

¿EVIDENCIA INSUFICIENTE?

La calidad y el uso de la evidencia en la acción humanitaria

Paul Knox Clarke y James Darcy

ALNAP es una red única para todo el sistema dedicada a mejorar el desempeño de la acción humanitaria a través del aprendizaje compartido.

www.alnap.org

¿Ha leído **La Guía de Introducción al Debate de ALNAP** que acompaña este estudio?

www.alnap.org/pool/files/discussion-starter-evidence-alnap-2014.pdf

Una copia electrónica de este estudio, la guía de introducción al debate y otros recursos relacionados están disponibles en el sitio web de ALNAP en www.alnap.org/ourwork/evidence.

Cita sugerida:

Knox Clarke, P. y Darcy, J. (2014) *¿Evidencia Insuficiente?: La calidad y uso de la evidencia en la acción humanitaria*. Estudio de la ALNAP. Londres: ALNAP/ODI.

© ALNAP/ODI 2014. Este trabajo está autorizado bajo la Licencia de Atribución No Comercial Creative Commons (CC BY-NC 3.0).

ISBN 978-0-9926446-4-2

Editado y corregido por Angela Hawke

Traducido por Milene Fernández

Diseñado por Soapbox



Contenidos

Reconocimientos	3
Introducción	5
1.1 ¿Por qué la evidencia es importante para la acción humanitaria?	5
1.2 Acerca de este documento	6
¿Qué significa “evidencia” en el contexto de acción humanitaria y cómo juzgamos su calidad?	9
2.1 Antecedentes: términos, conceptos y desafíos	10
2.2 ¿Para qué sirve la evidencia? Las propuestas humanitarias y los requisitos de la evidencia	12
2.3 Criterios para evaluar la evidencia humanitaria	15
¿Cumple la evidencia humanitaria actualmente con los criterios de calidad?	20
3.1 Los desafíos de generar evidencia en situaciones humanitarias	20
3.2 Evidencia de los sistemas de alerta temprana	22
3.3 Evidencia de la evaluación de necesidades	25
3.4 Evidencia del monitoreo	34
3.5 Evidencia de las evaluaciones y ensayos controlados	42
¿Se usa actualmente la evidencia para guiar operaciones humanitarias y elaboración de políticas?	50
4.1 La calidad no garantiza el uso	50
4.2 Uso de la evidencia de alerta temprana	50
4.3 Uso de la evaluación y monitoreo de la evidencia	54
4.4 Uso de la evidencia de evaluación	59
Conclusiones y recomendaciones: ¿Cómo se puede mejorar la calidad y el uso de la evidencia en la acción humanitaria?	73
5.1 Mejorando la calidad de la evidencia	73
5.2 Mejorando el uso de la evidencia en la toma de decisiones humanitarias	76
Anexo 1: Metodología	80
Anexo 2: Lista de entrevistados y colaboradores	81
Bibliografía	82

Reconocimientos

La Secretaría General de ALNAP realizó este documento en base al artículo *Evidence and Knowledge in Humanitarian Action (Evidencia y Conocimiento en la Acción Humanitaria)* desarrollado en la 28ª Reunión Anual de ALNAP (marzo de 2013). El documento original fue escrito por James Darcy y Paul Knox Clarke, con la asistencia en investigación por parte de Miriam Bradley y Alexandra Warner. Leah Campbell brindó asistencia de investigación al presente estudio. Los autores quisieran expresar su agradecimiento a todos aquellos que participaron en la Reunión Anual, y particularmente a aquellos que hicieron presentaciones, cuyas experiencias y conocimientos forman un importante elemento de este documento revisado. Los autores también quisieran agradecer a todos aquellos que aceptaron ser entrevistados para la realización de este documento, cuyos nombres se pueden encontrar en el Anexo 2. Finalmente, los autores están en deuda con Richard Garfield, Joanna Macrae y Anthony Redmond, quienes como revisores pares, hicieron sugerencias que han mejorado significativamente la calidad de este informe. Los autores siguen siendo completamente responsables por los puntos de vista expresados y de cualquier error u omisión en el texto.



Introducción

1.1 ¿Por qué la evidencia es importante para la acción humanitaria?

El fracaso para generar y utilizar la evidencia en la política y la respuesta hace que la acción humanitaria sea menos efectiva, menos ética y tenga menos capacidad para rendir cuentas.

Las respuestas efectivas de la acción humanitaria dependen tanto del conocimiento como del financiamiento o la logística. Con certeza, sin conocimiento, no se puede esperar que las respuestas tengan éxito: "La habilidad de la comunidad humanitaria internacional para recopilar, analizar, difundir y actuar en base a la información clave es fundamental para lograr respuestas efectivas" (Mahmood et al., 2010).

Igualmente importante es la evidencia si sostenemos principios humanitarios en un sector firmemente afianzado en valores: "Si cree en la imparcialidad, tiene que basarse en la evidencia. No puede ser imparcial si no sabe cuál es el rango de opciones" (Peter Walker, comunicación personal, marzo de 2013). Un reciente estudio del Departamento para el Desarrollo Internacional (DFID) sobre los desafíos de recopilar evidencia en situaciones de emergencia resalta la necesidad de seguir marcos éticos al hacerlo y también señala que es "poco ético efectuar intervenciones que en el mejor de los casos no han sido probadas, no son efectivas o en el peor de los casos, provocan un daño real." (DFID, 2012: 11).

Finalmente, la evidencia es importante para la rendición de cuentas. Las organizaciones donantes, los estados afectados y la sociedad civil tienen una expectativa legítima de saber cómo – y cuán correctamente – se gasta el dinero en nombre de las personas afectadas por las crisis. Para cumplir estas expectativas, las organizaciones humanitarias deben ser capaces de probar que las necesidades existen y demostrar que han hecho elecciones informadas y deliberadas acerca de la forma más eficiente y efectiva de responder. También deben ser capaces de brindar evidencia sobre el impacto de las elecciones que hagan.

Este documento (y la reunión de ALNAP en el que se basa) está respaldado por la percepción de que "actualmente, las decisiones humanitarias por lo general se basan en una información deficiente" (DFID, 2012: 5) y están "guiadas por una anécdota en lugar de una evidencia" (Mazurana et al., 2011:1). Incluso cuando la evidencia está disponible, las decisiones parecen estar guiadas por la convicción personal o por consideraciones políticas o de recaudación de fondos. Entonces es importante que la evidencia disponible sea de la mejor calidad, pero también es importante que los responsables de la toma de decisiones *usen* esta evidencia. Las fallas colectivas recurrentes para responder firmemente frente a una gran



Las respuestas efectivas de la acción humanitaria dependen tanto del conocimiento como del financiamiento o la logística.



evidencia de inminente crisis (especialmente de los sistemas de alerta temprana de situaciones de hambruna en el África Subsahariana) confirman que *generar* dicha evidencia es solo una parte del desafío. Lo mismo es aplicable a la evidencia de experiencias pasadas: un tema recurrente de las evaluaciones es que el sistema internacional y sus organizaciones individuales tienen problemas para aprender las lecciones del pasado y aplican su evidencia a las prácticas humanitarias actuales (Sandison, 2006; Hallam, 2011).

Dichas preocupaciones no son exclusivas de la acción humanitaria. Muchas áreas de la política pública – incluyendo la salud, educación y desarrollo internacional – han reconocido en los últimos 15 años la necesidad de integrar la evidencia a la práctica de un modo más sistemático. Stern et al. (2012) describen el incremento del movimiento de la Política Basada en la Evidencia (EBP), que argumenta que los elaboradores de políticas deben tomar decisiones basadas en la evidencia en lugar de basadas en una ideología o en respuesta a intereses especiales (véase también Davies et al., 2000; Nutley et al., 2007; NAO, 2003). Este mayor énfasis en la evidencia ha generado un debate sobre el significado real del término, sobre lo que se considera como "evidencia" y sobre los mejores enfoques para generar evidencia. Este documento considera algunas de estas preguntas en el contexto específico de la respuesta internacional ante emergencias humanitarias.

La generación y el uso de evidencia en la acción humanitaria no es un tema nuevo: los debates sobre la evidencia formaron parte del impulso para una mejor calidad y rendición de cuentas en la acción humanitaria en la década de los 90. Desde entonces, ha existido un sinnúmero de iniciativas técnicas para mejorar la calidad de la información proporcionada por la alerta temprana, los mecanismos de valoración y evaluación que pretenden generalmente asegurar que las respuestas humanitarias "se basen en las necesidades", y se apoyen en la evidencia de "qué funciona".

También han existido intentos para asegurar un amplio uso de la evidencia en la acción humanitaria. Muchas organizaciones han intentado fortalecer sus propios sistemas para el aprendizaje organizacional y para la gestión del conocimiento, mientras que el sistema en su totalidad ha respaldado esfuerzos colectivos para el aprendizaje conjunto, codificación y el establecimiento de normas (Walker y Purdin, 2004; Young y Harvey, 2004; Barnett, 2005). Sin embargo, aún existe un modo de mejorar tanto la calidad como el uso de la evidencia en el sector humanitario.

EN RESUMEN, EL DOCUMENTO ENCUENTRA QUE:

- A pesar del progreso durante los últimos 20 años, parece que hay espacio para la mejora en la calidad y en el uso de evidencia en la acción humanitaria internacional.
- La evidencia es importante: el uso de la evidencia de buena calidad mejora la efectividad y la rendición de cuentas de la acción humanitaria, además cumple con la ética y los principios humanitarios.
- Los debates sobre la evidencia en la acción humanitaria pueden ser vistos en el contexto de debates más amplios alrededor de la generación y el uso de evidencia en la política pública.

1.2 Acerca de este documento

El presente documento revisa la calidad de la evidencia disponible actualmente para apoyar la acción humanitaria. Se enfoca principalmente en la evidencia generada por el "sistema humanitario internacional formal"¹ a través de la alerta temprana, evaluación de las necesidades, monitoreo y retroalimentación, evaluación y valoración del impacto. También considera hasta qué grado los actores en el sistema humanitario realmente utilizan la evidencia para guiar sus decisiones operacionales y la formulación de sus políticas. Este documento propone el uso de seis criterios principales para la calidad en la evidencia:

- Exactitud
- Representatividad
- Relevancia
- Generalidad de las conclusiones
- Atribución
- Transparencia alrededor del contexto y los métodos.

El documento se dirige a los elaboradores de políticas humanitarias, a los responsables de la toma de decisiones humanitarias (en niveles estratégicos y operacionales), y a aquellos involucrados en la recopilación de información y el análisis en organizaciones humanitarias. También está destinado para ser leído por académicos y estudiantes de acción humanitaria, como una contribución al gran debate y discusión sobre este tema.

Las preguntas específicas consideradas en este documento son:

En la sección 1: ¿Por qué la evidencia es importante para la acción humanitaria?

En la sección 2: ¿Qué significa "evidencia" en el contexto de acción humanitaria y cómo juzgamos su calidad?

En la sección 3: ¿Cumple la evidencia actual con los criterios de calidad?

En la sección 4: ¿Se usa actualmente la evidencia para guiar operaciones humanitarias y elaboración de políticas?

En la sección 5: ¿Cómo se puede mejorar la calidad y el uso de la evidencia en la acción humanitaria?

El documento se basa en una revisión de la literatura, incrementado por las entrevistas y en el contenido de las presentaciones hechas en la 28ª Reunión Anual de ALNAP "Evidencia y Conocimiento en Acción Humanitaria" (Washington DC, marzo de 2013). En el Anexo 1 se proporcionan más detalles de la metodología.

¹ Un término utilizado para describir a las ONG nacionales e internacionales, los organismos de las Naciones Unidas, los miembros de la Cruz Roja/Media Luna Roja, y las entidades gubernamentales patrocinadoras involucradas en el financiamiento, programación e implementación de la asistencia humanitaria internacional. El concepto del sistema humanitario formal que se usa en este documento retoma el utilizado en el documento Estado del Sistema Humanitario de ALNAP: véase (ALNAP, 2012: 16).

Recuadro 1. Definiciones

Para los propósitos de este documento, adoptamos las siguientes definiciones de trabajo.

Términos claves

Evidencia: información que ayuda a corroborar y probar/refutar la verdad sobre una propuesta específica.

Información: cualquier dato que pueda informar el entendimiento o la creencia, presentado en un contexto que brinda significado a los datos. La información puede ser verdadera o falsa. La información solo se vuelve evidencia cuando se relaciona con una propuesta específica.

Conocimiento: "creencia verdadera justificada". En el contexto de este documento, se entiende que el conocimiento deriva de la observación directa o de un conjunto de evidencia para informar un verdadero entendimiento de un tema particular.

Términos de investigación

Sesgo: cualquier forma de error sistemático (no aleatorio).

Hipótesis: una propuesta, respaldada por una cantidad limitada de evidencia, ofrecida como una explicación potencial para un fenómeno. Por lo general, una hipótesis es el punto de partida para una mayor investigación: se puede recopilar mayor evidencia para respaldar más (o refutar) la hipótesis.

Cualitativo: (de investigación, análisis, datos) basado en la narración, en vez de en cifras. La investigación cualitativa tiende a relacionarse con las conductas y motivaciones humanas.

Cuantitativo: (de investigación, análisis, datos) basado en cifras, en vez de la narración. La investigación cuantitativa se basa en el análisis (estadístico) de una serie de datos.

Fiabilidad: Se relaciona con la consistencia de los resultados del mismo experimento cuando se repite: es una medida del grado hasta el cual el mismo experimento producirá el mismo resultado bajo las mismas condiciones en diferentes ocasiones.

Método científico: un método o procedimiento que ha caracterizado a la ciencia natural desde el siglo XVII, que consiste en la observación sistemática, medición y experimento, así como la formulación, prueba y modificación de las hipótesis (Oxford English Dictionary).

Validez: un término con una variedad de significados relacionados, los cuales involucran hasta qué grado un argumento o conclusión experimental está "bien fundamentado". Generalmente en las ciencias sociales y en la investigación cualitativa, la validez se conecta con la fuerza de la relación entre una conclusión o un análisis y los datos observados (LeCompte and Goetz, 1982) o más ampliamente, con la fuerza de la relación entre una declaración y el fenómeno social al cual se refiere (Hammersley, 1990). En la ciencia experimental, la validez tiene una serie de significados específicos. La validez interna es una medida de fuerza de las relaciones causa y efecto observada a través de las pruebas estadísticas: el grado hasta el cual se puede atribuir un efecto particular a una causa particular. Por otro lado, la validez externa se refiere al alcance hasta el cual una relación de causa y efecto identificada en una situación particular se puede generalizar a otras situaciones.



¿Qué significa “evidencia” en el contexto de acción humanitaria y cómo juzgamos su calidad?

2.1 Antecedentes: términos, conceptos y desafíos

El presente documento se basa en unos cuantos conceptos principales que son importantes de aclarar desde el principio. El significado de **conocimiento** ha sido muy debatido por los filósofos a lo largo de los siglos y ese debate es el centro del tema filosófico de la epistemología (que concierne a la naturaleza y el alcance del conocimiento). Para nuestros propósitos, tomamos el significado de una "creencia verdadera justificada", en otras palabras, una creencia que se basa en alguna forma de "hecho". El conocimiento difiere de la información, que tomamos con el significado de cualquier dato que pueda informar entendimiento o creencia, presentada en un contexto que le da sentido.

Dentro del campo de la epistemología, la escuela empírica del pensamiento tiene una particular importancia para nuestro análisis. Para los empíricos, el conocimiento proviene de la experiencia basada en el sentido ("la evidencia de los propios ojos"). El empirismo se opone al idealismo, la tradición, la autoridad: la idea que algo es correcto porque la teoría lo dice, porque así lo hemos hecho siempre o porque eso nos dicen los expertos que hagamos. Esta es la actitud que ha caracterizado mucho el pensamiento científico occidental durante los últimos 400 años y que subyace al método científico: formando hipótesis y posteriormente probando – y modificando estas hipótesis según se requiera – en base al experimento y observación. En las ciencias sociales, el empirismo por lo general se reformuló como "positivismo", a través del cual las propuestas que no se pueden probar utilizando la observación y la experiencia no tienen sentido, a menos que sean ciertas por definición o por inferencia lógica.

Sin embargo, los enfoques científicos sociales no están limitados a los enfoques empíricos o positivistas incorporados en las principales corrientes de tradiciones occidentales. Desde mediados del siglo XX, el positivismo se ha cuestionado por puntos de vista conflictivos en las ciencias sociales. En términos generales, esto ha implicado un desafío ante la existencia de la verdad "absoluta" y el conocimiento separado del observador y con ello, la idea de estricta objetividad. Además ha involucrado un desafío para los modelos "lineales" de causa y efecto. En cambio, se hace hincapié en lo relativo y subjetivo, en las percepciones y conductas de las personas y en explicaciones de conductas y resultados informados social y políticamente más complejos. Este enfoque tiende a privilegiar a los métodos cualitativos sobre los enfoques cuantitativos con su estricta concentración en lo que se puede medir. Los enfoques contemporáneos para la antropología social y cultural se ajustan en gran medida a este paradigma e informan a muchas otras áreas de pensamiento científico, social y político.

La idea de "evidencia", como se utiliza en este documento, se relaciona estrechamente con el empirismo y el método científico. La evidencia es información que ayuda a demostrar la verdad o falsedad de una hipótesis o propuesta dada. En muchos casos, estas propuestas están relacionadas con la existencia o ausencia de una condición (desnutrición o una enfermedad epidémica) que se pueden verificar mediante la observación repetida y sistemática, llevadas a cabo según las metodologías probadas. Sin embargo, en algunos casos, las propuestas se relacionarán con las conductas o creencias de grupos de personas (generalmente diversos) y en el presente documento, aunque se siga ampliamente el enfoque del desarrollo y prueba de las propuestas; necesitaremos reconocer que una "realidad" única y objetiva no existe, y que nuestras propuestas en realidad se relacionan a ideas múltiples y quizás conflictivas. A su vez, esto requiere que investiguemos el valor de una variedad de diferentes tipos de información como evidencia y utilicemos una variedad de metodologías no experimentales con el fin de considerar el valor de esta información para respaldar la propuesta.

“
La evidencia es información que ayuda a demostrar la verdad o falsedad de una hipótesis o propuesta dada.

”

Si la evidencia humanitaria difiere en tipo, también difiere en calidad. Los datos nutricionales recopilados utilizando un método válido de evaluación nutricional es la mejor evidencia para la existencia de una crisis nutricional que un informe anecdótico de un gran número de niños debilitados (aunque ambos son un tipo de evidencia). Sin embargo, debemos ser cuidadosos de no asumir que las diferencias en la calidad de la evidencia se basan simplemente en la naturaleza de la evidencia (cualitativa y cuantitativa): un cuestionario estructurado, administrado a una muestra probabilística de la población, puede decirnos menos sobre las formas de mejorar un programa de medios de vida que una serie de entrevistas semiestructuradas bien conducidas. Una reciente publicación del DFID, *Assessing the strength of evidence (Evaluando la solidez de la evidencia)* "evita construir una jerarquía de diseños y métodos de investigación... [y] reconoce que diferentes diseños son más o menos apropiados para los diferentes contextos y diferentes preguntas de investigación" (DFID, 2013: 7). El desafío – al que nos referimos en la Sección 2.3 – es identificar los criterios para la calidad de evidencia de una variedad de diferentes tipos de información. En muchos casos, seremos capaces de generar una evidencia más sólida utilizando enfoques de "método combinado," que utiliza "un amplio espectro de evidencia que usa y triangula varios diseños y métodos de investigación" (ibid: 8).

Finalmente, debemos señalar que la calidad de la evidencia que estaremos preparados para aceptar también será diferente, dependiendo en gran parte de las circunstancias en las que nos encontramos. A veces, una evidencia anecdótica puede ser todo lo que tengamos para continuar y puede ser suficiente para desencadenar acciones de cierto tipo; aunque la acción generalmente será obtener más evidencia rigurosa en la cual basar próximas decisiones acerca de la acción.

EN RESUMEN:

- En este documento, definimos la evidencia como información que informa sobre una propuesta específica.
- Esta comprensión de la evidencia se determina principalmente en un entendimiento empírico y ampliamente "científico" del conocimiento.
- Sin embargo, la naturaleza de la acción humanitaria hace que lidiemos frecuentemente con "realidades" subjetivas y construidas socialmente, y

como resultado debemos estar dispuestos a considerar diferentes tipos de información como evidencia.

- La evidencia difiere en calidad: sin embargo, esta diferencia no se relaciona con el hecho que la evidencia sea de naturaleza cuantitativa o cualitativa.
- El umbral para el uso de evidencia dependerá de las circunstancias: en algunas circunstancias debemos estar preparados para aceptar evidencia de menor calidad.

2.2 ¿Para qué sirve la evidencia? Las propuestas humanitarias y los requisitos de la evidencia

Como se describe en la Sección 2.1, la evidencia comprendida apropiadamente es evidencia *para* lograr algo; específicamente, es la información o el análisis que respalda una propuesta o afirmación en particular. Dado que el concepto de evidencia está estrechamente relacionado con la idea de propuestas, necesitamos considerar la naturaleza de estas propuestas en el contexto humanitario y lo que podría ser requerido para demostrar su veracidad o falsedad. Sugerimos que es posible identificar dos grupos de propuestas vinculadas que son importantes para los procesos de toma de decisiones de los actores en el sistema humanitario formal:

- Tipo A: propuestas sobre la existencia de necesidades humanitarias que resultan de una crisis y;
- Tipo B: propuestas sobre "qué funciona" para hacer frente a estas necesidades

Dentro de estos dos "grupos" habrán varias propuestas específicas y cada una requerirá de evidencia (con frecuencia de diferentes tipos) para respaldarlas.

Propuestas sobre la existencia de necesidad

El primer grupo de propuestas (Tipo A) tratan con la existencia de necesidades que la comunidad internacional debe abordar. Dentro de este grupo, las propuestas específicas claves son:

- Existe una situación que es "crítica" para aquellos afectados, o existe una situación que aunque no es fundamental, podría llegar a serlo si no se aborda.
- La situación requiere la intervención de actores externos.

Estas propuestas se encargan principalmente de probar que se requiere una intervención humanitaria. La evidencia para respaldar dichas propuestas generalmente proviene de una alerta temprana y de evaluaciones; con frecuencia se basa en la medición de ciertos indicadores para describir la situación en relación a los aspectos claves de la vida y el sustento de las personas. En muchos casos, los resultados de estas mediciones se comparan con umbrales aceptados de crisis para mostrar que se ha cumplido con las condiciones requeridas ante una situación de "crisis" real o inminente.

