regla de decisión adaptativa para este problema. 2) Redacte el párrafo del informe trimestral sobre este resultado, según la estructura meta, logrado, desviación, explicación, consecuencia, medida.

Corrección. Regla: «Si más del 15 % de las transferencias de una aldea no se retiran 10 días después del envío, el equipo verifica en un plazo de 72 horas la liquidez de los agentes y el acceso; si se confirma la causa, se organiza un punto de retiro móvil o un agente adicional antes del ciclo siguiente, y se informa al donante en el informe mensual». Párrafo: «Meta: 12 000 transferencias retiradas. Logrado: 9 840 (82 %). La desviación procede de 3 aldeas donde más de la mitad de las transferencias no se retiraron, debido a la falta de liquidez de los agentes y al corte de la carretera de Y. Consecuencia: unas 2 160 transferencias pendientes y un riesgo de recurso al endeudamiento para los hogares afectados. Medidas: dos puntos de retiro móviles desde la semana 12, acuerdo con el operador para abastecer a los agentes, información a los hogares por mensaje de voz; las transferencias siguen disponibles y se seguirán hasta su retiro completo».

Para recordar

- Gestión adaptativa: supuestos explícitos, indicadores de alerta, reglas de decisión escritas, revisiones periódicas, acuerdo con el donante.
- Revisión posterior a la acción: lo previsto, lo realizado, por qué la diferencia, qué hacer de otra manera; sin buscar culpables.
- Una lección es específica y accionable; si no, no es más que una impresión.
- Informe: meta, logrado, desviación, explicación, consecuencia, medida; beneficiarios únicos distintos de las participaciones.
- La sostenibilidad se juzga por las condiciones reunidas (responsable, recursos, competencias, demanda), no por una declaración.

Cuestionario

1. ¿Qué formulación es una lección aprendida útil?
   A. Hay que reforzar la coordinación
   B. Anunciar las distribuciones por radio y mensaje de voz 5 días antes, porque el 40 % de los hogares no estaba informado
   C. Hay que comunicar mejor
   D. El proyecto fue un éxito

2. En una revisión posterior a la acción, la regla principal es:
   A. Identificar al responsable del error
   B. Hablar de los hechos sin buscar culpables, y terminar con acciones con un responsable
   C. Redactar un informe largo
   D. Invitar al donante

3. Un informe cuenta 5 000 participantes en la formación y 5 000 en la distribución, es decir, «10 000 beneficiarios». El error es:
   A. Un posible doble conteo: hay que distinguir beneficiarios únicos y participaciones
   B. Ningún error
   C. Un error de cálculo de porcentaje
   D. Un error de tipo de cambio

4. ¿Qué cambio exige generalmente una enmienda formal con el donante?
   A. Cambiar de método de sensibilización
   B. Modificar el objetivo específico o las zonas de intervención
   C. Ajustar en dos semanas el calendario de una actividad
   D. Cambiar el orden del día de una reunión

5. El cierre de un programa se prepara:
   A. El último mes
   B. Desde el diseño con una estrategia de salida, y se pone en marcha al menos tres meses antes del final
   C. Después del informe final
   D. Solo si el donante lo pide

Respuestas del cuestionario : 1. B · 2. B · 3. A · 4. B · 5. B