La mayoría de crisis humanitarias, al menos como se ven mediante los ojos de los organismos humanitarios profesionales, se conforman de sub-crisis que son específicas del sector: por ejemplo, el efecto de las inundaciones catastróficas pueden enfrentar crisis de salud pública, seguridad alimentaria y refugio. Muchos de estos sectores tienen sus propios indicadores y umbrales, aunque en muchos casos las respuestas se basarán en menor medida en el grado hasta el cual los

síntomas han alcanzado umbrales explícitos y en mayor medida en el juicio de quien toma las decisiones. La elección del umbral y en algunos casos la falta de umbrales, origina preguntas sobre qué constituye una crisis y qué evidencia se necesita para demostrarla.

El desafío de identificar "una crisis" es particularmente crucial cuando la propuesta trata con el potencial de una futura emergencia, en lugar de una descripción de la situación actual. En este caso, la evidencia necesita mostrar que la situación se desarrolla de un modo en el que se considere que las condiciones de emergencia existirán en el futuro. En algunos casos, esto se puede hacer por medio de una comparación de información acerca de la situación actual con "desencadenantes" predeterminados para una acción temprana. Sin embargo, en muchos casos estos desencadenantes no existen. De hecho, existen argumentos legítimos sobre si los desencadenantes genéricos, que son independientes de contextos específicos, son realistas.

Idealmente, las propuestas de Tipo A también exigen una evidencia para mostrar que la crisis no se puede abordar efectivamente, a menos que la comunidad internacional se involucre. Sin embargo, en la práctica no siempre se solicita esta evidencia antes que se realice una intervención.

Propuestas sobre la efectividad de la respuesta

El segundo grupo de propuestas (Tipo B) se relaciona con cómo los actores internacionales deben responder a una crisis real o inminente. Dentro de este grupo, las propuestas específicas claves son:

- Una intervención de un tipo específico será (o fue) efectiva en prevenir o reducir los efectos de la crisis (o cualquier situación como una crisis) en formas definidas.
- La intervención propuesta es (o fue) la más apropiada disponible en el contexto, considerando la probable eficacia, la eficiencia de los costos, la relevancia según es percibida por la población local, etcétera. (la intervención de "mejor opción").
- La intervención se puede suministrar en la base propuesta, cumpliendo con las normas mínimas acordadas, en otras palabras es viable.

Estas propuestas radican en el centro de la elección de la respuesta al decidir cómo intervenir en una situación específica. La primera compara la eficacia de una intervención sin hacer nada. La segunda depende de demostrar por qué la intervención propuesta es la mejor elección entre las opciones disponibles y tiene una creciente importancia para los organismos humanitarios ya que miden el valor relativo de varios métodos de respuesta. En el pasado, muchos pudieron haber utilizado un rango bastante limitado de "respuestas predeterminadas", pero estas actualmente son desafiadas por los avances en el aprendizaje acerca de opciones alternativas. Por ejemplo, el incremento en la documentación de lecciones acerca de las alternativas hacia enfoques estándar para la inseguridad alimentaria y los medios de vida (ej. distribución de dinero en efectivo en lugar de ayuda alimentaria), está poniendo una creciente presión en los organismos para justificar su elección de respuesta (Maxwell et al., 2012).

La tercera propuesta se refiere al criterio utilizado para determinar cuál es la "mejor" intervención: por ejemplo, puede ser el costo-eficiencia; costo-eficacia;



El desafío de identificar "una crisis" es particularmente crucial cuando la propuesta trata con el potencial de una futura emergencia, en lugar de una descripción de la situación actual.



aceptación por parte de la población; la capacidad de llegar a todas las personas necesitadas; el grado en que la intervención se articula con otras intervenciones para cumplir todas las necesidades en forma coordinada; o alguna combinación de estas.² La evidencia relacionada con estos diferentes criterios requerirá diferentes grupos de información y diferentes métodos para la recopilación y el análisis.

Las variantes de estas propuestas, que consideran acciones pasadas, también son cruciales para la evaluación al considerar si una respuesta fue exitosa. En general, esto requiere evidencia de los resultados de la intervención y presenta el frecuente desafío complicado de probar que esta intervención específica conllevó a aquellos resultados específicos.

La evidencia requerida para abordar las propuestas humanitarias

Tanto las propuestas de Tipo A como de Tipo B tienden a requerir dos amplios tipos de evidencia: específica y general. La evidencia específica nos indica qué está sucediendo en un momento determinado y en un lugar determinado. La evidencia general se obtiene estudiando una serie de emergencias similares, con el propósito de poder decir algo razonable acerca de "cómo se desarrollan las situaciones de hambruna" o "cómo funcionan las intervenciones de salud pública". Al considerar ambos grupos de evidencia juntos, podemos situar qué está sucediendo en un contexto más amplio y podemos decir con cierta certeza, por ejemplo, si estas condiciones constituyen o no una crisis cuando se comparan con otras situaciones en otra parte, o si este es el tipo de situación en el que un tipo de respuesta particular es frecuentemente exitosa.

En algunos casos, la evidencia general – sobre cómo se desarrollan las situaciones de hambruna o cómo se puede brindar una mejor manera de refugio a los refugiados – se codifica en umbrales, desencadenantes o procedimientos estándar, todos los cuales se basan en la idea de que lo que ha pasado en otros contextos se puede extrapolar con certeza a este contexto específico. Estos desencadenantes y procedimientos pueden incrementar ampliamente la velocidad de respuesta y puede crear consenso de gran importancia sobre si se requiere una respuesta y qué forma debe tomar. Sin embargo, como cada situación de emergencia es diferente, debemos tratar a la evidencia más general –así como sus umbrales, desencadenantes y procedimientos resultantes– con algo de escepticismo y siempre asegurarse de que sea importante para la situación específica de interés.

También vale la pena destacar que la evidencia en torno a ambos grupos de propuestas rara vez se presenta en forma de medidas directas. Generalmente, nuestra evidencia toma la forma de indicadores: partes de información que no describen ni miden directamente el tema de interés, pero que señalan o "indican" el problema. Si bien los precios, estados nutricionales o rendimiento de los cultivos, no proporcionan una medida directa de inseguridad alimentaria, todos pueden apuntar a la misma. En general, los indicadores se utilizan en base a que son (relativamente) fáciles de medir y tienen una relación razonablemente fiable y consistente con el objeto que queremos conocer. Sin embargo, rara vez son decisivos y con frecuencia se tienen que usar en combinación. Una preocupación clave de la evidencia en la acción humanitaria, particularmente en las propuestas (Tipo A) de diagnóstico,

² Generalmente los criterios acordados para determinar el éxito de una intervención son aquellos propuestos en los criterios de evaluación de la OCDE DAC (véase Beck, 2006).

es la selección de indicadores: cuando se seleccionan los indicadores incorrectos, es posible tener una información perfectamente buena que no es relevante en absoluto para la propuesta específica.

Antes de continuar, es importante hacer una salvedad: los dos grupos de propuestas considerados en este documento y la evidencia que se utiliza para respaldarlos, surge de los requisitos operacionales de las organizaciones humanitarias internacionales. Aunque forman el enfoque de este documento, estos dos grupos de propuestas no necesariamente abarcan o admiten todo el conocimiento de una crisis existente. El documento se redacta principalmente desde el punto de vista de los actores humanitarios internacionales, pero debemos recordar que otros grupos involucrados en la emergencia, y en particular las personas afectadas de forma más directa, pueden tener otras preguntas, y se podrían requerir otras formas de evidencia y conocimiento para responderlas.

“
Cuando se seleccionan los indicadores incorrectos, es posible tener una información perfectamente buena que no es relevante en absoluto para la propuesta específica.
”

EN RESUMEN:

- La mayor parte de la evidencia recopilada por las organizaciones humanitarias internacionales está diseñada para tratar propuestas sobre la existencia de necesidad; o propuestas sobre la mejor forma de responder a esta necesidad.
- La evidencia relacionada con estas propuestas viene en diferentes formas y se relaciona a diferentes temas.
- Las decisiones humanitarias generalmente se toman comparando la evidencia de una situación específica con evidencia “general” sobre cómo se desarrollaron situaciones similares o cómo se abordaron en el pasado, pero se necesita precaución: la información “general” de respuestas previas no siempre puede ser relevante.
- Mucha de la información utilizada por los organismos humanitarios viene en forma de indicadores, pero se necesita tener cuidado en su selección para asegurar que tengan una relación fuerte y consistente con la condición que se está midiendo.

2.3 Criterios para evaluar la evidencia humanitaria

Nos encontramos con varios problemas al tratar de corroborar las propuestas señaladas en la Sección 2.2. Primero, necesitamos convencernos a nosotros mismos que las propuestas son verdaderas y se basan en la evidencia disponible. Sin embargo ¿cómo sabemos si podemos confiar en la evidencia? Segundo, ¿cómo juzgamos si la evidencia realmente respalda la propuesta en cuestión? En esta sección, consideramos algunos de los temas importantes y proponemos seis criterios para lograr una “buena evidencia”.

La evidencia en los contextos que consideramos viene tanto como: datos, ya sean cuantitativos o cualitativos, directos o indirectos; y como el resultado del análisis de dichos datos, como conclusiones de las valoraciones, las evaluaciones o los estudios. Al considerar si la información se puede utilizar como “evidencia”, es decir si aprueba o rechaza una propuesta, necesitamos ser conscientes que no es solo la calidad de datos que importa, sino también la calidad del análisis. Es muy posible que los buenos datos sean mal analizados, lo cual conlleve a conclusiones erróneas. Antes de decir que la información cuenta como evidencia, tanto los datos como

los métodos usados para analizar estos datos, se deben medir con los criterios específicos de calidad.

Como se sugirió anteriormente, la evidencia de diferentes tipos se puede describir en términos de un número de atributos claves, incluyendo veracidad, credibilidad, exactitud, fiabilidad y validez. Desafortunadamente, estos términos se usan en diferentes formas a través de diversas disciplinas y sectores, que dificulta la generalización de su significado. Para complicar más el asunto, algunos de estos términos se utilizan para describir la calidad de los datos y los métodos por los cuales los datos son recopilados y analizados.

En la presente hemos intentado resumir los seis principales criterios que pueden ser utilizados para evaluar la calidad de la evidencia en contextos humanitarios.

EXACTITUD



- i. **Exactitud:** *ya sea o no que la evidencia es un buen reflejo de la verdadera situación y un "verdadero" registro de lo que se está midiendo.* Las medidas antropométricas que se han realizado correctamente y las declaraciones de informantes claves que son una verdadera explicación de lo que creen los informantes, ambas pasan la prueba de exactitud. La exactitud puede verse comprometida por el informante o el observador/investigador a través de una falsificación consciente, sesgo inconsciente, mala interpretación o uso incorrecto de los instrumentos de medición.

REPRESENTATIVIDAD



- ii. **Representatividad de la evidencia:** *el grado en que la evidencia representa exactamente la condición del grupo de mayor interés.* Por ejemplo, la información que pasa la prueba de exactitud refleja la situación en un pueblo o los puntos de vista de un grupo de la población no puede representar a todos los pueblos o grupos de población en la área. Esta es una pregunta del grado en que una muestra representa a toda la población.

RELEVANCIA



- iii. **Relevancia:** *el grado en el que una parte de la información se relaciona con la propuesta que se destina a probar o refutar.* Una parte de la información puede ser precisa y representativa, pero aún así puede ser irrelevante para el argumento o la propuesta que se hizo. El tema de **relevancia** se relaciona con los indicadores. Estos no son medidas directas de una condición, sino que miden algo que señala o indica esa condición. Las medidas antropométricas proporcionan indicaciones confiables, por ejemplo, del estado nutricional: existe una fuerte relación entre el indicador y la condición, así que estas medidas son relevantes como la evidencia para probar la propuesta de que la "desnutrición existe en esta área". Sin embargo, las mismas medidas son menos efectivas como un indicador de la escasez de alimentos, ya que la relación entre la circunferencia del brazo y la falta de alimentos es menos fuerte (la desnutrición puede ser provocada por enfermedades o prácticas de alimentación deficientes). Como resultado, las medidas de circunferencia braquial (MUAC) son menos relevantes como evidencia para probar la propuesta de que "se requiere ayuda alimentaria". Con mayor generalidad, la prueba de *relevancia* se debe aplicar a cualquier información que se proporcione como evidencia en respaldo de un argumento particular o propuesta.

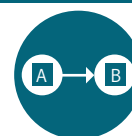
iv. Generalizabilidad de las conclusiones: *el grado en que la evidencia de una situación específica se puede generalizar más allá de esa respuesta a otras situaciones (y particularmente a situaciones en otros países, en otros momentos), y por lo tanto usada como evidencia más general de cómo se desarrollará una situación o del mejor tipo de respuesta. El tema de generalizabilidad es una particular preocupación para los elaboradores de políticas, que intentan crear políticas con aplicación global de la evidencia generada en contextos específicos. También está en el centro de este debate sobre los indicadores y umbrales globales comunes para desastres.*

GENERALIZABILIDAD



v. Atribución del análisis: *si el análisis demuestra una relación causal clara y certera entre dos condiciones o casos. Esto es importante para la calidad probatoria de la evaluación: donde los evaluadores identifican ciertos cambios en circunstancias que ocurrieron después de una intervención, ¿qué tan seguros pueden estar en que estos cambios que fueron el resultado de esa intervención específica y no de algo más?*

ATRIBUCIÓN



vi. Transparencia alrededor del contexto y los métodos: *el grado hasta el que está claro por qué, cómo y para quién se ha recopilado la evidencia. Como hemos visto, la evidencia es la información que se relaciona a una propuesta específica. Como tal, la información es solo evidencia en el contexto de una pregunta específica, hecha por una organización o grupo en particular. Un observador solo puede medir la calidad probatoria de cualquier información si conoce la propuesta con la que se relaciona la evidencia, quién quiso probar la propuesta y cómo recopilaron la evidencia. Sin información sobre el contexto, es imposible saber si la evidencia es relevante o generalizable. De igual manera, la información solo se puede aceptar como evidencia cuando los métodos utilizados para recopilarla y analizarla, así como cualquier limitación en el ejercicio, se realizan explícitamente. Solo en base a esta información es que el usuario puede determinar la exactitud, representatividad, relevancia, generalizabilidad y atribución de la "evidencia".*

TRANSPARENCIA
ALREDEDOR DEL
CONTEXTO Y LOS
MÉTODOS

Esta clasificación de criterios por ningún motivo es autoritaria y hay un número de otros enfoques para examinar la calidad de la evidencia. En un documento de 2005 sobre el uso de la evidencia para la elaboración de políticas de Louise Shaxson en Spencer et al. (2003), se sugieren cinco componentes que, juntos definen la "contundencia" o fuerza de la evidencia en términos de política: credibilidad, fiabilidad, objetividad (falta de sesgos), enraizamiento³, y generalizabilidad (Shaxson, 2005). El grupo de Organizaciones no Gubernamentales para el Desarrollo British Overseas (BOND) propone cinco principios para evaluar la fuerza de la evidencia producida por las ONG (para el desarrollo y humanitarias): ya sea que las perspectivas de las personas que viven en pobreza están incluidas en la evidencia; o la evidencia sea generada a través de métodos que son justificables dada la naturaleza de la evaluación; el grado en el cual los datos son triangulados; el grado en que la evidencia explora cómo ocurre el cambio y si la evidencia muestra los detalles de las fuentes de los datos y métodos utilizados, los resultados logrados y cualquier limitación en los datos o conclusiones.

³ Enraizamiento" se trata de si el problema que está siendo tratado por la evidencia verdaderamente representa la totalidad del asunto tratado o si hay otros aspectos que pueden y deben ser explorados. La idea se relaciona con la transparencia alrededor del contexto y los métodos señalados anteriormente.

John Gerring, al proponer un marco unificado para metodologías sociales científicas, sugiere algunos criterios que aplican a todos los argumentos de la ciencia social (propuestas), incluyendo verdad, precisión, generalidad, coherencia, conmensurabilidad y relevancia. Pero como él observa, "Penosamente, el vocabulario asociado con el tema de la metodología está lleno de ambigüedad. Términos claves... significan diferentes cosas en diferentes tradiciones de investigación y en diferentes contextos de investigación" (Gerring, 2011: 16). Entonces no debería sorprendernos que una sola lista de criterios acordados sea difícil de producir. Sin embargo, esperamos que los seis criterios dados, que reflejan el consenso general de los profesionales humanitarios y los académicos en la 28ª Reunión Anual de ALNAP⁴, brinden una herramienta útil a los actores humanitarios al considerar la fuerza de la evidencia que utilizan.

EN RESUMEN:

- Diferentes grupos de criterios existen para juzgar la calidad de la evidencia, aunque con frecuencia se refieren a un tipo específico de evidencia (cualitativa/cuantitativa) y por lo tanto no son generalmente aplicables a toda la evidencia humanitaria.
- Este informe utiliza seis criterios para juzgar la calidad de la evidencia generada y utilizada en la acción humanitaria: *exactitud; representatividad; relevancia; generalizabilidad; atribución; y claridad alrededor del contexto y los métodos.*

⁴ Una lista inicial, que incluye ampliamente los cinco primeros criterios, fue presentada en la reunión y debatida en el plenario. Aunque no hubo ninguna objeción a los criterios en sí, había un sentido general de que estos no incluían a quienes pertenecía esta evidencia, y para qué propósito era recolectada. Esta cuestión ha sido tratada en el sexto criterio. Los delegados en la reunión también sugirieron que se incluyera la "puntualidad" como un criterio. Esto no ha sido incluido en este documento, ya que parece que para los autores está más relacionado con la utilidad de la evidencia que con la calidad intrínseca de la evidencia en sí misma

3

¿Cumple la evidencia humanitaria actualmente con los criterios de calidad?

3.1 Los desafíos de generar evidencia en situaciones humanitarias

Hay un número de graves desafíos para la generación de buena evidencia en el sector humanitario, muchos de los cuales se originan del contexto en el cual operan los actores humanitarios. Los actores internacionales generalmente se involucran en respuestas donde las instituciones estatales son bastante débiles. Como resultado, los datos secundarios son muy limitados en las áreas donde se encuentran los trabajadores humanitarios: incluso los datos básicos de la población pueden ser difíciles de hallar (Recuadro 2).



Desafortunadamente, las circunstancias que hacen que la recopilación de información sea tan importante son precisamente aquellas que la hacen tan difícil.



Donde se encuentren disponibles buenos datos secundarios, la naturaleza de las emergencias humanitarias – que generalmente involucran altos niveles de mortalidad e importantes movimientos migratorios – significa que esa información puede caducar pronto. Las cifras de la población, tasas de desnutrición, información sobre refugios: todo puede cambiar rápidamente, creando un requisito para recopilar información continuamente a medida que la situación avanza.

Desafortunadamente, las circunstancias que hacen que la recopilación de información sea tan importante son precisamente aquellas que la hacen tan difícil. Muchas respuestas humanitarias ocurren en situaciones donde el acceso físico es altamente restringido, limitando las posibilidades para la recopilación de datos. Al mismo tiempo, la distribución de grandes cantidades de bienes y servicios en contextos donde hay una inmensa necesidad, inevitablemente tienen implicaciones políticas; sobre todo en "emergencias complejas" donde ya existen desavenencias políticas y facciones entre diferentes grupos dentro de una área. Bajo estas circunstancias, los agentes humanitarios no pueden asumir que la información, y especialmente la información de informantes claves, es objetiva o exacta. Tampoco pueden asumir automáticamente que ellos mismos están libres de sesgos.

Recuadro 2. El asunto de las cifras

Algunas de las cifras básicas que cuentan para las propuestas humanitarias son notablemente arbitrarias. En particular, las estimaciones y las cifras de la población dadas para las "cifras de afectados" o las "cifras de beneficiarios alcanzados" se ven afectadas por un alto grado de incertidumbre y una falta de claridad en la definición. Esto amenaza con debilitar la credibilidad de las propuestas que incluyen estas cifras.

A menudo, la raíz de este problema se basa en la falta de certeza acerca de las cifras referenciales (denominador) de la población. La mejor fuente de datos referenciales de población normalmente es el censo, pero los datos del censo pueden no estar disponibles a nivel local y cuando el censo se diseña y se utiliza como herramienta política, las cifras que proporcionan no son fiables. Aunque este no sea el caso, las cifras del censo generalmente están desactualizadas y mientras existan métodos aceptados para crear proyecciones de población, estos dependen de supuestos críticos que pueden ser difíciles de realizar con certeza.

A falta de datos del censo, los agentes humanitarios han recurrido a una variedad de enfoques: conteos rápidos, ya sea en el terreno o utilizando imágenes satelitales de casas/edificios; conteo de personas; listado y creación de perfiles; monitoreo de flujo y encuestas aleatorias de los hogares. Sin embargo, estos enfoques consumen tiempo y varían en exactitud. Desafortunadamente, tienden a ser más exitosos cuando ya existen los datos del censo.

Una consecuencia lamentable de esta falta de datos de referencia es que la diversidad dentro de una sociedad generalmente se oculta. Particularmente, las diferencias de género y edad que son un elemento importante de la mayoría de censos, rara vez se recopilan entre los grupos sociales evaluados y los datos desglosados por edad y sexo siguen siendo la excepción en vez de la norma.

Fuentes: Demographic Assessment Techniques in Complex Humanitarian Emergencies: Summary of a Workshop (2002) Committee on Population, National Academies Press, US; 'Sex and Age Matter', Tufts University (2011); State of the Humanitarian System, ALNAP (2012); Guidelines on Data Issues in Humanitarian Crisis Situations, UNFPA (2010); Technical Brief: Estimation of Affected Population Figures (ACAPS, 2012).

Las operaciones humanitarias también están marcadas por las acentuadas diferencias de poder que existen entre los extranjeros (quienes controlan los recursos y comprenden el sistema humanitario) y las personas afectadas por la emergencia. Los últimos pueden estar en gran necesidad. Al mismo tiempo, tienden a tener solo un entendimiento limitado de cómo funciona el sistema y tienen muy poco tiempo para completar cuestionarios o participar en las entrevistas. Esto no solo crea importantes desafíos prácticos para la recopilación de los datos primarios, sino que también conlleva a la preocupación de desafíos éticos: ¿hasta qué punto es legítimo solicitar información a las personas si ellas mismas no pueden beneficiarse directamente?, ¿hasta qué punto se puede solicitar un consentimiento informado a las personas afectadas sobre el uso y divulgación de información en un sistema que no comprenden, del cual esperan que les proporcione respaldo material? Como argumenta Bradt, "La reunión de datos y las consiguientes intervenciones humanitarias son procedimientos invasivos con consecuencias imprevistas. Las buenas intenciones no justifican los malos resultados" (Bradt, 2009:15).

Un mayor desafío ético se origina del hecho que la generación de evidencia requiere de fondos (y tiempo) que en la mayoría de los casos, provienen de los presupuestos que se podrían gastar en el apoyo directo a las personas afectadas, creando un difícil equilibrio entre el rigor probatorio y la necesidad para responder.



Las operaciones humanitarias en Afganistán y Zimbabue difieren bastante tanto en contexto como en naturaleza. El establecimiento de la evidencia "que funciona" en estas circunstancias puede ser extremadamente complejo.



Finalmente, las operaciones humanitarias con frecuencia se caracterizan por la necesidad de obtener información en muy cortos lapsos de tiempo. Es importante no exagerar este punto: los estados donde se ha producido el mayor gasto humanitario en los últimos cinco años se ven afectados por conflictos crónicos (ALNAP, 2012), y gran parte de la labor humanitaria que se produce en estos lugares toma la forma de programas de larga ejecución donde hay tiempo para recopilar y analizar la información (Macrae, 2013). Sin embargo, en situaciones de emergencia repentinas como el terremoto de Haití en el año 2010 o el tifón Yolanda (Haiyan) en el año 2013, el tiempo es escaso y esto afecta la capacidad de los actores humanitarios para recopilar suficiente buena evidencia a fin de respaldar la acción efectiva en las primeras fases de una respuesta.

Todas estas restricciones surgen, principalmente, a partir del contexto de la acción humanitaria en el proyecto individual o nivel de programa. Sin embargo, también existen problemas a nivel global o sistémico, no menos importantes por el hecho que las operaciones humanitarias en Afganistán y Zimbabue difieren bastante tanto en contexto como en naturaleza. El establecimiento de la evidencia "que funciona" en estas circunstancias puede ser extremadamente complejo en un sector donde las variables sociales y culturales son cruciales para el éxito de muchas operaciones.

El resto de esta sección considera las maneras en que las organizaciones abordan estos desafíos, mirando cómo se genera la evidencia y el conocimiento en la práctica del sector humanitario, así como preguntas relacionadas sobre la calidad de esta evidencia. Nuestro enfoque se centra en la práctica actual del sector ya que se relaciona a los elementos "estándar" de programas humanitarios: los procesos que generan evidencia de los sistemas de vigilancia y alerta temprana, evaluación de necesidades y procesos de monitoreo, así como evaluaciones de proyectos o programas. Consideramos la naturaleza de la evidencia generada por cada uno, especialmente como se juzgan frente a los criterios de "calidad" establecidos anteriormente en la Sección 2.

EN RESUMEN:

- Generalmente existe una necesidad urgente de información en situaciones de emergencia y crisis.
- Sin embargo, solo se dispone de datos referenciales limitados en muchos contextos en los que se requiere asistencia internacional.
- Recopilar información en emergencias puede ser extremadamente difícil, con limitaciones para el acceso; información que puede ser utilizada como herramienta política; y consideraciones éticas importantes que deben tomarse en cuenta.

3.2 Evidencia de los sistemas de alerta temprana

Los sistemas de alerta temprana se han descrito como "combinaciones de herramientas y procesos integrados en estructuras institucionales... [y] compuestos de cuatro elementos: conocimiento de riesgos, un servicio de monitoreo técnico y alerta, difusión de alertas significativas a las personas en riesgo y concientización

pública y preparación para actuar...".⁵ No todos los procesos de alerta temprana son parte de un sistema tan integrado. Algunos procesos de alerta temprana en situaciones de hambruna proporcionan una recopilación de datos independientes y análisis sin ser preceptivo acerca de la acción. Otros son menos formales y se basan más en la comunidad.

Los sistemas de alerta temprana combinan nuevos datos sobre una situación en desarrollo con conocimiento histórico para predecir el resultado probable para una área dada durante un lapso de tiempo. Si bien los datos son específicos del contexto, el análisis se basará tanto en el conocimiento del contexto (ej., tipos de medios de subsistencia) y el conocimiento de sucesos previos en este u otros contextos para realizar predicciones específicas de la situación. Con frecuencia ese análisis se basa en la tendencia: en otras palabras, su fuerza depende de la capacidad para establecer un caso convincente para una tendencia emergente; que si se deja sin verificar; conllevará a resultados catastróficos (propuesta de Tipo A anteriormente mencionada).

La alerta temprana contra la hambruna es el mejor ejemplo conocido de este tipo. El caso que se realiza es generalmente complejo, combinando datos de diferentes tipos, como los patrones de lluvia, rendimiento de los cultivos, precios de los alimentos, términos de comercio, ingresos de los hogares y niveles de desnutrición. Otros casos pueden ser más simples, como la alerta temprana que hace seguimiento a la trayectoria probable de un ciclón que se acerca o la repentina probabilidad de una inundación, o incluso un terremoto inminente o erupción volcánica, aunque estos son más difíciles de predecir.⁶ La alerta temprana puede ganar tiempo para realizar una acción preventiva, preparatoria o evasiva.

La alerta temprana es una de las áreas de práctica humanitaria que más ha avanzado en los últimos 20 años. Desde los sistemas de alerta temprana contra la hambruna como la Red FEWS y el sistema de Clasificación Integrada de las Fases (IPC) de la FAO en África, hasta los sistemas de seguimiento de ciclones en Asia y el Caribe, los avances tecnológicos combinados con sistemas nacionales y transfronterizos coordinados y la movilización comunitaria eficaz han ayudado a reducir la vulnerabilidad. Combinados con otros elementos de preparación (ej., refugios ante una inundación en Bangladesh, planes de ayuda alimentaria en el Cuerno de África, mecanismos de respuesta comunitaria en la India y América Central), estos sistemas han salvado muchas vidas.

La evidencia de la mayoría de los sistemas de alerta temprana generalmente parece tener una buena calificación en la *exactitud* cuando se compara con los criterios para la fortaleza de la evidencia. Sin embargo, muchos sistemas se enfrentan con temas de *relevancia*, relacionando los datos que se han recopilado con la propuesta de que una crisis que tendrá un gran impacto en las vidas y en los medios de subsistencia se aproxima.

En las crisis repentinas cuando (por ejemplo) los ciclones y las inundaciones se pueden rastrear de cerca, existe mayor posibilidad de decir con cierto grado de certeza si las circunstancias probablemente conllevarán a un impacto humano

⁵ Fuente AlertNet: www.trust.org/alertnet/news/early-warning-of-disasters-figures

⁶ UNISDR, 2006.



La complejidad dificulta la presentación de un argumento claro y convincente a los responsables de la toma de decisiones.



catastrófico y que áreas se verán afectadas. Sin embargo, en casos de desastres de evolución lenta, como las crisis de seguridad alimentaria relacionadas con la sequía puede ser mucho más difícil. En estas crisis, se combinan factores múltiples para determinar el impacto en las personas y en sus medios de subsistencia. Esto hace que la elección y la ponderación de los indicadores sea extremadamente difícil: ¿son el rendimiento de los cultivos o el precio de los cereales los mejores indicadores de una futura inseguridad alimentaria?, ¿y cómo se relacionan? La compleja interacción de factores también significa que los indicadores negativos no siempre conllevan a malas situaciones y a menudo entran en juego factores imprevistos. Esta complejidad también dificulta la presentación de un argumento claro y convincente a los responsables de la toma de decisiones.

La respuesta tardía ante la inseguridad alimentaria y los medios de subsistencia en el Cuerno de África durante los años 2010 y 2011 resaltó estas dificultades y enfocó la atención en varios sistemas de alerta temprana que operan en la región. En general, los observadores concluyeron que los sistemas "proporcionaron información precisa y oportuna que permitió a aquellos en posiciones de poder planificar y responder" (Hillier y Dempsey, 2012: 14; véase también Majid, 2011; Slim, 2012). Un informe conjunto de Oxfam GB y Save the Children señaló que "los informes de FEWSNET y FSNWG fueron calificados de "muy bueno" a "excelente" en términos de su exactitud para predecir la severidad y el inicio de la crisis"⁷ (Hillier y Dempsey, 2012: 14).

Sin embargo, incluso en estos sistemas efectivos, los datos dispersos conllevaron a algunas personas a cuestionar la *exactitud* y *representatividad* de los informes de alerta temprana (Levine et al., 2011). Este desafío estaba compuesto por el hecho de que los métodos utilizados para analizar los grupos de información múltiples e interrelacionados no eran transparentes; hubo una falta de *claridad alrededor del contexto y métodos* (ibid). Los autores del informe *System Failure: Revisiting the problem of timely response to crises in the Horn of Africa (Falla en el sistema: revisión del problema de la respuesta oportuna a las crisis en el Cuerno de África)* reconocen que una forma de abordar este problema de análisis e interpretación sería identificar "el indicador perfecto que no se pueda manipular y señalará cuál es la intervención necesaria en cada distrito" (Levine et al., 2011: 11). Sin embargo, en la resolución de un grupo de problemas en torno a la evidencia, esto crearía otro problema más, ya que ni un solo indicador pudo capturar con precisión la inseguridad de los medios de subsistencia en todos los contextos (en términos de la tipología señalada anteriormente, ningún indicador único sería *relevante* en todos los casos: un problema de *generalizabilidad*).

Paradójicamente, los debates sobre las mejores metodologías e indicadores para utilizar en la alerta temprana – debates que tienen por objeto mejorar la calidad de la evidencia – parece haber hecho que los responsables de la toma de decisiones sean más, en vez de menos escépticos, acerca del valor de los sistemas de alerta temprana. Volveremos al tema del uso de la evidencia en la Sección 4 y particularmente el uso de la evidencia de sistemas de alerta temprana.

⁷ Hasta cierto punto esto puede ser porque la Alerta Temprana trata con incertidumbre; con situaciones que pueden existir en el futuro, en lugar de situaciones que actualmente se pueden medir. Esta incertidumbre conlleva a que los responsables de la toma de decisiones soliciten un mayor nivel de evidencia y por lo tanto de certidumbre, antes de que actúen.

EN RESUMEN:

- La alerta temprana es una área que ha visto una importante inversión y mejora en los últimos años.
- El desafío probatorio clave en la alerta temprana es seleccionar los indicadores importantes que pueden predecir condiciones de manera fiable, sobre todo donde la alerta temprana hace frente a las crisis (como las hambrunas) que son el resultado de muchos factores que interactúan en el tiempo.
- Para tener éxito, los sistemas de alerta temprana necesitan ser claros sobre sus métodos, sobre todo para el análisis y la ponderación de muchos tipos de evidencia distinta.

3.3 Evidencia de la evaluación de necesidades

La "evaluación de necesidades", como describimos aquí, abarca una amplia variedad de actividades que tienen el propósito de identificar si se requiere la asistencia externa y si es así, señalar el tipo, cantidad y el período de la asistencia. Entonces, en el contexto de este documento, la mayoría de las evaluaciones de necesidades tienen como propósito brindar evidencia relacionada con las propuestas en torno a la existencia de una crisis (propuestas de Tipo A) y, en menor medida, en torno a la mejor manera de hacer frente a una crisis (Tipo B).

Las restricciones genéricas afectan las evaluaciones de necesidades para evidenciar la recopilación identificada en la Sección 3.1: falta de datos básicos; dificultades de acceso y falta de tiempo. Sin embargo, también enfrentan desafíos específicos que surgen de la naturaleza misma de la tarea. Sobre todo, las emergencias tienden a generar una amplia variedad de diferentes tipos de necesidad. Algunas de estas – y particularmente aquellas relacionadas con la salud y la nutrición– se pueden pronosticar y se pueden capturar mediante la medición de indicadores objetivamente verificables utilizando métodos y protocolos estándar. Sin embargo, otras necesidades, como la inseguridad alimentaria, son categorías conceptuales, que tienden a ser medidas por diversas combinaciones de indicadores. Como señalamos en la sección sobre alerta temprana, estos indicadores generalmente diferirán de un lugar a otro y como resultado, hay poco acuerdo común sobre cuál se debe utilizar y cómo se deben ponderar. Hay otras necesidades que tienen un fuerte componente subjetivo: sentimientos de amenaza o inseguridad, por ejemplo, señalan las necesidades que merecen atención pero son difíciles de capturar a través de métodos cuantitativos de medición (Mahmood et al., 2010).

Como resultado, las evaluaciones de las necesidades requieren, idealmente, individuos o equipos expertos en una variedad de diferentes disciplinas, con capacidad para recopilar, analizar y conciliar una variedad de diferentes tipos de evidencia. En la práctica, este requisito de sintetizar la evidencia y "llenar" cualquier vacío de evidencia deja que muchas evaluaciones de necesidades dependan en gran parte del conocimiento contextual de los individuos involucrados en la respuesta (véase Poole y Primrose, 2010; Assessment Capacities Project (ACAPS), 2012; ICRC y IFRC, 2008).

El conocimiento contextual también es importante para determinar la escala de la evaluación requerida. En algunos casos – donde los desastres repetitivos del mismo tipo han tenido consecuencias muy similares – el conocimiento contextual puede permitir que los agentes humanitarios eviten o minimicen la evaluación y vayan

directo al programa. También puede ayudar a identificar los mejores métodos y enfoques requeridos para realizar una evaluación en cualquier lugar determinado. La disponibilidad de la información y del personal capacitado así como la habilidad para acceder a la área de emergencia, cambiará de una situación a otra. Como resultado, "una medida no encaja en todo... no hay un paquete estándar único (evaluación) que se pueda desplegar en cada desastre. Más bien, las mejores prácticas acerca de buenos enfoques de evaluación se necesitan adaptar para cubrir las necesidades de información específicas para cada crisis. (Assessment Capacities Project (ACAPS), 2012: 6).

El peligro aquí es que sin un enfoque único de evaluación estandarizado, cada organismo en cada emergencia creará su propia metodología de evaluación para sus propias evaluaciones separadas. Una cierta cantidad de duplicación en la evaluación – particularmente donde se planifica y se coordina esta duplicación – probablemente sea algo bueno: las múltiples evaluaciones pueden incrementar la calidad probatoria de la información permitiendo que los resultados se comparen. Sin embargo, la situación en el pasado ha tendido a ser una donde la "mayoría de organismos tenían sus propios formatos de encuesta no estandarizados... que generalmente produjeron resultados conflictivos o repetitivos" (Darcy et al., 2013: 24). Obviamente, este no es un uso eficiente de recursos escasos. Diferentes organismos recopilan información similar, aunque fallan al llenar vacíos de información obvios⁸; las personas en comunidades afectadas por desastres pasan tiempo valioso respondiendo las mismas preguntas en múltiples ocasiones. Además, los informes resultantes generalmente se utilizan por el organismo que los creó porque no se comparten (Mills, 2005); no se comparten oportunamente (Darcy y Garfield, 2011); no son importantes para otras partes interesadas (Bradt, 2009), o no están en un formato que sea útil para otros actores (Poole y Primrose, 2010).



Parece que la respuesta se basa en la creciente coordinación entre las evaluaciones y en la creación y aceptación de normas comunes y metodologías para la recopilación de datos.



Parece que la respuesta se basa en la creciente coordinación entre las evaluaciones y en la creación y aceptación de normas comunes y metodologías para la recopilación de datos (DFID 2012: 31; Poole y Primrose, 2010: 1) que, si bien se pueden adaptar a contextos específicos, utilizan un lenguaje, indicadores y enfoques similares más amplios. Afortunadamente, en los últimos años se han visto grandes esfuerzos para incrementar la coordinación y desarrollar normas amplias, respaldadas por actores claves como IFRC/ICRC, el Grupo de Trabajo de Evaluaciones de Necesidades (NATF) de IASC y el Proyecto sobre Capacidades de Evaluación (ACAPS).

En términos de coordinación, parece que existe una tendencia generalizada hacia la aceptación de las evaluaciones conjuntas y coordinadas, que como Garfield et al. han señalado, tienen el potencial de incrementar la eficiencia y de proporcionar una base más sólida para la planificación y acción coordinada (aunque con los costos potenciales en términos de lentitud y gasto) (Garfield et al., 2011). En un ejemplo, mencionado en el informe State of the Humanitarian System (Estado del sistema humanitario), una evaluación descubrió que la evaluación conjunta de necesidades realizada por organismos que forman parte del Proyecto de Fortalecimiento de Capacidades en Situaciones de Emergencia (ECB) en el oeste de Java y Sumatra

⁸ Como la recopilación de datos separados por edad y género, véase (Mazurana et al., 2011).

permitieron una respuesta con una mayor cobertura geográfica y una mayor expansión sectorial (Wilson et al., 2010).

Mientras tanto, una revisión de la orientación producida por diversos "actores claves" en la evaluación y de recientes informes de evaluación a disposición del público⁹ sugiere un creciente consenso metodológico en torno a los enfoques: particularmente los enfoques para evaluaciones multisectoriales de fase temprana (aquellas evaluaciones realizadas en las primeras tres a cuatro semanas después de un desastre). El consenso que se ha empezado a generar parece ser sobre las evaluaciones basadas en una revisión de datos secundarios, aumentado en lo posible por la información recopilada de los informantes claves y de los grupos de discusión de personas afectadas utilizando un formato de cuestionario. En la mayoría de los casos, las comunidades que se visitan para las entrevistas se seleccionan de acuerdo a una muestra intencional, en lugar de una muestra aleatoria. Si bien la orientación de la ICRC/IFRC ya ofrece un formato de cuestionario sugerido (ICRC e IFRC, 2008), las directrices de la Evaluación Multisectorial Inicial Rápida (MIRA) de NAFT siguen siendo provisionales (IASC, 2012a) y están en un proceso de mejora (Joyce Luma, comunicación personal, abril de 2013). Mientras tanto, las ACAPs no brindan un formato de cuestionario, pero pueden considerarlo en el futuro (Richard Garfield, comunicación personal, abril del año 2013). En todos los casos, estos formatos no son prescriptivos, pero pretenden ser un "punto de partida" (con la expectativa que) los equipos agregarán preguntas" (Joyce Luma, comunicación personal, abril de 2013).

Teniendo en cuenta el estado actual de la evaluación de las necesidades humanitarias en contraste con los criterios de calidad probatoria propuestos en la Sección 2.3, quizás es inevitable que las evaluaciones de necesidades – las cuales formarán la base de la distribución de cantidades limitadas de recursos escasos y muy necesarios – encuentren problemas de *exactitud*. Las evaluaciones de necesidades proporcionan incentivos obvios para que las personas exageren la gravedad de una situación para incrementar el nivel de respaldo proporcionado. Al mismo tiempo, los Gobiernos pueden subestimar la magnitud de las necesidades para negar que exista una crisis.

Por lo general, la guía de evaluación de necesidades sugiere que este desafío de *exactitud* se aborda mejor al triangular la información del informante clave o de un grupo de discusión con los resultados de otros grupos o informantes, con fuentes secundarias, o, en algunos casos, con "constantes" como el consumo calórico necesario (véase Seaman et al., 2000). Una evaluación de la situación en Alepo en Siria brinda un ejemplo de rigurosa triangulación: se solicitó que los encuestadores evalúen la fiabilidad de los informantes claves y triangulen las respuestas de los informantes entre sí, con fuentes secundarias y con su propia observación. Cuando ocurrieron discrepancias que no se podían resolver, la información del informante

9 Los informes revisados fueron: The Philippines Multi-Cluster Needs Assessment for Tropical Storm Washi: Multi-Sector Initial Rapid Assessment, Pakistan Floods, 2012; Evaluation Initiale Rapide Multi-Cluster sur les Inondations au Moyer Chari, Tandjilé, Mayo Kebbi Est – Chad (todos estaban disponibles en Humanitarian Response info: <http://assessments.humanitarianresponse.info/mira-reports>. Un cuarto informe, disponible en la misma página, no se consideró) y Joint Rapid Assessment of the Northern Governorates of Yemen; Joint Rapid Assessment of Northern Syria – Aleppo City; y Rapid Initial Needs Assessment (de los cuales todos se encontraron al utilizar el término de búsqueda "informe de evaluación" en el sitio web de ACAPS. Otros informes adicionales identificados por la búsqueda no se consideraron).

no se utilizó (Assessment Working Group for Northern Syria, 2013). Una evaluación del año 2011 en Yemen tuvo un enfoque similar (ACAPS, 2011). Sin embargo, otros informes revisados para este documento no especificaron los métodos utilizados para establecer la veracidad de las respuestas del informante clave y de los grupos de discusión.

Un segundo desafío y con igual grado de importancia para la *exactitud* de las evaluaciones proviene de la confianza que es generalmente colocada en asesores expertos y analistas para "llenar los vacíos" o sintetizar diversas fuentes de información. Como resultado, los hallazgos de las evaluaciones pueden estar sesgados – de forma consciente o, mayormente, de forma inconsciente – por los prejuicios de los profesionales de ayuda en el terreno. Una vez más la triangulación tiene un rol que desempeñar aquí: La guía de la Cruz Roja/Media Luna sugiere establecer diversos equipos de evaluación que verifiquen el trabajo de cada uno, mientras la guía de IASC enfatiza la importancia de explicitar cualquier sesgo y vacíos de conocimiento en el informe de evaluación y explicar cómo se llegó a esas conclusiones. En este contexto, también es importante hacer que las evaluaciones estén disponibles para el público con el fin de permitir que los hallazgos sean considerados y cuestionados por las autoridades fuera del país. Aunque todos los informes considerados para este documento abordaron el problema de los "sesgos del evaluador", fue interesante ver cómo algunos informes de evaluaciones de los últimos años ya son accesibles por internet.

“

Es importante hacer evaluaciones disponibles para el público con el fin de permitir que los hallazgos sean considerados y cuestionados por las autoridades fuera del país

”

Otro desafío para la *exactitud* de las evaluaciones es el hecho de que, en las emergencias, la situación (y las necesidades resultantes) puede cambiar extremadamente rápido. Una evaluación puede brindar una verdadera representación de la situación en un pueblo el día en que el equipo los visite, pero podría estar desactualizada para el momento en que el informe sea completado. El enfoque señalado en la guía provisional de IASC tiene como objetivo asegurar la continuidad de la información de la evaluación a lo largo del tiempo a través de un proceso gradual, en el cual las evaluaciones multisectoriales rápidas son seguidas por evaluaciones detalladas y dirigidas por los clústers. (IASC, 2012a).

Sin embargo, hasta la fecha los clústers solo han tenido un progreso desigual en el desarrollo de una guía sobre estas evaluaciones detalladas, y estas aún no han sido realizadas de forma consistente (Richard Garfield, comunicación personal, abril de 2013). Tampoco una sola nueva ronda de evaluaciones sectoriales resolvería el problema – lo que parece tanto conceptual como práctico: la comunidad humanitaria todavía tiende a ver las evaluaciones como un evento "único", en vez de verlas como procesos continuos, y la evaluación efectiva sufre de las mismas limitaciones que el monitoreo de programas dando como resultado (véase más adelante) – particularmente una falta de financiamiento y soporte institucional.

Las zonas que ven emergencias regulares y donde uno esperaría se hayan desarrollado enfoques efectivos y consistentes sobre las evaluaciones de necesidades durante los últimos 20 años, aún son atendidas de forma deficiente (véase Poole y Primrose, 2010; Slim, 2012). Incluso en esos contextos (poco frecuentes) donde las evaluaciones secuenciales han sido realizadas, el uso de diferentes cuestionarios y métodos de evaluación han hecho imposible comparar resultados y por lo tanto comprender los cambios en las necesidades con el paso del tiempo. La situación descrita en la respuesta de la Evaluación de Tiempo Real (RTE)

del Cuerno de África puede representar la mayoría de operaciones humanitarias: aunque la calidad de las evaluaciones fue buena, el sistema "carecía de monitoreo continuo... (la información) estaba frecuentemente desactualizada para el momento en que era procesada y leída por el HTC (Equipo Humanitario de País) (Slim, 2012: 10).

El segundo elemento de calidad probatoria considerado en este informe es la *representatividad*: en este caso, el grado en el cual la información recopilada para la evaluación es representativa del grupo más grande que requiere ayuda. Como se sugirió anteriormente, la mayoría de evaluaciones de necesidades utilizan un muestreo intencional para seleccionar las comunidades, hogares o individuos que participarán en la evaluación. La guía de NATF y ACAPS enfatiza que, como resultado, los hallazgos de la evaluación "no pueden representar toda la población afectada por el desastre" aunque pueden utilizarse para "comprender las preocupaciones, cuestiones y necesidades más urgentes" (IASC, 2012: 21) y en muchos casos, formarán la mejor evidencia disponible para una intervención.

Un muestreo probabilístico aleatorio – que permite la extrapolación de resultados estadísticamente importantes provenientes del muestreo poblacional más amplio – es mucho menos habitual en las evaluaciones de emergencia, aunque esto sí ocurre. Las evaluaciones de necesidades en Pakistán y en Haití han intentado utilizar técnicas de muestreo aleatorio (ACAPS, 2010; UNOCHA y NDMA de Pakistán, 2012): la encuesta de Pakistán utiliza datos demográficos georeferenciados para establecer una muestra, la evaluación de Haití que depende del muestreo espacial, evalúa las áreas geográficas en base a una cuadrícula dibujada sobre las zonas afectadas, sin hacer referencia a la densidad o la distribución de la población. Ambos casos son instructivos. En Pakistán, el muestreo representativo solo fue posible porque el Gobierno tenía los datos demográficos bastante actualizados y porque tuvieron suficientes personas capacitadas disponibles. Mientras que esto incrementó la *representatividad* de los resultados, los encuestadores no pudieron alcanzar el 23% de pueblos muestreados. En Haití, el desplazamiento de la población, combinado con una falta de datos de la población, significó que mientras la muestra fue aleatoria, no fue necesariamente representativa (Garfield et al., 2011). Si bien sería insensato descartar un muestreo aleatorio en todos los contextos, los "métodos estadísticos (para un muestreo en las primeras etapas de las evaluaciones) normalmente no son factibles debido a las limitaciones de tiempo y acceso"(ICRC y IFRC, 2008), así como por la falta de información referencial, financiamiento y habilidades del personal.

Además de identificar a un grupo que pueda representar a toda la población – mediante un muestreo intencional o probabilístico – las evaluaciones de necesidades también necesitan recopilar datos de forma separada para diferentes sub-grupos en la población, que experimentarán diferentes tipos y niveles de necesidad. La evaluación efectiva y la comprensión de la necesidad requiere de un entendimiento de las diferentes situaciones de las mujeres, niños, ancianos, personas discapacitadas y aquellos que son socialmente marginados (Mazurana et al., 2011; Jacobsen y Furst-Nichols, 2011; Human Rights Watch, 2011; Nakagawa y Shaw, 2004).

En los pocos casos donde es factible y conveniente utilizar métodos estadísticos, el problema se puede abordar mediante el muestreo estratificado (si se utilizan métodos de muestreo aleatorio). Al utilizar técnicas de muestreo intencional, ACAPS sugiere que los datos secundarios se utilicen para identificar diferentes grupos económicos y sociales, así como diferentes zonas de medios de subsistencia. Sobre

esta base, las evaluaciones incrementan el tamaño de la muestra hasta que la situación para los diferentes grupos se haya descrito (ACAPS, 2012b). Sin embargo, aunque algunas de las evaluaciones de necesidades consideradas para esta revisión habían realizado evidentes intentos para obtener datos exhaustivos sobre las necesidades específicas de los grupos vulnerables – particularmente, la evaluación de Yemen considera las necesidades específicas de las mujeres en cada sector – el panorama general sigue siendo muy similar al descrito por Mazurana y otros autores: uno de "uso de SADD limitado, ad hoc y esporádico (Datos Disgregados por Sexo y Edad) en la Fase I y Fase II (evaluaciones)" (Mazurana et al., 2011: 3).

Las evaluaciones humanitarias también pueden enfrentar problemas relacionados con la *relevancia* de la información que incluyen y con los indicadores que utilizan para identificar la necesidad. Los informes de evaluación generalmente incluyen información porque está disponible y no porque traten de probar o refutar la existencia de las necesidades. Igualmente, y quizá más preocupante, las evaluaciones en el pasado han utilizado indicadores que han tenido solo una relación muy débil o indirecta con la situación que se está evaluando. Por ejemplo, los datos sobre rendimiento de las cosechas probablemente nos indiquen poco acerca de la seguridad alimentaria en áreas donde las cosechas en cuestión no son una fuente importante de alimentos para la población local. En la evaluación, "la comprensión del contexto es esencial para derivar el significado de... indicadores, ya que los medios de subsistencia y las estrategias de adaptación variarán" (Catley et al., 2008).

Probablemente el desafío probatorio más crítico para las evaluaciones surge en la área de la *claridad alrededor del contexto y los métodos*. El problema de fondo en la presente es uno de supuestos. Las evaluaciones de necesidades con frecuencia se perciben, e incluso a veces ellas mismas se presentan, como descripciones exhaustivas y objetivas de la situación en una zona como consecuencia de una situación de emergencia. Bajo este supuesto, las evaluaciones de necesidades se ven como la respuesta a la pregunta: ¿Cuál es la situación aquí? suponiendo que la información resultante se utilice por un amplio grupo de beneficiarios interesados. Evidentemente, la realidad es bastante diferente. La mayoría de las evaluaciones de necesidades en realidad responden a una variación de la pregunta: "¿Qué y cuánto necesitamos proporcionar? y la información principalmente es para el uso de los actores que recopilaron la información con el fin de argumentar el financiamiento y para planificar las operaciones.

Este problema de *transparencia alrededor del contexto y los métodos* conlleva a tres desafíos probatorios diferentes pero relacionados. El primero es una tendencia para suponer (en vez de probar) que existe una necesidad y que esta necesidad no se puede satisfacer mediante los esfuerzos de los actores locales: el Gobierno; las organizaciones de sociedad civil y las mismas personas afectadas. El resultado puede ser brindar asistencia que, en el mejor de los casos, no se necesita y en el peor de los casos "genera pasividad... debilita la iniciativa (y)...alimenta la dependencia e impotencia" (Anderson et al., 2012: 22).

Los cuestionarios de evaluación en el pasado han tendido a ser estructurados alrededor de estas suposiciones (con preguntas como: "¿cuántas personas carecen de un refugio adecuado? y "¿cuántos hogares en tu comunidad enfrentan la escasez de alimentos hoy en día?") y a menudo han pasado por alto las capacidades de estructuras locales y existentes para responder a la situación. La guía más reciente

sobre la evaluación de necesidades enfatiza la necesidad de hacer preguntas abiertas que no suponen la existencia de una necesidad (IASC, 2012b). La guía también aclara que el propósito de la evaluación no es planificar operaciones, sino determinar si una operación es necesaria o no (ICRC y IFRC, 2008). Estos cambios parecen, en cierto grado, haber influenciado la práctica de las evaluaciones: la mayoría (aunque no todas) de las evaluaciones consideradas en esta revisión tomaron hasta cierto punto en cuenta las capacidades locales para responder a la necesidad. Quizás no resulte sorprendente que aquellas evaluaciones que se realizaron en colaboración con las autoridades u organizaciones (como las de Pakistán y Chad) parecían ser más fuertes en este ámbito (OCHA, 2012; UNOCHA y NDMA de Pakistán, 2012).

El segundo desafío se relaciona con el tamaño de la necesidad. El hecho de que las evaluaciones son la base para las decisiones de financiamiento y que no sean descripciones neutrales de la situación posterior a la emergencia, podría brindar un incentivo para que los organismos exageren la verdadera dimensión de la necesidad (para incrementar el financiamiento) o resten importancia a la verdadera dimensión de la necesidad (para ser "realistas" acerca del financiamiento disponible) (Darcy y Hofmann, 2003). Un ejemplo reciente de esto fue la falla de la evaluación de necesidades en Somalia en 2011, que fue parcialmente el resultado de que la "presión se mantuviera baja, mostrando una tendencia para formar las solicitudes humanitarias en relación a lo que... los donantes podrían dar" (Slim, 2012: 10).

Las evaluaciones, que describen las necesidades que el Estado no puede cubrir, también son actos intensamente políticos y como resultado las cifras pueden ser "rutinariamente reducidas" (Slim, 2012: 14) en la etapa de análisis (véase también UNOCHA y NDMA de Pakistán, 2012; Darcy et al., 2013). Nuevamente, el sistema humanitario parece ser consciente de la necesidad de realizar evaluaciones más imparciales: el establecimiento de ACAPS fue incentivado en parte por el deseo de proporcionar evaluaciones independientes a los responsables de la toma de decisiones. Sin embargo, aún queda mucho camino por recorrer antes que las evaluaciones de necesidad independientes estén disponibles para todas las emergencias: por ahora, "el sistema [aún] carece de un grupo cifras concluyentes sobre todas las personas en necesidad" (ALNAP, 2012: 42).

El tercer y último desafío gira en torno a la comprensión de la necesidad. Diversos estudios han demostrado que las percepciones locales de ayuda frecuentemente difieren de modo notable de aquellas de los actores internacionales (véase específicamente Dijkzeul y Wakenge, 2010). Particularmente, la comprensión y la priorización de las "necesidades" para las personas afectadas por los desastres difieren de un contexto a otro (Anderson et al., 2012; Harragin y Chol, 1998). A pesar de esto, la mayoría de evaluaciones de necesidades se organizan "sectorialmente", con preguntas que cuestionan ciertos tipos de necesidad predeterminados. Como resultado, estas brindan evidencia relacionada a las necesidades que los actores internacionales son capaces de abordar, en lugar de evidencia de la necesidad en sí.

De igual manera, las definiciones impuestas externamente de "grupos vulnerables" de quienes se espera que necesiten asistencia pueden diferir considerablemente de las percepciones y comprensión de las mismas personas afectadas por los desastres. El informe de ALNAP *El estado del sistema humanitario* del año 2012 encontró que "a través de diversos contextos y tipos de emergencias, los



La comprensión y la priorización de las "necesidades" para las personas afectadas por los desastres difieren de un contexto a otro.



evaluadores observaron un problema con nociones preconcebidas de vulnerabilidad, las cuales condujeron a intervenciones inapropiadas" (ALNAP, 2012: 50).

Existen pasos para abordar este tema. Por ejemplo, el DFID identificó recientemente un requisito para que los agentes humanitarios "involucren sistemáticamente a los beneficiarios en la recopilación y el uso de datos", teniendo en cuenta que esto puede requerir nuevas habilidades y enfoques (DFID, 2012: 31), mientras que en el terreno, las iniciativas como la Escala sobre Necesidades Percibidas en Contextos Humanitarios de Emergencia (HESPER) se diseñan para fortalecer un mejor entendimiento de la necesidad. El desarrollo de la tecnología de comunicaciones también ofrece oportunidades bastante mejoradas para que los agentes humanitarios accedan a los puntos de vista de las personas afectadas (Recuadro 4).

Sin embargo, se requiere de más trabajo. Si bien ha pasado algún tiempo desde el veredicto condenatorio de la Coalición de Evaluación del Tsunami (Telford y Cosgrave, 2006), el informe de ALNAP *El estado del Sistema Humanitario* del año 2012 encontró que "las encuestas sobre el terreno realizadas a los receptores de ayuda para este informe... revelaron que dos tercios de los encuestados dijeron que no estaban satisfechos o que solo estaban parcialmente satisfechos con la cantidad y la calidad del paquete completo de ayuda que habían recibido", y que una de las razones claves para esto fue el fracaso para comprender el contexto local (ALNAP, 2012) véase también (Grünwald et al, 2010).

Reconocer la comprensión local de "necesidad" en el diseño de un programa es un paso importante para asegurar programas efectivos (Featherstone, 2013). Sin embargo, la importancia de escuchar a los más afectados directamente y tomar en cuenta su realidad tiene un impacto que va más allá de la efectividad inmediata. Esto se demuestra por los resultados del "Proyecto de Escucha" de la CDA (Proyecto de Aprendizaje Colaborativo), que consideró conversaciones con más de 6,000 receptores de ayuda internacional, quienes dieron un mensaje muy consistente.

El mensaje es que las evaluaciones diseñadas por y para los actores internacionales contribuyen a una situación en la que, en emergencias, demasiada ayuda viene muy rápido: una situación que muchas personas ven como cumplir las necesidades institucionales de las organizaciones internacionales, en lugar de las necesidades en el terreno. Al mismo tiempo, se supone que las personas afectadas son esencialmente receptores pasivos de ayuda, que como hemos visto, frecuentemente subyace la idea que la "evaluación de necesidades" puede conllevar a una dependencia de largo plazo, y requisitos para una asistencia continua y creciente con el paso del tiempo. (Anderson et al., 2012).

Aquí un recordatorio importante de que las preguntas que hacemos y la evidencia que utilizamos para responderlas, no son neutrales. Como hemos visto, la evidencia siempre responde a una propuesta y alguien ha creado esa propuesta en base a ciertos requisitos y supuestos. Nuestra elección de preguntas, fuentes y métodos no solo reflejan el sistema humanitario que tenemos, sino que también crea el sistema humanitario del futuro.

EN RESUMEN:

- No hay un único grupo de preguntas de evaluación que pueda ser utilizado en todos los casos, y como resultado, ha habido una gran variedad en la naturaleza, métodos y calidad de evaluaciones de emergencia.
- Este problema está siendo mayormente abordado por medidas para estandarizar metodologías (mientras se deja espacio para la personalización) y para realizar más evaluaciones conjuntas/coordinadas.
- El conocimiento del contexto local y el juicio individual son elementos importantes en la mayoría de las evaluaciones.
- Las evaluaciones tienden a ser vistas como eventos únicos, pero son más efectivas cuando forman parte de un proceso de recolección de información y análisis a más largo plazo.
- Una variedad de enfoques están en uso para mejorar la *exactitud* de las evaluaciones.
- La mayoría de evaluaciones utiliza muestreos intencionales, y esto puede dificultar su *representatividad* de la población en conjunto.
- Un pequeño número de evaluaciones ha utilizado un muestreo probabilístico: sin embargo, esto requiere de recursos que frecuentemente no están disponibles (principalmente el personal capacitado y los datos de referencia).
- Generalmente, las evaluaciones siguen siendo deficientes en la representación de necesidades específicas de grupos sociales marginalizados y vulnerables.
- Muchas de las evaluaciones son diseñadas para satisfacer las necesidades de información específica de las organizaciones internacionales y están estructuradas alrededor de sus categorías conceptuales, conduciendo a evaluaciones que son imprecisas o no son relevantes para las verdaderas necesidades de las personas afectadas

Recuadro 3. Evidencia y normas humanitarias

La experiencia de establecer normas mínimas para la respuesta humanitaria nos dice bastante sobre la disponibilidad y la calidad de la evidencia en el sector humanitario.

Los equipos que desarrollan normas humanitarias como el Proyecto Esfera y las Directrices y Normas de Emergencia para el Sector Ganadero (LEGS) generalmente requieren de dos tipos de evidencia. La primera de estas, es la "evidencia referencial" que se utiliza para desarrollar indicadores que permitirán que los resultados de un programa se midan en comparación a la norma. Por ejemplo, si una norma sugiere que las personas deben recibir "suficiente agua para cubrir las necesidades básicas", se requerirá que la evidencia referencial sugiera cuántos litros de agua por persona son suficientes, en la mayoría de contextos. Esto permite que los organismos utilicen la norma para medir el efecto de su programa en contraste con un nivel mínimo de idoneidad.

El segundo tipo de evidencia trata sobre "qué funciona". En el manual del Proyecto Esfera, la evidencia sobre "qué funciona" se utiliza para sugerir acciones claves: "las actividades y aportes sugeridos" (Proyecto Esfera, 2011) que en la mayoría de los casos, conllevarán a que los organismos logren las normas en sus programas. En relación a una "evidencia referencial", el Proyecto Esfera basó algunos indicadores – notablemente aquellos relacionados con los requisitos humanos fisiológicos – en cifras provenientes de la investigación empírica. Por ejemplo, el indicador de 2,100 kilocalorías por día se basa en una serie de estudios proporcionados en conjunto por el Comité Internacional de Nutrición (CIN) en el año 1995 y posteriormente fue aprobado para su uso en situaciones de emergencia por la Organización Mundial de la Salud (OMS/ACNUR, 1997). Sin embargo, en muchos casos este tipo de evidencia no estaba disponible: tanto el Proyecto Esfera como las LEGS dependieron ampliamente de la experiencia directa de profesionales y de las opiniones de expertos que revisaron y debatieron la información que estaba disponible y llegaron a una decisión (Catley, 2013; Damerell, 2013).

También fue difícil conseguir la evidencia sobre lo que funciona. Cuando comenzó el Proyecto LEGS, no habían publicaciones revisadas por pares que abordara los resultados o impactos de los programas de ganadería en emergencias humanitarias, y las evaluaciones e informes de los organismos operacionales que estaban disponibles tuvieron muy poco valor probatorio. Generalmente, las evaluaciones se enfocaron en la producción (el número de ganado vacunado) en lugar de enfocarse en los resultados e impactos (los efectos de la vacunación en la morbilidad de las enfermedades y el impacto en los medios de subsistencia humanos). Por lo tanto, no podían brindar evidencia sobre lo "que funcionó" en términos de abordar temas humanitarios. Si bien esto puede reflejar las dificultades metodológicas generales en el establecimiento de la atribución, también refleja las fallas en el diseño del proyecto: muchos de los proyectos que se evaluaron no habían aclarado los resultados – en términos de salvar vidas y de proteger los medios de subsistencia – que esperaban lograr.

3.4 Evidencia del monitoreo

El monitoreo – la recopilación sistemática y continua de tipos específicos de información – toma varias formas en las operaciones humanitarias. Para los propósitos de esta revisión, hemos identificado cuatro grandes tipos de actividad de monitoreo:

- El monitoreo de la situación, que se relaciona a la evaluación de necesidades y responde a las propuestas de Tipo A señaladas en la Sección 2.2, que describen el contexto y que muestran si se requerirá o no una intervención.
- El monitoreo de programas, que se relaciona a las propuestas de Tipo B ("qué funciona"), que recopila información sobre la implementación y el desempeño de un programa humanitario.
- El monitoreo organizacional, que demuestra el desempeño de una organización en todos sus programas, y
- El monitoreo sistémico, que considera el desempeño global del sistema humanitario internacional en conjunto.

Monitoreo de la situación

Idealmente, las evaluaciones de necesidades humanitarias – particularmente en zonas donde las situaciones de emergencia son frecuentes o crónicas – formarían parte de un proceso más exhaustivo de monitoreo y comprensión de las necesidades humanitarias. En lugar de ser casos "únicos", las evaluaciones de necesidades pueden ser ocasionadas por, y marcar una intensificación de, el monitoreo continuo de la situación. Sin embargo, aunque existe un acuerdo general sobre la importancia del monitoreo de la situación (por ejemplo, véase ACAPS, 2012a; CICR y la FICR, 2008; Garfield, 2013), "la evidencia de la evaluación (de monitoreo) continua hace bastante falta – y el evidente vacío para intervenciones basadas en la evidencia" (Darcy et al., 2013).

Los sistemas de monitoreo de la situación existen: Médicos Sin Fronteras (MSF) de Bélgica ejecuta un sistema de vigilancia de enfermedades en la República Democrática del Congo (RDC) que monitorea mediante una red de "antenas"; mientras que en Etiopía, hay un sistema de vigilancia nutricional conjunto diseñado para identificar "zonas claves" de desnutrición que involucra al Gobierno, organismos nacionales e internacionales. Sin embargo, estos sistemas de monitoreo tienden a enfocarse en sectores específicos, y son, en todo caso, pocos y muy espaciados. Como se ha señalado en la sección anterior, el monitoreo de rutina de indicadores contextuales claves no sucede en Sudán del Sur (Poole y Primrose, 2010), Somalia (Slim, 2012) o DRC (Darcy et al., 2013), a pesar de una historia de compromiso humanitario que se remonta a más de dos décadas.

Esta falta de monitoreo de la situación parece ser, en gran medida, un problema de cultura e incentivos organizacionales. Las inversiones continuas, a largo plazo en los esfuerzos de monitoreo no encajan bien en los cronogramas a corto plazo de los programas humanitarios, e incluso cuando ha existido la oportunidad para recopilar datos episódicos durante un período de años, los organismos humanitarios, en muchos casos, han cambiado continuamente el método y los datos recopilados a lo largo del camino. De hecho, se han hecho pocos intentos para monitorear las situaciones por más de uno o dos años.

En algunos (aunque relativamente pocos) contextos los programas de "desarrollo" han respaldado la recopilación de información y actividades de monitoreo que pueden originar respuestas humanitarias (Mahmood et al., 2010). Sin embargo, existe mayor probabilidad que las inversiones en actividades a más largo plazo de esta naturaleza sucedan cuando el enfoque de ayuda internacional se traslade desde una provisión intermitente de ayuda en emergencia hasta un apoyo constante a las poblaciones vulnerables. Este es el caso en Etiopía, donde "lo que una vez fue una máquina de respuesta humanitaria masiva está siendo lentamente reinventada como una ayuda integrada, una máquina de bienestar y de reducción de la pobreza" (Darcy et al., 2013).

Monitoreo de programa

DAI mismo tiempo, la falta de un monitoreo de la situación efectivo también se relaciona a las debilidades en el segundo tipo de monitoreo, el monitoreo de programa. El tema aquí se relaciona con el problema probatorio de la *claridad alrededor del contexto y el método*: en torno a lo que se pregunta, por qué y para quién. En la misma forma en que las evaluaciones de necesidades de los organismos

tienden a no proporcionar medidas imparciales de necesidad total en una situación dada, los sistemas de monitoreo de los programas organizacionales tienden a concentrarse en el grado en que los aportes han sido exitosamente convertidos en resultados, más que en la forma en que estos resultados han influenciado los niveles de necesidad. Una organización humanitaria calculó que solo entre el 13% y el 32% de sus sistemas de monitoreo capturaban información sobre los efectos reales de los programas en el contexto humanitario (Willet, 2013).



Los sistemas de monitoreo de los programas organizacionales tienden a concentrarse en el grado en que los aportes han sido exitosamente convertidos en resultados, más que en la forma en que estos resultados han influenciado los niveles de necesidad.



Es comprensible que los donantes – quienes con frecuencia son los que hacen las preguntas que los sistemas de monitoreo están diseñados para responder – requieren información sobre cuánto dinero se gasta: sin embargo, en algunas circunstancias, esto puede conducir a un monitoreo que ampliamente pasa por alto si un programa humanitario hace realmente una diferencia ante la necesidad humanitaria (Argren, 2013). El problema se agrava por el hecho que la mayor parte del monitoreo se realiza en cada organismo de manera independiente: como resultado, puede ser difícil identificar los cambios acumulativos causados por los programas humanitarios en cualquier zona.

De hecho, el monitoreo del programa comparte muchos de los desafíos probatorios de la evaluación de necesidades: puede ser difícil establecer indicadores; para saber si la información recopilada es exacta y representativa; y para comparar los diferentes tipos de información. Sin embargo, el monitoreo del programa difiere en una forma importante: es un proceso continuo, en lugar de un proceso "único". Esto conlleva a desafíos probatorios adicionales, pero también puede mejorar la calidad de la evidencia generada por los mecanismos de monitoreo.

Hemos señalado que el acceso físico a ambientes inseguros plantea un desafío para la recopilación de mucha información humanitaria. Particularmente este desafío es agudo para el monitoreo del programa, que requiere un acceso frecuente y regular a las operaciones. El acceso limitado puede dificultar bastante la seguridad de la *exactitud* o *representatividad* de cualquier parte de la información, y de grupos de datos que van surgiendo con el paso del tiempo.

La mejora de las telecomunicaciones puede ayudar a que los organismos accedan a la información de las zonas inseguras (véase Recuadro 4) y los organismos también están empleando soluciones "humanas" innovadoras para abordar este problema. En el este de Myanmar, el Comité Internacional de Rescate (IRC) se ha unido con otros organismos para crear un sistema en el que, cada vez que un organismo tiene acceso a una zona, el sistema monitorea los proyectos de todos los organismos en la zona. Este enfoque reduce la posibilidad de sesgos por parte del monitor (y por lo tanto mejora la *exactitud*); mientras se incrementa el tamaño de la muestra, lo que permite una mayor triangulación y más verificaciones para ver si las condiciones monitoreadas son representativas. En Somalia, UNICEF ha tomado un enfoque bastante diferente para un problema similar, utilizando la validación de terceros para asegurar la calidad de la información que se recibe de los sistemas de monitoreo. Cuando las verificaciones de calidad básicas sugieren que pueden haber problemas con la información del monitoreo, UNICEF envía monitores independientes contratados para realizar una investigación de seguimiento. Estos monitores son "asignados al azar", por lo que es más difícil para ellos fabricar información. Nuevamente, el enfoque tiene la capacidad de mejorar radicalmente la *exactitud* de la información (Fetouh et al., 2013).

El monitoreo en la acción humanitaria tiende a incluir menos a las personas afectadas que algunas otras fases del ciclo del programa (Alexander et al., 2013), aunque las puede incluir de una manera más eficaz. La naturaleza de actividades de monitoreo continuas y a largo plazo permiten la inclusión más significativa, pero menos simbólica: hay tiempo para establecer relaciones y entablar un diálogo bidireccional que tiende a ocurrir en evaluaciones de necesidades o en una valoración. El reciente crecimiento en los mecanismos de quejas y retroalimentación, que suelen ser un elemento del monitoreo del programa, ha conducido a muchas organizaciones a experimentar con múltiples canales de retroalimentación, con el propósito de establecer sistemas que sean contundentes, de contexto específico y permitan que los grupos marginados proporcionen información sobre su experiencia de la ayuda (Bonino et al., próximo; Fenton et al., 2011). Al igual que en la evaluación de necesidades, existe cierta evidencia de que estos mecanismos mejoran la calidad probatoria de la información de monitoreo (Featherstone, 2013; Laybourne y Obrecht, 2013).

Recuadro 4. Tecnología y evidencia

Los avances en la tecnología de la información ofrecen la posibilidad de mejorar en gran medida la habilidad de los agentes humanitarios para recopilar y analizar datos y para obtener evidencia utilizada.

La tecnología de la información significa que los datos pueden ser reunidos rápidamente después del inicio de una crisis, y con regularidad a medida que se desarrolla la crisis. La tecnología de la información (IT) también hace que la evidencia sea más accesible y fácil de interpretar. Sin embargo, sin sistemas para procesar adecuadamente estos datos, esto puede volverse abrumador. Los agentes humanitarios necesitan desarrollar la capacidad para analizar, comprender y utilizar la evidencia reunida. Donde los recursos son escasos, o la capacidad es limitada, opciones de baja tecnología pueden ser más efectivas.

A medida que las nuevas tecnologías siguen desarrollando su potencial masivo para respaldar a la acción humanitaria basada en evidencia, una variedad de otras consideraciones necesitarán ser consideradas. Aunque se puede reunir grandes cantidades de datos rápidamente, no todos son considerados como evidencia útil. Al mismo tiempo, el incremento de la cantidad de datos reunidos puede generar expectativas que los agentes humanitarios no pueden cumplir. Y una creciente dependencia en la tecnología puede hacer que las organizaciones sean vulnerables a las fallas tecnológicas comunes en contextos posteriores a desastres.

Una preocupación clave se relaciona a la calidad probatoria de la información "multitudinaria" (es decir, información obtenida a través de un gran número de personas que utilizan la telefonía móvil o los servicios basados en Internet). Los avances tecnológicos hacen posible recopilar los datos de actores nuevos y diversos. Por un lado esto hace que sea más fácil triangular la información y representar un panorama completo; por otra lado la exactitud de la evidencia, y credibilidad de algunas fuentes puede ser cuestionada. Además, hay preguntas acerca de la representatividad. Si la evidencia es reunida a través de mensajes de texto, por ejemplo, la fuente de información es limitada para aquellos que tienen acceso a teléfonos celulares

El Proyecto de Mapeo Ushahidi en Haití brindó conocimiento preciso de la situación y una evaluación llegó a la conclusión que fue muy relevante para la alerta temprana. Sin embargo, pocos agentes humanitarios actuaron en base a esta información. Ellos tenían sospechas de la gente y estaban preocupados por la calidad probatoria y la representatividad. Los datos eran inapropiados para la rigidez de los requisitos de la información de las organizaciones encargadas de responder. Muchos temen que la evidencia de las plataformas multitudinarias sea sesgada o manipulada. Otros consideran que todos los datos tienen sesgos inherentes y la toma de decisiones humanitarias con frecuencia yace en la "evidencia" anecdótica. Tener datos sesgados puede ser mejor que no tener ningún dato.

La tecnología tiene el potencial para mejorar la rendición de cuentas a las personas afectadas por la crisis al mejorar el flujo de la información y la comunicación bidireccional, y para los donantes al brindar pistas claras de información. Esto puede hacer que la reunión de evidencia sea más participativa y más simple de llevar a cabo a gran escala. Sin embargo, también trae consigo problemas de protección de datos. El geoetiquetado (grabación de la ubicación donde se tomó la información/imágenes) puede ser útil en el análisis; sin embargo, tiene riesgos significativos de seguridad y privacidad. De manera similar, aunque los registros digitales ahorran tiempo y reducen errores de duplicado, muchos son aprehensivos sobre el almacenamiento de grandes cantidades de datos en servidores o sistemas basados en la nube. Los agentes humanitarios deberían desarrollar directrices para asegurar que la información se utilice de forma ética y segura.

Otra área donde la acción de todo el sistema puede ser útil es en el desarrollo de plataformas comunes. La evidencia es más útil cuando es de fácil acceso, y aunque la amplia variedad de aplicaciones de datos personalizados hace que los datos sean más exhaustivos para sus usuarios, son de uso limitado para las personas que usan otro sistema o plataforma.

Los agentes humanitarios deben reconocer las oportunidades que presenta la nueva tecnología para mejorar la calidad y el uso de la evidencia, y desempeñar un rol activo para asegurar que, a medida que se introduce la tecnología, aborde el tema de la calidad de datos, la confidencialidad, el acceso y la interoperabilidad.

Fuentes: Avila, R. et al (2010); Coyle, D. y Meier, P. (2009); OCHA (2013b); Smith, G. et. al. (2011); World Vision (2013); IFRC (2013).

Monitoreo Organizacional

Un creciente número de organizaciones humanitarias a lo largo del sector, además de los programas individuales de monitoreo, están tratando de monitorear su desempeño general como actores actores humanitarios, agregando información a nivel de país o a nivel de programa para brindar un panorama global. La naturaleza de lo que se monitorea difiere de una organización a otra: Mercy Corps considera el desempeño en contraste con la misión organizacional (Willet, 2013); Acción contra el Hambre (ACH) en contraste con el criterio de evaluación de la OCDE DAC (Guerrero et al., 2013); mientras que los Servicios de Auxilio Católico (CRS) se enfocan principalmente en las medidas de rendimiento. (Carr, 2013). Estas diferencias se reflejan en la naturaleza de los datos recopilados, y el grado en que se incluyen los datos adicionales (además de aquellos proporcionados por el monitoreo

del programa y los sistemas de evaluación). Mientras Mercy Corps ha sido muy clara en que su sistema 'Mission Metrics' utilizará solo datos de los sistemas de monitoreo existentes, Oxfam GB añade informes de programas con entrevistas del personal del programa. (Walden, 2013).

Estas iniciativas se enfrentan a algunos desafíos probatorios comunes. Dada la gran diversidad de contextos en los que una organización puede trabajar y los muy diferentes tipos de programas que puede emprender, puede ser extremadamente difícil establecer categorías comunes para la medición de todos los programas (Carr, 2013). De hecho, incluso el término "beneficiario" puede tener distintos significados en diferentes contextos. Cuando se recopila la información de un sub-grupo de programas, también pueden haber problemas con la *representatividad*: ¿hasta qué grado aquellos programas representan el trabajo de la organización en conjunto?

Sin embargo, quizás el principal desafío al igual que en otras áreas, es la *claridad en torno al contexto y los métodos*: ¿Cómo se prevé el uso de la información y por quiénes?. Como ha sugerido un experto: "Cómo lo estamos haciendo no es una buena pregunta de investigación" (Willet, 2013). Los organismos generan grandes cantidades de información, y su análisis requiere de recursos que se podrían utilizar en otros lugares. El monitoreo del desempeño organizacional se puede utilizar para el desarrollo de la estrategia; para la defensa; para informar a los donantes; y – cuando la información se despliega a nivel del país – para gestionar el desempeño. Dadas las implicaciones de recursos de estos enfoques, es importante fortalecer los sistemas de monitoreo organizacionales para responder a preguntas claras y explícitas.

El monitoreo de desempeño en toda la organización, si es bien realizado, tiene el potencial de respaldar la mejora en las organizaciones individuales. Sin embargo, los incrementos en el tamaño, el alcance y la complejidad de las respuestas humanitarias en la última década han llevado a un aumento de interacción e interdependencia entre las organizaciones humanitarias, ya que intentan hacer frente a situaciones que requieren de respuestas a gran escala, de múltiples fases y multisectoriales. En este contexto, el monitoreo y la mejora en el desempeño de organizaciones individuales es un paso importante en la mejora de la respuesta humanitaria. Sin embargo, idealmente, también podríamos monitorear el desempeño del sistema de ayuda humanitaria en su conjunto.

Monitoreo del sistema

Existe una variedad de marcos y sistemas para monitorear el desempeño humano. Estos tienden a concentrarse en el desempeño de ciertos grupos de actores dentro de un sistema humanitario más amplio, como la iniciativa de la Buena Donación Humanitaria, que se concentra en el desempeño de los donantes frente a principios específicos (véase DARA, 2011); en el desempeño relacionado con elementos específicos de la respuesta humanitaria, tales como la Asociación Internacional de Responsabilidad Humanitaria, que se concentra en la rendición de cuentas a las personas afectadas (véase HAP, 2010); o en el desempeño relacionado con respuestas específicas, como la respuesta internacional al tsunami del Océano Índico (Cosgrave, 2007).

Sin embargo, hasta hace poco no ha habido "ni siquiera un intento por evaluar el desempeño del sector humanitario en su conjunto, ya sea de forma única o regular" (Ramalingam et al. 2009: 73). No existe un marco común acordado para



Puede ser extremadamente difícil establecer categorías comunes para la medición de todos los programas.



monitorear el desempeño humanitario internacional, y si bien existen (como hemos visto) un gran número de mecanismos de monitoreo y de presentación de informes específicos por organismo, estos están fragmentados y recogen una amplia variedad de diferentes tipos de información por diferentes motivos.

En el 2009, ALNAP como parte del Proyecto de Desempeño Humanitario de la Red, sugirió un marco "balanceado, exhaustivo y coherente para el desempeño humanitario que pueda ser debatido y acordado entre los organismos" (Ramalingam et al., 2009: 7) y propuso un "cuadro de mando integral" humanitario, como la base del monitoreo en todo el sistema. Hasta la fecha, ha habido poco progreso en la idea de un marco de desempeño común, a pesar de los recientes debates acerca de la eficacia humanitaria, relacionados con la Cumbre Global Humanitaria en el 2016 se ha visto un renovado interés en el tema (OCHA, 2013). Mientras tanto, ALNAP ha impulsado un proceso de monitoreo del desempeño del sistema humanitario en contraste con los criterios de OECD DAC¹⁰ con un informe bienal del *Estado del Sistema Humanitario* (ALNAP, 2010 y 2012). Este proceso tiene como objetivo describir el tamaño, la forma y el desempeño del sistema humanitario formal; identificar las tendencias en el desempeño; y, quizás lo más importante, aclarar los problemas claves y los vacíos que deben ser llenados por los esfuerzos de mejora de todo el sector.

Aunque estos informes han sido muy bien recibidos (John Mitchell, comunicación personal, junio de 2013; Registro de Evaluación y Monitoreo ALNAP); el proceso también ha arrojado una serie de desafíos vinculados a la recopilación y el análisis de la evidencia (particularmente alrededor de las propuestas de Tipo B: "qué funciona") relacionados a las múltiples organismos y organizaciones en múltiples contextos.

El primer desafío señala dificultades observadas previamente en las secciones sobre alerta temprana y evaluación: esto se relaciona con la relevancia de los indicadores de medición del informe. Los criterios de OCDE DAC utilizados en el informe, como la "adecuación" y "coherencia" son categorías conceptuales con un componente subjetivo, que en la mayoría de los casos, solo se puede medir mediante el uso de una variedad de diferentes tipos de datos. Para hacer las cosas más difíciles, no existe un acuerdo general sobre qué grupos de indicadores se pueden o se deben usar para medir cada criterio (y, como se mencionó, puede que no sea posible establecer un único grupo de indicadores acordados, teniendo en cuenta la variedad en los contextos en los que trabajan los agentes humanitarios).

Como resultado, la selección de indicadores para el informe depende bastante de la opinión de expertos. ALNAP utiliza un enfoque de tres etapas en un intento por disminuir cualquier sesgo en el análisis. El primer informe preliminar fue recopilado por investigadores expertos con un buen conocimiento del sistema. Este informe preliminar fue revisado de manera crítica por la Secretaría General de ALNAP y luego fue revisado por separado por un Grupo Asesor Independiente, compuesto por personas con una variedad de perspectivas dentro y fuera de la Red ALNAP. La diversidad de los antecedentes representados por el Grupo Asesor también ayuda a asegurar que dichos temas que son "difíciles de medir" no se ignoren en el informe final (Knox Clarke, 2013a).

10 Para más detalles sobre cómo los criterios de Evaluación de OCDE DAC se aplican a la acción humanitaria, véase Beck (2006).

El segundo desafío – en relación con la exactitud de la información y análisis – es la gran falta de datos con la que estamos familiarizados. Para diversas zonas que son monitoreadas, existen muy pocos datos convincentes disponibles. Por ejemplo, en la área de cobertura "el sistema carece de un grupo de cifras concluyente sobre el total de personas que necesitan ayuda humanitaria en cualquier momento dado" (ALNAP, 2012:42). Cuando existen los datos, estos se generan normalmente por actores dentro del mismo sistema y por eso son potencialmente propensos al sesgo. En gran medida, este sesgo puede ser abordado por la triangulación, pero es importante triangular con la información que proviene de los actores fuera del "sistema formal", para desafiar cualquier sesgo "sistémico" o "cultural" que puede existir dentro de un sistema como en su conjunto. En este sentido, la inclusión de las encuestas de los beneficiarios y las entrevistas con el personal de las Autoridades Nacionales de Gestión de Desastres, (NDMAs) son elementos importantes de la metodología de investigación.

Como se mencionó anteriormente, cualquier medición del desempeño global en el sector humanitario – ya sea a nivel de un solo organismo o a nivel de todo el sistema – también enfrentará dificultades en torno a la *representatividad*: ¿en qué medida la muestra de la información refleja con exactitud el panorama completo? En iteraciones anteriores del *Estado del sistema humanitario*, las revisiones de las evaluaciones y las entrevistas de los informantes claves se han basado en muestras intencionales, mientras que las visitas de terreno (para las "verificaciones de datos") y las encuestas de los beneficiarios se han basado en muestras de conveniencia. En la siguiente fase del proyecto, las visitas y las encuestas serán seleccionadas utilizando un enfoque intencional, en un intento por mejorar el grado en que la información obtenida de estas fuentes pueda representar con mayor precisión la acción humanitaria en su conjunto (John Mitchell, comunicación personal, julio de 2013).

EN RESUMEN:

- El monitoreo comprende una variedad de actividades relacionadas, pero separadas: monitoreo de situación; monitoreo de programa; monitoreo organizacional y monitoreo de sistema.
- El sistema humanitario es débil en el monitoreo de situación: han existido pocos intentos exitosos para monitorear las situaciones por períodos mayores a un año o dos.
- El monitoreo de desempeño se puede enfocar mucho en resultados, y no lo suficiente en los resultados y en el impacto, además tiene problemas con la exactitud, particularmente en situaciones donde el acceso es un desafío.
- Los organismos han encontrado una variedad de formas innovadoras para abordar estos problemas, muchas organizaciones en la Red ALNAP están intentando monitorear su desarrollo humanitario a nivel mundial. Particularmente, el monitoreo organizacional es difícil cuando se usa información existente, pero la claridad sobre qué se está monitoreando, y cómo se usará la información, es importante para el éxito.
- ALNAP monitorea el desempeño de todo el sistema humanitario en el informe bial *Estado del sistema humanitario*, pero esto es un desafío, en parte porque mucha de la información necesaria no se ha recopilado, y en parte debido a las dificultades para generalizar desde un grupo muy diverso de situaciones.

3.5 Evidencia de las evaluaciones y ensayos controlados

Las evaluaciones de la acción humanitaria se enfrentan a desafíos similares para el monitoreo y las evaluaciones de necesidades. El lector ya debe estar a estas alturas familiarizado con dichos desafíos: las evaluaciones se llevan a cabo con frecuencia en ambientes carentes de datos, politizados y complejos, donde el acceso físico es limitado, las poblaciones son móviles y existen diferentes actores, quienes pretenden reafirmar su punto de vista sobre lo sucedido. Estos contextos dificultan la creación de una base de evidencias exacta, representativa de la experiencia de la población afectada y no sesgada por una interpretación subjetiva.

En general, la guía para la evaluación de la acción humanitaria sugiere que las evaluaciones tienen mayores probabilidades de generar evidencias contundentes cuando emplean enfoques de "métodos combinados". Por ejemplo, las directrices de evaluación y el monitoreo de la FICR sugieren que los datos cualitativos permiten solo una *generalización* limitada y pueden percibirse como datos de baja credibilidad, mientras que los métodos cuantitativos pueden ser costosos y "excluyen explicaciones y opiniones sobre el porqué de alguna situación". Por lo tanto, "se suele recomendar un enfoque de métodos combinados que puede emplear las ventajas de ambos" (FICR 2011: 35). Asimismo, la unidad de evaluación de MSF sugiere que "una combinación de métodos cualitativos y cuantitativos permite obtener por lo general los mejores resultados" (MSF 2012: 7), mientras que las directrices del Programa Mundial de Alimentos (PMA) sugieren que "se deben emplear tanto los datos cualitativos como cuantitativos ya que se complementan entre sí" (PMA, n.d.: 23).

“
Las evaluaciones
tienen mayores
probabilidades de
generar evidencias
contundentes
cuando emplean
enfoques de "métodos
combinados."
”

Con respecto a estas directrices, el PMA y la Oficina del Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) culminaron recientemente una serie de evaluaciones sobre el impacto del método combinado en la contribución de la asistencia alimentaria para soluciones duraderas en situaciones prolongadas de refugiados (Steen et al., 2013). Dichas evaluaciones emplearon datos cuantitativos de encuestas en hogares,¹¹ basadas en muestras aleatorias de la población de refugiados, y datos cualitativos de entrevistas, grupos de discusión y observación. El informe de síntesis indica que "el principal garante de la validez de los hallazgos se encuentra en la gama muy amplia de fuentes y métodos empleados para reunir la evidencia y su triangulación (Steen et al., 2013: 3). Se ha empleado también el muestreo probabilístico – que mejora la *representatividad* de los hallazgos – para evaluaciones en contextos pastoriles que contaban con listas disponibles de familias (Abebe et al., 2008).

Sin embargo, la evaluación humanitaria por lo general "emplea principalmente métodos cualitativos" (Buchanan-Smith y Cosgrave, 2012). Una revisión informal realizada para el informe de evaluaciones comprendidas entre el 2010 y 2012 en la Biblioteca de Recursos de ALNAP sugiere que el uso de enfoques de métodos combinados es poco común, y que la mayoría de evaluaciones se ha basado en enfoques ampliamente cualitativos para la generación de evidencias, y sobre

11 Medición de seguridad alimentaria mediante indicadores estándar del Puntaje de Consumo de Alimentos (PCA), Puntaje de Diversidad Dietética en el Hogar (HDDS) e Índice de Estrategia de Supervivencia para enfrentar la Crisis (CSI).

todo en entrevistas – a menudo con informantes claves – así como mediante la observación personal.

La mayoría de evaluaciones emplean técnicas de muestreo intencional para identificar a los entrevistados, basadas en la triangulación de fuentes (y, hasta cierto punto, en la triangulación de las observaciones de diferentes evaluadores) para establecer la *exactitud*. La orientación hacia los enfoques cualitativos y discursivos se observa particularmente en las evaluaciones cuyo fin principal es el aprendizaje (en vez de la rendición de cuentas), ya que este tipo de evaluación enfatiza la importancia de la experiencia subjetiva y la participación de las partes interesadas claves en el proceso de evaluación como una precondition para el aprendizaje y cambio.

Aunque se tiende algunas veces a considerar que los enfoques cuantitativos generan evidencias “concluyentes” y que los enfoques cualitativos generan evidencias “cuestionables”, el uso de enfoques cualitativos no debe ser una excusa para la falta de rigor. Los seis criterios descritos en la Sección 2.3 para la evidencia contundente son tan importantes para la información reunida mediante entrevistas como para los datos de encuestas aleatorias. Por desgracia, muchas evaluaciones humanitarias no emplean métodos cualitativos reconocidos académicamente, ni cumplen con las normas básicas de calidad relacionadas a la *exactitud*, *representatividad* y *relevancia*. Hace poco, una organización miembro de ALNAP revisó sus propias evaluaciones comparándolas con los principios de evidencia de las ONG de BOND (Bond, n.d.). Sus hallazgos sugieren que existe una posibilidad significativa de mejora:

De acuerdo con la herramienta BOND, la evidencia en las...evaluaciones arrojó un resultado de baja calidad en casi todas las categorías identificadas. Esto indicó que mientras la calidad de los proyectos que se están evaluando puede ser buena, la baja calidad de la evidencia dificultó concluir ello a partir de los informes. Se obtuvieron diversas conclusiones a partir de una aparente falta de evidencia y se presentaron varias anécdotas como ilustraciones de hallazgos más generales. Los informes contenían opiniones, aunque rara vez sirvieron como ejemplos o evidencias para corroborar algún dato. Asimismo, se observó con frecuencia una falta de información cuantitativa básica (ej. números de hombres/mujeres/adultos mayores/jóvenes) y no se presentó ningún método ni metodología analítica convincente. (Parker y Sanderson, 2013).

Estos hallazgos pueden no representar a todas las organizaciones humanitarias, aunque – nuevamente en base a una revisión informal de la Biblioteca de Recursos de ALNAP – tampoco parecen ser particularmente inusuales.

Debido a la importancia que conlleva asegurar la calidad probatoria de las evaluaciones, varias organizaciones han instaurado sistemas para revisar la calidad de la evaluación en los últimos años: el PMA tiene dicho sistema¹², y la UNOCHA

12 La Evaluación del Sistema de Aseguramiento de Calidad, o EQAS, que establece procesos con etapas incorporadas para el aseguramiento de calidad y modelos para los productos de evaluación. La Evaluación incluye también el aseguramiento de calidad de los informes de evaluación (informes iniciales, completos y de resumen) basados en listas de verificación estandarizadas...Este proceso de aseguramiento de calidad no interfiere con los puntos de vista e independencia del equipo de evaluación, aunque asegura que el informe proporcione la evidencia necesaria de una manera clara y convincente y determine sus conclusiones en dichas bases (FAO Y PMA, 2013).

planifica implementar uno (Jock Paul, comunicación personal, abril de 2013). Al mismo tiempo, ALNAP está probando actualmente una Guía para la Evaluación de la Acción Humanitaria, que presenta enfoques para llevar a cabo enfoques cualitativos más contundentes (Buchanan-Smith y Cosgrave, 2013).

“
La falta de una perspectiva del beneficiario constituye otra área de particular interés en muchas evaluaciones.”

La falta de una perspectiva del beneficiario constituye otra área de particular interés en muchas evaluaciones, a pesar de la gran confianza en las entrevistas como una fuente de información evaluadora. Cuando Beck y Buchanan-Smith realizaron meta evaluaciones para ALNAP, encontraron que casi tres cuartas partes de las evaluaciones revisadas entre 2001 y 2004 no habían consultado a los beneficiarios o habían incluido solo consultas mínimas (Beck y Buchanan-Smith, 2008).

A pesar de algunas excepciones destacadas, las evaluaciones aún tienden a subestimar la experiencia de las poblaciones afectadas como una fuente de evidencia: la edición del año 2012 de *State of the Humanitarian System (El estado del sistema humanitario)* (ALNAP, 2012) identificó la consulta a los destinatarios como una de las áreas más débiles del desarrollo humanitario, mientras que la Asociación Internacional para la Rendición de Cuentas en el Ámbito Humanitario (HAP) en el *Humanitarian Accountability Report (Informe de Rendición de Cuentas)* de 2013 indica que la participación del beneficiario es “sumamente inusual en la evaluación, incluso si la tendencia actual es impulsar la participación del beneficiario en esta etapa” (Alexander et al., 2013: 29). Los enfoques de impacto participativo analizados más adelante brindan un ejemplo interesante, incluso diferente, sobre cómo se puede elaborar la evaluación en torno a la experiencia de personas afectadas por los desastres.

Causalidad y transferibilidad – Dos desafíos particulares para las evaluaciones

Además de los otros ensayos probatorios indicados en la Sección 2, los evaluadores necesitan abordar el desafío de atribución. Para una evaluación, no es suficiente representar una situación con exactitud; es necesario también mostrar la relación entre la intervención específica, o series de intervenciones, y la situación descrita (es decir, propuestas de Tipo B). Como resultado, las evaluaciones deben ser rigurosas en su enfoque a la *atribución*.

Muchas evaluaciones de acción humanitaria abordan este desafío de causalidad en base a un marco lógico. En este enfoque, se diseña un proyecto según una cadena causal, que establece un tipo de hipótesis. Si se ofrecen algunos productos (una cierta cantidad de pozos de perforación que producen una cantidad y calidad específica de agua) y algunas suposiciones son ciertas (las personas obtienen agua potable de esta fuente y no en otro lugar), entonces la suposición “lógica” es que se obtendrán algunos resultados positivos, como una disminución de enfermedades de origen hídrico. En este caso, si se demuestra en una evaluación que se ofrecieron productos, que los resultados se produjeron posteriormente y además si los entrevistados establecen una relación entre los productos y los resultados que desestiman las explicaciones alternativas; se deduce entonces que el producto provoca por lo general el resultado.¹³

13 El enfoque tiene repercusiones interesantes del racionalismo, en vez del empírico, entendimiento de la verdad, dependiendo como ya lo hace en una secuencia lógica, en vez de una observación directa de la relación causal.

Este enfoque, quizás imperfecto, se ha inclinado al control de cualquier evaluación que se centre en el nivel de resultados. Sin embargo, resulta menos provechoso cuando la evaluación considera el impacto de la intervención humanitaria, ya que la cadena causal entre el producto y el impacto¹⁴ suele ser más compleja y ambigua, además puede no haberse establecido previamente

La medición del impacto de la acción humanitaria es desafiante por varias razones: falta de capacidad; elevada rotación de personal; rechazo al fracaso publicitario; así como dificultades técnicas para establecer líneas de base y grupos de control, y dilucidación entre el impacto de una sola intervención y el impacto más amplio de una operación. Todo esto actúa en contra de la evaluación del impacto contundente, y por ende, las evaluaciones del impacto siguen siendo bastante inusuales en el sector humanitario (Proudlock y Ramalingam, 2009). Dependiendo de dónde sucedan, el significado de “impacto” con frecuencia difiere de un organismo a otro (ALNAP, 2012; ACH, Mimeo). No obstante, se ha producido un reciente e importante mayor interés en la evaluación del impacto humanitario, provocado al menos de forma parcial por las ganas de determinar los enfoques que ofrecen el mejor uso del dinero en un ambiente financiero limitado.

Fuera del sector humanitario, existe un debate constante sobre los enfoques más válidos para establecer evidencias contundentes de *atribución*. Los defensores de enfoques experimentales – por lo general ensayos controlados aleatorios (RCT) – sostienen que ellos representan la “norma de oro” para establecer la causalidad; o, con mayor modestia, “la peor forma de diseño, salvo todos los demás que se han intentado” (Bickman y Reich, 2009). Las críticas apuntan al costo de los RCT – cuyo monto generalmente oscila entre \$200,000 a \$900,000 (Banco Mundial n.d, citado en Bradt, 2009) – indicando que los enfoques no experimentales y menos costosos se emplean con frecuencia en una variedad de disciplinas científicas para establecer la causalidad más allá de la duda razonable (Scriven, 2009). Ellos sugieren que los resultados de los RCT son realmente efectivos solo cuando las intervenciones son “estables y relativamente simples, además...producen efectos relativamente rápidos y grandes en relación a otras influencias potenciales” (Piccioto, 2012). Asimismo, las críticas sostienen que los RCT son de uso limitado en la elaboración de políticas ya que no se pueden aplicar a otros niveles (Schwandt, 2009) y porque rara vez explican cómo una intervención conlleva a impactos específicos (Piccioto, 2012).

Es tal vez posible concluir que los RCT son definitivamente útiles cuando responden la pregunta clínica en formato específico “PICO” para la cual han sido diseñados (Bradt 2009).¹⁵ Fuera del ámbito clínico, los RCT suelen ser más efectivos para establecer la *atribución* “donde la cadena causal entre el agente y el resultado es bastante corta y simple, y donde los resultados se pueden llevar con seguridad a otros niveles” (Victora et al., 2004, citado en Dijkzeul et al., 2013).

Se han producido algunos intentos para emplear los RCT en el ámbito humanitario con el fin de obtener evidencia del impacto. El DFID financió un RCT en Malawi para

14 Entendido con frecuencia como la relación final de una cadena causal: una intervención de desarrollo, directa o indirectamente, planificada o accidental (OCDE-DAC, 2012).

15 PICO representa a los pacientes (o población) que reciben intervención; la intervención en estudio; la comparación (la intervención alternativa que se considera); los resultados clínicos que se buscan. La consideración de estos cuatro factores conlleva a la creación de una pregunta “comprobable”. Las preguntas PICO son por lo general de particular relevancia para temas como determinar cuál será la terapia más efectiva

analizar diferentes composiciones de alimentos instantáneos en el tratamiento de la desnutrición severa aguda (Kerak et al., 2009 en Buchanan-Smith y Cosgrave, 2012), mientras que ACH realizó unos RCT sobre el impacto de alimentos instantáneos empleados en programas de alimentación suplementaria para la desnutrición infantil (Huybregts et al., 2012). En Liberia, el IRC ha llevado a cabo una evaluación del impacto de un programa de reconstrucción dirigido por la comunidad mediante un diseño experimental (Fearon et al., 2008): un ejemplo interesante, y bastante inusual a la fecha, es el uso de enfoques experimentales en el trabajo humanitario fuera de los sectores de salud y nutrición.



Se espera un aumento del número de ensayos controlados, revisiones sistemáticas y otros enfoques que priorizan las metodologías experimentales, aunque se producirán muchas situaciones donde los enfoques experimentales no serán convenientes o factibles.



Se espera un aumento del número de ensayos controlados, revisiones sistemáticas y otros enfoques que priorizan las metodologías experimentales, apoyados por organizaciones tales como 3IE y EvidenceAid, que intentan incrementar el rigor y complejidad de la generación de evidencia en el sector humanitario. A medida que este trabajo progresa, se contará con mayor información sobre intentos iniciales. La experiencia de ACH en Chad confirma que los RCT en contextos humanitarios requieren un largo período para su diseño e implementación, y son vulnerables a cambios en situaciones donde participan los grupos. Su éxito depende de los investigadores altamente capacitados (en este caso, mediante una asociación con la Universidad de Gante y el Instituto de Medicina Tropical, Amberes) y de una comunicación efectiva y frecuente con los participantes, la población en general y el personal de la organización en el terreno (Puett y Salpeteur, 2013). Debido a los altos costos que implican, es importante seleccionar temas lo suficientemente estratégicos: en efecto, donde el gasto se compensa con el costo de no realizar la investigación (Chloe Puett, comunicación personal, abril de 2013).

Al mismo tiempo, se producirán muchas situaciones donde los enfoques experimentales no serán convenientes o factibles. Ambos, Stern en un reciente informe para DFID (Stern, 2012) y Rogers en un informe para Interaction (Rogers, 2012) sugieren diseños y métodos alternativos, incluyendo enfoques “cuasi experimentales”, enfoques basados en casos, y enfoques basados en teoría. Ambos apuntan a la conveniencia del uso de una variedad de enfoques para considerar la causalidad e indicar la importancia, en el contexto humanitario, de considerar el grado hasta el cuál las intervenciones contribuyen a cambios, en vez de intentar atribuir un cambio solamente a cualquier intervención única.

Los profesionales humanitarios están experimentando cada vez más con una variedad de métodos que abordan el tema de la *atribución* (Recuadro 5). Las evaluaciones de impacto del PMA/ACNUR, mencionadas anteriormente, emplearon grupos de comparación “oportunistas”, como los grupos oficiales y no oficiales de refugiados, o campamentos que recibieron diferentes tipos de ayuda. Ante la falta de un marco formal lógico para el programa, los evaluadores también reconstruyeron la lógica de las intervenciones, para determinar si el estímulo (ayuda alimentaria) conllevó en realidad al impacto de la forma originalmente prevista.¹⁶ El proyecto ECB ha desarrollado y analizado una metodología que promueve el cambio de la contribución de intervenciones humanitarias, empleando la estadística descriptiva combinada con datos de la entrevista (Few et al., 2013).

16 Curiosamente, la evaluación pudo concluir que “no se mantuvieron importantes suposiciones con respecto a la ayuda humanitaria para los refugiados” (Steen et al., 2013: 14).

Como se indica anteriormente, la evaluación del impacto es una de las pocas áreas donde se han realizado intentos organizados para desarrollar una evaluación en torno a las percepciones de personas afectadas. El Instituto Feinstein de la Universidad de Tufts ha establecido pautas sobre la evaluación del impacto participativo: estas apuntan a dificultades prácticas y éticas para establecer grupos de control que examinen la *atribución*, y en cambio, se centran en el uso de herramientas participativas para evaluar la contribución relativa del cambio de los factores relacionados y no relacionados al proyecto (Catley et al., 2008).

Sin embargo, el enfoque participativo no solo es una forma de evitar dificultades prácticas. Las evaluaciones participativas reconocen que el impacto – siendo en gran parte sobre el más largo plazo y consecuencias frecuentemente inesperadas de una intervención – no encaja claramente en la lógica del proyecto generado externamente, y por lo general, se entienden mejor a partir de la experiencia de aquellos que reciben la ayuda.

Mientras “los sistemas tradicionales M y E suelen poner demasiado énfasis sobre “nuestros indicadores” y no “sus indicadores”” (Catley et al., n.d.: 21), el enfoque participativo emplea indicadores de impacto identificados por los beneficiarios previstos. Asimismo, el Método de Primer Impacto en la Población (PFIM) solicita a las comunidades identificar los cambios más importantes en sus vidas y las causas de estos cambios (véase, por ejemplo, O’Hagen y McCarthy, 2012). En Somalia, la FAO empleó un enfoque similar: cuya evaluación de impacto se basó en un enfoque iterativo y descriptivo: las personas que recibieron la ayuda compartieron sus experiencias del programa, y los impactos que tuvieron, en las entrevistas semiestructuradas que se reforzaron por una variedad de herramientas participativas. Luego, los temas emergentes retroalimentaron a los grupos de la comunidad para un debate adicional y verificación posterior (Tessitore, 2013).

Recuadro 5. Diseños de investigación para investigar la atribución

Diseño experimental:

En un diseño experimental, se asignan aleatoriamente a los participantes, ya sea a un grupo que recibe servicios del programa o a un grupo de control que no los recibe, y que sirve como una “hipótesis de contraste”. Luego se comparan los resultados de estos dos grupos. El diseño del experimento permite que cualquier diferencia en el resultado entre el grupo receptor y de control se atribuya a los servicios recibidos.

Diseño cuasi experimental:

Los estudios cuasi experimentales pretenden también demostrar la atribución mediante la comparación de resultados, aunque no consideran la designación aleatoria de los participantes a los grupos. En vez de ello, comparan los resultados entre grupos que recibieron servicios y grupos similares que no los recibieron (un “experimento natural”); o un grupo antes y después de una intervención.

Enfoques basados en teoría:

Estos enfoques no pretenden demostrar la atribución mediante comparaciones de grupos receptores y de hipótesis de contraste. En vez de ello, analizan la teoría subyacente de causalidad mediante la cual, los diseñadores del programa esperan que algunas actividades conlleven a ciertos resultados. En un enfoque basado en teoría, las series de suposiciones en el diseño del programa que relacionan el estímulo, contexto y resultado se tratan como hipótesis, que pueden ser puestas a prueba mediante una variedad de métodos cuantitativos y cualitativos.

Enfoques basados en casos:

Estos enfoques se basan en un estudio sobre lo que sucedió realmente en casos específicos: identificar los factores que conllevaron a algunos resultados y luego comparar dentro de los casos, o entre casos, con el fin de realizar “generalizaciones analíticas”.

Fuentes: definiciones basadas en Stern et al. (2012); Morra Imas y Rist (2009); Leeuw (2012).

EN RESUMEN:

- Si bien muchas de las pautas para la evaluación humanitaria apuntan hacia los beneficios del uso de enfoques de métodos combinados, la mayoría de las evaluaciones se realizan solo con métodos cualitativos.
- Las evaluaciones humanitarias emplean por lo general muestreos intencionales para tratar de asegurar su *representatividad* y triangulación para aumentar la *exactitud*.
- Sin embargo, muchas evaluaciones humanitarias no hacen un buen uso de estos métodos cualitativos y, por ende, la calidad de la evidencia que presentan es con frecuencia deficiente.
- Muchas evaluaciones no incluyen de ninguna forma significativa las percepciones de personas afectadas por las crisis.
- Algunos de los trabajos más efectivos que se esfuerzan por estructurar la evaluación en torno a las percepciones de personas afectadas se han observado en la área de evaluación de impacto.
- La *atribución*/contribución son desafíos particulares para las evaluaciones, que pretenden investigar las relaciones causales entre un programa y una situación posterior.
- Se ha producido recientemente cierto aumento en los enfoques “experimentales” para establecer la *atribución*, y los organismos están también empezando a considerar una variedad de otros enfoques no experimentales para la evaluación de la contribución.



¿Se usa actualmente la evidencia para guiar operaciones humanitarias y elaboración de políticas?

4.1 La calidad no garantiza el uso

La sección anterior considera la calidad de la evidencia generada dentro del sistema humanitario internacional. Si se dan a conocer las decisiones mediante la evidencia, es, por supuesto, fundamental que dicha evidencia sea de la mayor calidad posible. Sin embargo, la calidad sola no garantiza el uso de la evidencia. Los procesos para toma de decisiones son poco frecuentes, si no inexistentes, y completamente racionales: son influenciados por una variedad de consideraciones políticas y organizacionales, y los sesgos de los responsables de la toma de decisiones.

Si se empleará la evidencia para informar sobre la acción humanitaria, se debe entonces abordar no solo el proceso de recopilación de información y análisis, sino también el proceso de toma de decisiones. La experiencia de la profesión médica es instructiva en este aspecto: el avance hacia “la medicina basada en evidencia” que se llevó a cabo en los años 90 se produjo en gran parte por los cambios en los procedimientos organizacionales y diferentes actitudes entre los distintos médicos (Anthony Redmond, comunicación personal, octubre de 2012): “la toma de decisiones basadas en la evidencia no requiere por lo general datos científicos adicionales en sí mismos, sino más bien...recursos financieros y voluntad política” (Bradt, 2009: 3).

Teniendo en cuenta esta idea, la presente sección evalúa hasta qué grado los responsables de toma de decisiones humanitarias han empleado los diferentes tipos de evidencia y conocimiento, e investiga los factores que parecen limitar, y apoyar, el uso de la evidencia en la toma de decisiones.

4.2 Uso de la evidencia de alerta temprana

Como se indica en la Sección 3, se ha producido un progreso considerable en la generación de una evidencia de alerta temprana oportuna y exacta durante las últimas dos décadas, facilitando la elaboración de propuestas de Tipo A con cierto grado de confianza en muchos contextos. Sin embargo, el uso de dicha información ha provocado mayores complicaciones, particularmente en relación a la hambruna e inseguridad alimentaria. El problema de una desconexión entre la información de alerta temprana y las respuestas reales han sido bastante reconocidas (véase Buchanan-Smith y Davies, 1995), sin embargo el problema persiste en el centro de la política y en la toma de decisiones, por ejemplo, las crisis alimentarias de evolución lenta (Levine et al., 2011).

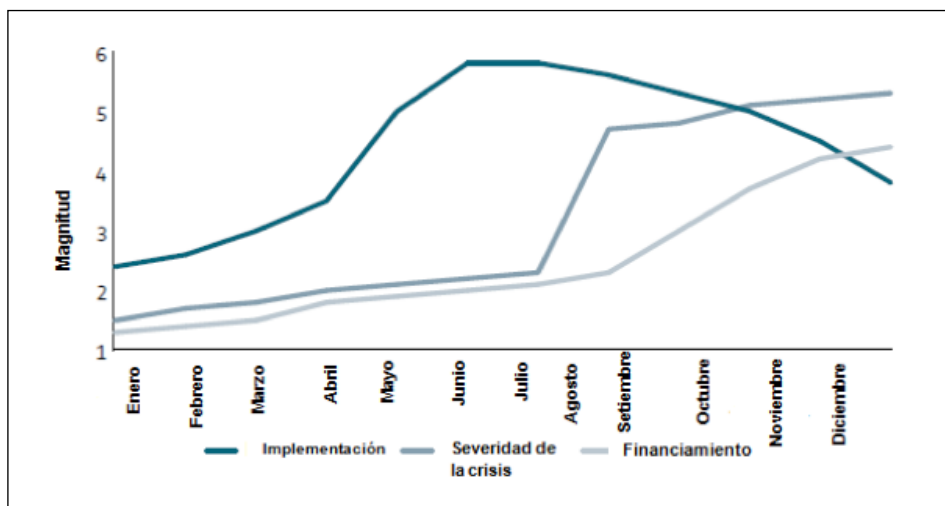


El problema de una desconexión entre la información de alerta temprana y las respuestas reales persiste en el centro de la política y la toma de decisiones.



Se pueden observar algunos de estos temas relacionados a la hambruna del año 2011 en la región central y meridional de Somalia. El Gráfico 1, obtenido de la Evaluación en Tiempo Real (RTE) de IASC sobre la respuesta ante la crisis en Somalia, establece la relación entre la severidad de la crisis y la respuesta de financiamiento e implementación de la respuesta.

Gráfico 1: Representación esquemática de la respuesta ante la hambruna en Somalia del año 2011.



Fuente: Evaluación en Tiempo Real de IASC sobre la respuesta internacional ante la crisis en Somalia del año 2011 (Valid International, 2012).

La característica más notable del Gráfico 1 son las brechas: primero, entre la severidad de la crisis¹⁷ (línea verde) y la disponibilidad de financiamiento (línea verde claro), y segundo, entre el financiamiento y la implementación de la respuesta (línea verde oscuro). Se contó con evidencia sólida de alerta temprana de la crisis inminente desde el último trimestre del 2010 en adelante (Red FEWS e IPC de la FAO). No obstante, un mayor aumento en el financiamiento se produjo solo con la declaración de hambruna en julio del 2011. Este fue el factor principal al que los autores de la evaluación en tiempo real se refieren como un “fracaso sistemático de respuesta temprana” que abarcó dos aspectos:

- La falta de acción preventiva para afrontar las causas próximas de vulnerabilidad mediante la intervención urgente de los medios de subsistencia que permitiría fortalecer la resistencia a corto plazo y reducir la necesidad de ayuda.
- La falta de ayuda temprana en gran proporción para afrontar los síntomas más graves de la crisis cuando se requirió dicha asistencia con mayor urgencia a inicios y mediados del 2011 (Valid International, 2012: 37).

Ante dicho caso, la planificación para una respuesta más amplia solo se llevó a cabo en julio del 2011, luego de la fase más crítica. Los autores de la RTE indicaron que: “Dado los plazos de entrega considerados, esto se debió dar a más tardar en enero/febrero si las acciones preventivas y de ayuda se hubieran entregado en el momento requerido”. Cuando las respuestas del programa empezaron a ser más amplias, la crisis ya había sobrepasado su punto máximo.

¹⁷ La “severidad de la crisis” como se representa en el documento, se basa en un conjunto de indicadores que incluyen los niveles de desnutrición aguda y los precios de alimentos en el mercado. Estas tres variables se reconocen como difíciles de cuantificar e inconmensurables, de modo que el Gráfico 1 es solo indicativo.

Durante los años 2010-2011, la situación en la región central y meridional de Somalia era complicada por los problemas políticos y de seguridad. Una buena parte de las áreas afectadas estaban en ese momento bajo el control del grupo Al Shabaab. Las principales organizaciones donantes se rehusaron a brindar apoyo a este grupo, mientras que el mismo Al Shabaab impidió que muchos organismos operaran en la región. Sin embargo, el patrón de alerta temprana y respuesta tardía se había producido antes, donde estas limitaciones no habían alcanzado el mismo grado (Hillier y Dempsey, 2012) y se repitió en otras partes del Cuerno de África. En otras palabras, las peculiares condiciones de Somalia no explican el fenómeno.

El problema no fue principalmente sobre la disponibilidad ni la calidad de la evidencia. Como se indica anteriormente, la evidencia de alerta temprana fue oportuna y se consideró, en general, de buena calidad, aunque la validez de algunos de los datos y análisis estuvo abierta al debate (Levine et al., 2011). "Por último, los sistemas de alerta temprana realizados, aunque los responsables de la toma de decisiones prefirieron mantenerse en reserva" (Hillier y Dempsey, 2012: 14).

Es posible que existan varias razones para el fracaso del sistema humanitario internacional en el uso de la evidencia de alerta temprana. Estas limitaciones se relacionan con la naturaleza de la evidencia disponible; los procesos para la toma de decisiones preponderantes en el sistema humanitario internacional; y los procesos operacionales más amplios del sistema en su conjunto.

La primera limitación – la naturaleza de evidencia disponible – se relaciona con el hecho que, incluso cuando las organizaciones sabían que la situación podría conllevar a impactos significativos y negativos en las vidas y medios de subsistencia de las personas en la región, no supieron qué hacer con esta información: cómo transformar dicho conocimiento en respuestas efectivas (Levine et al., 2011). Si bien existen ejemplos documentados de acción preventiva efectiva en situaciones similares (véase el debate sobre la intervención temprana en emergencias causadas por sequías en la Sección 4.4), la mayoría de las organizaciones parecen no haber sido conscientes de estos ejemplos, ni tener experiencia en su uso, habiéndose centrado, históricamente, en actuar luego de una crisis que ya es evidente. Como consecuencia, muchas organizaciones humanitarias tienen problemas para transformar la información de alerta temprana a programas concretos.

Dadas las propuestas humanitarias descritas anteriormente, este es un hallazgo importante: no fue posible emplear la evidencia en el Cuerno de África para una propuesta de "Tipo A" (donde existe un problema) sin tener evidencia para las propuestas de "Tipo B" (donde se requiere un enfoque efectivo para abordar el problema). La alerta temprana pudo haber sido más efectiva si hubiera estado acompañada de ejemplos de "buenas prácticas" programáticas de organismos con experiencia en esta área.

Sin embargo, esta no fue la única razón para la inactividad. La falta de acción para atender la crisis en Somalia también señala la segunda limitación: debilidades en la forma en que se toman las decisiones dentro del sistema humanitario. Hillier y Dempsey indican que "muchas personas 'en el terreno' dentro de la región... eran conscientes de la crisis inminente e intentaron levantar voces de alarma en enero y febrero de 2011, (aunque) no siempre pudieron llamar la atención "más allá de ellos mismos"" (Hillier y Dempsey, 2012: 4). La naturaleza mundial y atomizada del

sistema humanitario señala que las decisiones requieren con frecuencia el consenso de un gran número de actores en diferentes lugares y organizaciones; y no siempre es claro dónde y quiénes deben tomar dichas decisiones (Knox Clarke, 2013b). Además, las personas con diferentes cargos trabajan a menudo bajo distintos grupos de incentivos y se encargan de darle un valor diferente a la misma evidencia.

En el caso de Somalia, parece que (para simplificarlo) el personal en el terreno se comprometió a prestar atención a los datos de alerta temprana por la situación deteriorada que observaron en el terreno. Por otra parte, los miembros del HCT en Somalia, eran conscientes de las dificultades que implicaban la recaudación de fondos y estuvieron más reacios para comprometerse con la información (Slim, 2012). En Etiopía, algunos responsables de la toma de decisiones consideraron a las necesidades identificadas por los sistemas de alerta temprana como “insignificantes” para catalizar una respuesta efectiva de la organización donante (Majid, 2011). Mientras tanto, las organizaciones donantes en las capitales occidentales eran conscientes de la necesidad por hacer un mejor uso de los recursos de los contribuyentes y no estaban conformes con la liberación de fondos hasta que hubiera evidencia, – en forma de datos nutricionales – de que se estaba produciendo ya una crisis humanitaria ((Hillier y Dempsey, 2012; Levine et al., 2011). Al mismo tiempo, los gobiernos en la región se resistían a pedir ayuda internacional a menos que fuera absolutamente necesario, ya que el hacerlo significaba la aceptación de su fracaso.

En general, los actores tuvieron que decidir cuánta atención prestarían a evidencias específicas en una situación de incertidumbre, además estas decisiones estarían muy influenciadas por otras informaciones e incentivos organizacionales. Este es un problema común en las operaciones humanitarias – muchos autores han comentado sobre los altos niveles para evitar riesgos en el sistema humanitario (véase Buchanan-Smith y Scriven, 2011), con el cual se podría discrepar, suelen aislar el riesgo de las organizaciones humanitarias y colocarlo sobre las espaldas de las mismas personas afectadas por el desastre.

Durante los años 2010-2011, estos problemas de toma de decisiones estaban compuestos por una falta de “desencadenantes” (Levine et al., 2011): declaraciones claras sobre los niveles de necesidad – particularmente en sectores de agua y saneamiento – que podrían forzar automáticamente una decisión en la respuesta. En efecto, los sistemas de alerta temprana fueron parte de un amplio proceso de toma de decisiones: por desgracia, parecen haberse desarrollado significativamente mejor que otros elementos de dicho proceso. Se produjo cierta ambigüedad sobre temas claves, tales como quiénes deben tomar las decisiones; la importancia relativa de diferentes criterios para evaluar la evidencia en la toma de decisiones; e incluso cuándo y si se deben o no tomar dichas decisiones.

La tercera serie de limitaciones sobre el uso de evidencia de alerta temprana comprendió los procesos operacionales más amplios que transforman el financiamiento en acción. Por lo general, la liberación actual de fondos y el inicio de actividades se relaciona a la evaluación de necesidades, en vez de mecanismos de alerta temprana (Levine et al., 2011). Las evaluaciones de necesidades – que se concentran en necesidades inmediatas y reales en vez de necesidades anticipadas – suelen realizarse un tiempo después de las evaluaciones de alerta temprana, cuando la situación ya se ha deteriorado. Luego, se produce por lo general un retraso de varios meses antes de comenzar las operaciones en el terreno. Desde esta



Los altos niveles para evitar riesgos en el sistema humanitario suelen aislar el riesgo de las organizaciones humanitarias y colocarlo sobre las espaldas de las mismas personas afectas por el desastre.



perspectiva, no se emplean evidencias de alerta temprana para tomar decisiones sobre la intervención ya que no se realiza una designación formal para dirigir dichas decisiones.

Para remediar esto, Levine y otros autores proponen una serie de acciones de preparación concretas, en áreas como reclutamiento y adquisición, que los organismos podrían tomar como base de evidencia de alerta temprana para reducir su tiempo de respuesta. Asimismo, Hillier y Dempsey recomiendan una serie de acciones “libres de excusas” que se deben desencadenar por alertas tempranas, incluyendo el aumento de capacidades y la predisposición de reservas.

EN RESUMEN:

- A pesar de las mejoras en la calidad de los sistemas de alerta temprana, la evidencia que señala la probabilidad de una crisis no puede con frecuencia ser plasmada en acciones, como se demostró en la imposibilidad para responder a dicha evidencia en el Cuerno de África y en particular en la región central y meridional de Somalia durante el 2010 y 2011.
- El fracaso en el uso de la evidencia de alerta temprana se produce por diversas causas, aunque un factor clave radica en que la generación de evidencia de alerta temprana no forma parte de un proceso estructurado amplio de alerta temprana y respuesta:
- El “proceso comercial” actual de respuesta humanitaria relaciona la respuesta a las actividades de evaluación, mas no a la alerta temprana; mientras que el rol de la evidencia de alerta temprana para generar una respuesta no es aún claro.
- Resulta por lo general difícil determinar a los responsables de la toma de decisiones relacionadas a las respuestas tempranas, y por lo tanto, nadie tiene asignada una responsabilidad establecida para estas decisiones.
- Resulta por lo general difícil determinar las acciones que se deben tomar en base a la alerta temprana: las organizaciones humanitarias cuentan con mayor experiencia en cómo reaccionar ante las crisis, más que en anticiparlas.
- Además de estos fracasos estructurales y de procesos, los incentivos organizacionales (y en particular, el interés por evitar el riesgo de “desperdiciar dinero”) evitan cualquier acción que se pueda realizar antes de producirse una crisis.

4.3 Uso de la evaluación y monitoreo de la evidencia

Si bien es cierto que el lugar de alerta temprana en el proceso de toma de decisiones es a menudo confuso, no existe tal confusión sobre la función de las evaluaciones de necesidades que se designan específicamente para proporcionar evidencia en el uso de la toma de decisiones. Por lo general, esta evidencia se enfoca en las propuestas de Tipo A relacionadas con la existencia de necesidades y permite que las decisiones se realicen si se requiere asistencia; cuánto¹⁸; para quiénes; y por cuánto tiempo. La evaluación es con frecuencia considerada por los menos por dos grupos de responsables en la toma de decisiones: el personal del organismo que la

¹⁸ Aunque como se ha visto, una falla común de las evaluaciones es asumir que se requiere asistencia y enfocarse más en cuánta asistencia se debe proporcionar.

emplea para el diseño inicial de un programa y requisitos de financiamiento; y los donantes, que la emplean para determinar si los requisitos de financiamiento son válidos. Además, los profesionales humanitarios a menudo esperan hacer uso de la información obtenida de las evaluaciones de necesidades como evidencia para respaldar las decisiones en torno a las propuestas de Tipo B: el mejor tipo de ayuda que se pueda brindar en dicho contexto específico. Sin embargo, el último uso, que se detallará más adelante, conlleva a complicaciones.

Una revisión de la literatura sugiere que, cuando surgen decisiones relacionadas a las propuestas de Tipo A (ya sea para intervenir, dónde y en qué medida), el uso de las evaluaciones de necesidades no es consistente, y dónde se emplean las evaluaciones, la evidencia proporcionada comprende con frecuencia solo uno de varios factores que repercuten en las decisiones. Al principio, esto puede parecer sorprendente – las evaluaciones están diseñadas, después de todo, principalmente como herramientas para la toma de decisiones – aunque se explica por una serie de limitaciones relacionadas a las mismas evaluaciones y a la naturaleza de decisiones humanitarias.

De acuerdo con lo observado anteriormente, la calidad probatoria de las evaluaciones de necesidades humanitarias es desigual, y – ya que las evaluaciones casi nunca son realizadas por actores imparciales, aunque sí por los mismos organismos que las emplearán para justificar los requisitos de financiamiento – son a menudo tratadas con cierto escepticismo por los donantes (Bradt, 2009; Darcy et al., 2013). De acuerdo con esta perspectiva, el abordar los problemas con la calidad probatoria descrito en la Sección 3.1 (más arriba), mejoraría definitivamente el uso de las evaluaciones en la toma de decisiones.

Una segunda limitación inherente al proceso de evaluación (en vez de los factores que en gran parte superan el control del asesor) radica en la puntualidad. Para que las evaluaciones influyan la toma de decisiones, la información tiene que ser accesible cuando se tome la decisión.

Al revisar la aceptación de una evaluación inicial rápida de necesidades por ACAPS posterior al terremoto de Haití del 2010, Darcy y Garfield señalan el efecto de los retrasos al hacer que los resultados de la evaluación estén disponibles para los clústers y otros responsables de la toma de decisiones. Asimismo, concluyen que “no se sabe con certeza si el análisis se hubiera hecho antes, se podría haber informado sobre la toma de decisiones en torno a los planes revisados de Clúster o Flash Appeal. Lo que es cierto es que, incluso asumiendo que el análisis fuera relevante y creíble, se produjo demasiado tarde como para informar las decisiones de planificación inicial” (Darcy y Garfield, 2011).

Un enfoque en fases para la evaluación, como el descrito en la guía MIRA (IASC, 2012a) puede ayudar a abordar esta limitación, al asegurar que se cuente con cierta información en un periodo de 72 horas, con análisis cada vez más complejos y disponibles en las próximas semanas y meses. Es evidente que la información generada en las primeras 72 horas es bastante limitada: las evaluaciones de necesidades sobre crisis repentinas, posiblemente más que en cualquier otra área de esfuerzo humanitario, desafían a los profesionales para encontrar un equilibrio entre la calidad probatoria y la velocidad – y el objetivo debe ser encontrar evidencia “lo suficientemente buena” para respaldar la toma de decisiones (Lars Peter Nissen, comunicación personal, diciembre de 2012).

El tiempo es lo que más escasea en muchos contextos humanitarios. Esto no solo detiene la espera de evidencia por parte de los responsables de la toma de decisiones, sino que también les evita pasar por extensos informes de evaluación. Si se va emplear evidencia, la presentación juega un papel importante. ACAPS ha encontrado que los informes cortos que presentan la información de forma gráfica tienen mayor probabilidad de influir en la toma de decisiones (ACAPS, 2012). Un estudio realizado por el Grupo de Políticas Humanitarias (HPG) sobre los Diagnósticos Humanitarios indicó que, en general, “la manera en la que se presenta la información puede ser crucial para su aceptación y uso por los responsables de toma de decisiones”, además recomienda que las evaluaciones sean cortas, sucintas e incluyan “hechos determinantes” para llamar la atención (Darcy, 2009: 11).

En algunos casos, particularmente cuando el responsable de la toma de decisiones se encuentra dentro de la organización que realizó la evaluación, o por lo general está predispuesto a emplear la información de la evaluación, asegurándose que los informes sean contundentes, claros y oportunos, puede ser suficiente para asegurar que se emplee su evidencia. Sin embargo, en la mayoría de los casos, el grado hasta donde se utiliza la evaluación también se determinará por los factores que se relacionen más con la persona responsable de la toma de decisiones – y el contexto dentro del cual se toman las decisiones – más que la evaluación en sí. En particular, es más probable que la evidencia de la evaluación se emplee cuando sugiere un curso de acción que es posible o conveniente para los responsables de la toma de decisiones, y cuando sus suposiciones (con frecuencia inconscientes) se adecúan a la situación.

Varios autores señalaron hace poco la importancia de una “trayectoria de dependencia” para determinar el grado en que se empleará la información de evaluación para la toma de decisiones (Darcy et al., 2013; Maxwell et al., 2013). Básicamente, la idea de una trayectoria de dependencia reconoce que las decisiones humanitarias no se toman de la nada, sino que se toman en un contexto donde algunos elementos ya están establecidos (o por lo menos, difíciles de cambiar). La mayoría de decisiones son bastante influenciadas por las decisiones que ya se han tomado sobre la estrategia organizacional, capacidad o asignación de fondos.

Al decidir sobre temas como el tamaño, ubicación y naturaleza de las intervenciones, “la mayoría de decisiones (tomadas dentro de los organismos) parecen haberse tomado dentro de parámetros bastante precisos: la variedad de opciones limitada por preguntas decididas anteriormente sobre prioridades estratégicas, recursos disponibles y así sucesivamente” (Darcy et al., 2013: 7). Del mismo modo, al decidir sobre qué tipo de intervención se desarrollará, los organismos no empiezan de la nada. Por el contrario, “las opciones de respuesta también se ven influenciadas por la capacidad y los etos organizacionales del organismo implementador” (Maxwell et al., 2013: 7).

Incluso donde existe más libertad para la toma de decisiones – donde las decisiones anteriores no determinan la asignación de recursos – una evaluación de necesidades solo es “el inicio de la variedad de información, influencias y consideraciones que...(los responsables de toma de decisiones) deben sopesar” (Poole y Primrose, 2010: 5). Poole y Primrose señalan que los donantes deben tener en cuenta una serie de factores durante la asignación de fondos, y que estos pueden ejercer una influencia válida en las decisiones. Incluso si las evaluaciones

sugieren una necesidad de intervenir en un cierto lugar y a cierta escala, los donantes pueden elegir no hacerlo ya que las necesidades son mayores en otro lugar, o debido a las condiciones de seguridad deficientes o una falta de confianza entre socios que hace improbable el éxito de una intervención. Las consideraciones políticas y mediáticas pueden también influenciar en la decisión.

Por lo tanto, la realidad de toma de decisiones humanitarias requiere una ponderación consciente de diferentes grupos de evidencia – en este caso la evidencia relacionada a la necesidad, a las capacidades y a las condiciones de seguridad – con el fin de emitir un juicio. En estas circunstancias, es posible que las buenas decisiones se entiendan mejor como “informadas en la evidencia” que “basadas en la evidencia”: ya que los diferentes componentes de la evidencia tenderán a respaldar las propuestas competitivas, aunque igualmente verdaderas. Sin embargo, al mismo tiempo, existen con frecuencia una serie de elementos menos racionales y menos conscientes que influyen en la toma de decisiones.

Es posible que al trabajar bajo presión en situaciones donde se deben considerar grandes cantidades de información, los responsables de la toma de decisiones suelen recurrir a “atajos” –suposiciones simples, generalizaciones y reglas empíricas, que facilitan el orden y el sentido de la información: los psicólogos las denominan heurística. La heurística servirá de filtro, y tendrá como efecto general la disminución del grado que las personas notarán y se comprometerán con la información inesperada, inusual o desafiante (Fiske y Taylor, 1991; Macrae et al., 1993), particularmente cuando se requieran decisiones y no se cuente con mucho tiempo (Bohner et al., 1995). En el ámbito humanitario se ha observado la tendencia humana general por seguir los “modelos mentales” existentes – incluso cuando son desafiados por nuevas evidencias – (Darcy, 2009; Maxwell et al., 2013; Knox Clarke, 2013b), y parecen influenciar la aceptación de la evidencia de evaluación en tres formas específicas.

La primera está relacionada a lo que Peter Walker denominó “modelos de inicio”. Hemos observado que con frecuencia las evaluaciones de necesidades se fortalecen en una serie de suposiciones – la más importante a menudo es la suposición incorrecta, en la cual no existe una capacidad local para responder ante la crisis. Se puede decir lo mismo de las operaciones humanitarias de forma más general: comienzan con suposiciones – o modelos de inicio – sobre una situación y este modelo se convierte en un tipo de heurística: una manera de filtrar y ordenar la información. En este contexto, la evidencia de las evaluaciones de necesidades que rechaza el modelo de inicio será ignorada. Un ejemplo de ello es la situación en Afganistán, donde después de una intervención militar en el 2001, “el modelo fue el conflicto finalizado, no hay problema en trabajar con el Gobierno – ellos no son combatientes” (Walker, 2013).

Los modelos de inicio son bastante duraderos porque suelen a ser compartidos por un gran número de personas en el sector humanitario. Esto nos proporciona un segundo principio heurístico: de conocimiento social. Cuando una descripción se vuelve ampliamente aceptada, toda información que la desmienta batallará para cobrar importancia. Poole y Primrose señalan “la descripción conceptual ampliamente sostenida de una rápida transición para el desarrollo” (Poole y Primrose, 2010: 1) en Sudán del Sur, donde es difícil destacar las necesidades humanitarias. Se puede observar una dinámica similar en la República Democrática



Es posible que las buenas decisiones se entiendan mejor como “informadas en la evidencia” que “basadas en la evidencia”: ya que los diferentes componentes de la evidencia tenderán a respaldar las propuestas competitivas, aunque igualmente verdaderas.



del Congo (RDC) donde una “descripción post conflicto” conllevó a una situación en la cual “el Plan de Acción Humanitaria de la RDC no cuenta en la actualidad con un buen financiamiento, y las necesidades humanitarias permanecen tan preocupantes como siempre” (Darcy et al., 2013: 10).

Por último, y posiblemente más extensamente, tenemos el principio heurístico “más vale malo conocido”, donde los responsables de la toma de decisiones tratarán de interpretar situaciones a través de experiencias previas. Estas experiencias son un tipo de evidencia en sí – aunque la evidencia no siempre es relevante ante la nueva situación y no siempre es contundente. Los responsables de la toma de decisiones también tratarán de incumplir las formas “intentadas y probadas” para responder ante las circunstancias (Maxwell et al., 2013), o en caso contrario, solicitarán niveles más altos de evidencia antes de apoyar los “nuevos” cursos de acción (Darcy, 2009). Esta tendencia a descartar la información que sugiere condiciones poco comunes o cursos de acción pueden, desde luego, ser respaldada por incentivos organizacionales para evitar correr riesgos.

De todos los factores descritos anteriormente – la trayectoria de dependencia, las prioridades competitivas y las diferentes heurísticas – aplican a las decisiones relacionadas a las propuestas de Tipo A (ya sea si se requiere una respuesta) y las propuestas de Tipo B (que tipo de respuesta se lleva a cabo). Sin embargo, existe también una limitación en el uso de las evaluaciones específicas para las propuestas de Tipo B, y que constituyen el enfoque del trabajo de Maxwell sobre el análisis de respuesta: una desconexión entre “el tipo de información normalmente proporcionado por evaluaciones y el tipo de información que requiere la elección de respuesta” (Maxwell et al., 2013: 1).

El problema descrito no radica en la calidad de la evidencia, ni que la evidencia no se pueda aplicar o se ignore, sino más bien en que la evidencia requerida para actuar está incompleta. El proceso de evaluación de necesidades no está recopilando toda la evidencia que necesitan los responsables de la toma de decisiones; asimismo, los sistemas de alerta temprana no son capaces de proporcionar información sobre las acciones que las organizaciones humanitarias deben tomar para prevenir una crisis. Hasta cierto punto, esto puede ser un reflejo de los “silos” que existen entre los distintos especialistas en información en el sistema humanitario: las personas encargadas de la evaluación de necesidades se concentran estrictamente en las necesidades y dejan la “respuesta” al personal del programa, de la misma forma que los evaluadores trabajan con frecuencia alejados de las personas que gestionan el programa (Argren, 2013; Hallam y Bonino, 2013).

Cualquiera sea la causa, la consecuencia se centra en que los responsables de la toma de decisiones no consiguen toda la información que requieren. La importancia de recopilar evidencia, no solo para las propuestas generales, sino para formular preguntas específicas que sean de interés para los responsables de la toma de decisiones, es un tema que se analizará en la siguiente sección.

EN RESUMEN:

- El uso de las evaluaciones de necesidades en la toma de decisiones es inconsistente.
- Esto es, en parte, una consecuencia de los problemas con las propias evaluaciones: no siempre son vistas como imparciales; no siempre se realizan a tiempo; pueden ser inaccesibles y “difíciles de leer”; no siempre incluyen la información que el personal del organismo necesita para planificar los programas.
- Es también una consecuencia de la forma en la que se toman las decisiones:
 - Los responsables de la toma de decisiones limitan con frecuencia la libertad en la toma de decisiones, muchas opciones han sido descartadas por estrategias o disposiciones organizacionales.
 - Se necesita también tomar en cuenta una variedad de información cuando se tomen decisiones, la información sobre las necesidades solo es una parte de un gran grupo de información.
 - Los responsables de la toma de decisiones son influenciados por sesgos inconscientes, los cuales pueden llevarlos a ignorar la información de las evaluaciones.

4.4 Uso de la evidencia de evaluación

El número de evaluaciones humanitarias ha crecido cada vez más en la última década. La Biblioteca de Recursos de ALNAP – que de ninguna manera es exhaustiva – contiene más de 1,200 evaluaciones de acción humanitaria. En efecto, es una de las fuentes únicas más grandes para evidenciar “qué funciona” (y qué no) en la respuesta humanitaria internacional.

Como indican Telford y Cosgrave en la síntesis de la Coalición de Evaluación del Tsunami (TEC), en un contexto donde muchos informes de organismos “se concentran en el éxito, ignoran o restan importancia sobre el fracaso...(y) la cobertura mediática suele concentrarse en casos dramáticos únicos, en vez de en una revisión balanceada de la calidad total, la información más detallada sobre el desempeño del organismo se puede obtener en los informes de evaluación de organismos” (Telford y Cosgrave: Synthesis 2006: 108).

Aunque el número de evaluaciones ha estado creciendo, parece haber algún escepticismo en cuanto al grado en que realmente se emplean estas evaluaciones. El uso de la evidencia de evaluación relaciona tanto la práctica operacional como la elaboración de política en el sector humanitario. Sandison (2007), que sigue a Patton (1997), describe tres usos principales de los hallazgos de evaluación:

- i. Juzgar el mérito o valor de un programa (ej. rendición de cuentas a las partes interesadas; para informar las decisiones de financiamiento).
- ii. Mejorar un programa (ej. aprendizaje y desarrollo continuo).
- iii. Generar conocimiento (que se puede usar en programas posteriores y en cualquier lugar).

Como indica Sandison, ambos usos (i) y (ii) se pueden considerar como “instrumentales”: la evaluación se emplea para realizar cambios directos (con

frecuencia basados en las recomendaciones de los evaluadores) al programa, sector o sistemas que se están evaluando. La mayoría de evaluaciones en el sector humanitario se diseñan principalmente para este tipo de uso instrumental.

El tercer tipo de uso (iii) es más general y menos directo. Aquí la información contenida en la evaluación se añade a una base de datos general sobre un tema y se emplea para realizar cambios y mejoras indirectamente y a cierta distancia. Si bien esta área se puede considerar como la más apropiada en el campo de la investigación, las evaluaciones realizan con frecuencia una contribución importante para generar conocimiento (por ejemplo, véase los Documentos de lecciones de ALNAP: (Hedlund y Knox Clarke, 2011; Sanderson y Knox Clarke, 2012).

En cuanto a este estudio, es probable que los primeros dos tipos de uso tengan un menor umbral probatorio que el tercero. Si bien la evidencia informa sobre cambios en un programa específico o práctica específica, se deben cumplir las normas mínimas (idealmente) *de exactitud, representatividad y relevancia*, además abordar efectivamente los temas de *atribución*, la evidencia no necesita ser *generalizable*. La tercera categoría es diferente, y el alcance que se toma como contribuyente al conocimiento va más allá de la situación específica evaluada en torno al grado con el cual se cumplen todos los criterios para la calidad probatoria, incluyendo la *generalizabilidad*.

El uso instrumental de las evaluaciones

Con respecto al uso instrumental de las evaluaciones, las conclusiones de Sandison son bastante negativas: "Solo una minoría de evaluaciones son efectivas al introducir cambios evidentes o mejoras en el desempeño" (Sandison, 2006: 91). Además agrega que "el uso instrumental es la forma menos probable de utilización" (ibid.: 121). Sin embargo, también indica que "no conocemos ni cuántas evaluaciones se llevan a cabo, ni mucho menos cuántas se emplean", por lo tanto "la fuente de preocupación con respecto al no uso en el sector es casi anecdótica" (ibid.: 91).

En realidad, si bien no es difícil encontrar ejemplos de recomendaciones de evaluación que se han ignorado, y mientras "que la literatura presente un registro inconsistente y, en algunos casos, un deficiente del uso de evaluaciones" (Sandison, 2006), la situación no es de ninguna manera completamente negativa. Los evaluadores profesionales contactados durante la preparación de este documento señalaron de forma consistente a recomendaciones que ya se habían implementado. Por ejemplo, el informe de la TEC aseguró haber llevado a cabo mejoras en la capacidad de acceso en el sistema y haber generado ímpetu al trabajo del Equipo de Tareas de Evaluación de Necesidades. Asimismo, el Segundo Clúster de Evaluación condujo hacia un enfoque enriquecido en el compromiso con la autoridad local en respuestas internacionales. Además, los evaluadores que participaron en la Comunidad de Evaluación de Práctica de ALNAP y que contribuyeron al estudio de ALNAP *Using Evaluation for a Change (Utilizando la evaluación para lograr un cambio)* (Hallam y Bonino, 2013) fueron capaces de señalar las diversas situaciones donde se emplearon las evaluaciones e identificaron varias condiciones que respaldaron el uso de la evaluación.

La situación es similar cuando se observan los (relativamente pocos) registros cuantitativos de la implementación de recomendaciones de evaluación. La matriz de respuesta de la administración para la revisión intermedia de la OCHA del

Fondo Central de Respuesta a Emergencia demostró que, al año siguiente de la revisión, el 50% de las recomendaciones se implementaron (OCHA 2007). Cuando el PMA estudió el grado en que se habían incorporado las recomendaciones de evaluación, encontraron que el 54% habían sido implementadas mientras que el 65% se incluyeron en documentos posteriores. El PMA encontró que era más probable implementar las recomendaciones cuando fueran operacionales, en vez de estratégicas y donde la implementación requirió acción solo de un número limitado de personas. Las recomendaciones más amplias, o aquellas que requirieron una coordinación con los socios o las unidades de sedes centrales, tuvieron menos probabilidad de implementarse, al igual que recomendaciones con beneficios intangibles o aquellas que implicaron críticas del personal del PMA (PMA, 2005).

Estas observaciones sugieren que la evidencia proporcionada por las evaluaciones humanitarias se emplea con frecuencia para realizar cambios "instrumentales" en el financiamiento o en los programas, pero de una forma altamente selectiva. Para determinar si se debe emplear una evaluación, la calidad de la evidencia puede ser menos importante que el grado en que cualquier recomendación dada sea fácil de implementar.

Durante la última década, muchas organizaciones humanitarias – incluyendo el DFID, SIDA, UNICEF y PMA – han tratado de identificar las formas en las que pueden mejorar la aceptación y el uso de la evidencia evaluadora. ALNAP también ha publicado cuatro documentos sobre el tema, Hallam y Bonino (2013); Hallam (2011); Sandison (2005); y van de Putte (2001) que se basaron en estas experiencias y aquellas de los otros miembros de la Red. Esta investigación sugiere que las evaluaciones son más propensas a conllevar a cambios en la implementación del programa o financiamiento cuando ya exista un interés, o un debate en torno, al desempeño de un programa; cuando la producción de la evaluación coincida con una "ventana de decisiones" (como una extensión del programa); cuando los resultados se comuniquen en un formato apropiado y accesible para los responsables de la toma de decisiones y donde se cuenten con mecanismos para el "seguimiento".

Sin embargo, la única lección más importante que surge de estos estudios es posiblemente la importancia de comprometer a los responsables de la toma de decisiones operacionales en cada etapa del proceso de evaluación: desde la selección de las preguntas de evaluación hasta la recopilación de información para la implementación y el seguimiento. Esto ayuda a asegurar la *relevancia* de la evaluación ante las necesidades operacionales y fortalece la apropiación de los hallazgos. Al mismo tiempo – cuando la objetividad se ve como un elemento importante de rigor metodológico – la estrecha participación del personal del programa plantea interrogantes en torno a la objetividad de los hallazgos de la evaluación y puede conllevar a preocupaciones sobre la calidad probatoria de la evaluación. Como señala un autor, "existe generalmente una tensión entre la independencia de los departamentos de evaluación y su éxito en comprometer a los usuarios de la evaluación" (Foresti, 2007).

Los desarrollos recientes en la evaluación humanitaria han tratado de incorporar todo o parte de estos enfoques en un intento por incrementar la probabilidad del uso de evaluaciones. Se ha producido un creciente interés por la Evaluación en Tiempo Real (RTE) en un intento por generar información sobre el progreso de una operación que se puede emplear para una "corrección del curso" inmediata. En abril



La evidencia proporcionada por las evaluaciones humanitarias se emplea con frecuencia para realizar cambios "instrumentales" en el financiamiento o en los programas, pero de una forma altamente selectiva.



del 2011, el IASC incluyó Evaluaciones en Tiempo Real interinstitucionales como un componente necesario de todas las emergencias en todo el sistema (nivel 3). En este caso, la información no solo estaba disponible de manera oportuna, sino que el ejercicio de evaluación se relacionó explícitamente a la toma de decisiones, ya que la RTE se diseñó para “informar la reunión de los Directores a fines de un período de activación de 3 meses” (IASC, 2011). En la actualidad, el IASC está pensando poner en marcha Revisiones Operacionales en Tiempo Real, implementadas principalmente por Equipos Humanitarios de Países en primera instancia, una medida que puede ser destinada a incrementar la apropiación de los resultados.

En Haití, el Grupo URD ha implementado lo que Grünewald denomina “Evaluación Plus en Tiempo Real”. En este caso, un equipo realiza una serie de evaluaciones del mismo proyecto por un período de dos años, trabajando cerca al equipo del proyecto a lo largo del proceso. Las últimas evaluaciones se concentran mayormente en identificar el progreso hecho sobre las recomendaciones de las misiones anteriores y la identificación de nuevos desafíos. Como explica Grünewald: “Esto conlleva a un diálogo de gran alcance entre el evaluador y el personal del programa que transcurre a lo largo de la duración del proyecto...el evaluador pierde un grado de su independencia (aunque, felizmente no su objetividad) con el fin de convertirse en un agente de cambio...las adquisiciones en la mejora – que después de todo es el fin principal – hace que valga la pena” (Grünewald, 2012: para. 3).

ACH fomenta también el diálogo entre los evaluadores y el personal en el terreno, en un intento por incrementar el uso de lecciones de las evaluaciones. La organización ha cambiado su proceso de evaluación para asegurar que los evaluadores identifiquen a menudo las mejores prácticas como parte de su trabajo. Se le pide al personal del programa a considerar, debatir y trabajar en estas mejores prácticas, que se incluyen en una revisión de aprendizaje y se difunden a través de la organización (véase Learning Review (Revisión de Aprendizaje) de ACH, 2011). Por lo tanto, la objetividad evaluadora se mantiene y la unión crucial entre la evaluación y el aprendizaje organizacional se fortalece cada vez más (Guerrero, 2012; Allen, 2012).

Evidencia y políticas generales – el uso de evaluaciones e investigación

Esta sección analizará ahora los usos de evaluaciones no instrumentales y de investigación de forma más amplia. La pregunta general en este caso es cómo – y hasta qué grado – la evidencia de la acción humanitaria en contextos específicos informa sobre la práctica humanitaria a otros tiempos y en otros lugares. Por ejemplo, ¿hasta qué grado la evidencia de buena calidad de Afganistán en el año 2000 puede influenciar el programa en Zimbabue en el año 2012?

Existe una variedad de canales a través de los cuales la evidencia puede informar a la práctica. Por ejemplo, las síntesis de evidencia, como los “documentos de lecciones” mencionados anteriormente, pueden ser leídas por las personas en el terreno e informar directamente a sus acciones. Se puede hacer uso de la evidencia de las evaluaciones para crear normas, que posteriormente orienten la acción en el diseño del programa o en la entrega. En algunos casos, la evidencia se puede emplear para desarrollar los mecanismos de apoyo de decisiones o software y, por lo tanto, tendrán una influencia directa en las decisiones tomadas usando este software.

Sin embargo, el camino más importante de la evidencia a la práctica tal vez es mediante la política: la evidencia informa a la política, que a la vez orienta a la práctica a través de una organización. En esta sección, consideramos las formas en que la evidencia evaluadora y otras han influenciado en el desarrollo político, concentrándose particularmente en cuatro ejemplos: el programa basado en dinero en efectivo; la introducción de normas mínimas; la consideración de factores culturales en el programa humanitario; y la intervención temprana en emergencias relacionadas con la sequía.

Programas basados en dinero en efectivo

Uno de los desarrollos políticos más importantes en la última década ha sido la creciente aceptación y el apoyo del uso de dinero en efectivo, en lugar de las distribuciones de alimentos y otros bienes. Por ejemplo, la política de asistencia humanitaria de ECHO (Dirección General de Ayuda Humanitaria y Protección Civil de la Comisión Europea), junto con las directrices de la organización sobre el uso de dinero en efectivo, refleja cada vez más la amplia aceptación del programa de dinero en efectivo, y el organismo sobrepasó recientemente el límite de 100,000 euros en los programas de dinero en efectivo (DG ECHO, 2009). En el Reino Unido, la Revisión de la Respuesta Humanitaria ante Emergencias recomendó que el DFID "debe...convertir las respuestas basadas en dinero en efectivo en ayuda habitual y posición de recuperación ante sus socios" (Ashdown, 2011: 24). La USAID ha cambiado recientemente su política Título II de Alimentos para la Paz con el fin de incluir de manera explícita programas de transferencia de dinero en efectivo.

Dicha política de apoyo a nivel del donante ha llevado a un notable incremento en el financiamiento para los programas de dinero en efectivo dentro de las operaciones humanitarias: Development Initiatives informa que el gasto en los programas de transferencia de dinero en efectivo creció de \$74.9 millones en 2006 hasta \$188.2 millones en 2010 (Development Initiatives, 2012).

Aunque el financiamiento disminuyó luego en 2011, la tendencia general parece ser un incremento en el uso de dinero en efectivo, particularmente porque los grandes organismos planifican aumentar sus actividades en esta área cada vez más (Haley Bowcock, comunicación personal, diciembre de 2012). Como un enfoque del programa, "el trabajo en la ayuda humanitaria basado en dinero en efectivo ha cambiado...de radical y riesgoso a...convencional" (Ramalingam et al., 2009: 43).

¿Hasta qué grado la evidencia ha jugado un papel en esta variación política? Vale la pena que se repitan aquí los puntos claves del estudio de ALNAP que documentan la evolución del uso de dinero en efectivo en programas humanitarios (Ramalingam et al., 2009a). El estudio sugirió que aunque la historia del uso de dinero en efectivo en la respuesta de emergencia fue bastante prolongado, no fue hasta el año 2000 que estas experiencias dispersas se revisaron metódicamente en un documento único: *Buying Power: the use of cash transfers in emergencies (El poder de compra: el uso de transferencias de efectivo en emergencias)* (Peppiatt et al., 2000). Este documento fue seguido por el trabajo del Grupo de Políticas Humanitarias (HPG) en ODI, el cual publicó una serie de estudios que considera la utilidad del dinero en efectivo en contextos de emergencia. Muchos de aquellos involucrados en la adopción de programas de dinero en efectivo "citaron el trabajo de...HPG como crucial en... persuadir a un número de organismos a que iniciarán sus propios proyectos";

además, "la investigación fidedigna que documenta la viabilidad del dinero en efectivo en diversos escenarios...ayudó a la defensa interna y externa de las organizaciones" (Ramalingam et al., 2009a: 63–64). La investigación (con frecuencia en forma de estudios de casos) y las evaluaciones de programas de dinero en efectivo continúan llevándose a cabo y son cotejadas por el Consorcio de Aprendizaje de Programas de Transferencia en Efectivo (CaLP). En la actualidad, la página web de CaLP contiene 45 evaluaciones de programas de dinero en efectivo y 40 informes de investigación.

La evaluación de ALNAP fue que "la investigación y evaluación desempeñaron un rol importante" en la aceptación del programa de dinero en efectivo (Ramalingam et al., 2009a: 63). Los colegas en CaLP están de acuerdo sobre la importancia de la evidencia. Ellos observan la falta de evidencia en algunas áreas (particularmente alrededor del costo-eficiencia y costo-eficacia del dinero en efectivo, en vez de la ayuda en especie) como una limitación para una mayor aceptación del enfoque, por ello han elaborado hace poco un programa de investigación para abordar algunos de estos vacíos de evidencia.¹⁹ Al mismo tiempo, existe un reconocimiento general en el cual, la evidencia en sí no es suficiente para disipar las dudas y preocupaciones percibidas por muchos organismos en torno al uso de programas de dinero en efectivo, en especial durante emergencias complejas. Un artículo reciente de Degan Ali señala que, aunque "se disponía de evidencia de que las transferencias en efectivo fueron una opción viable y eficaz" en la región central y meridional de Somalia, el "rechazo de la comunidad humanitaria hacia el riesgo los hizo reacios al uso de programas de dinero en efectivo en etapas tempranas" y por lo tanto, "a pesar de una historia comprobada de eficiencia en la región, la decisión (final) del uso de dinero en efectivo fue más un resultado de pensamientos correctos y una falta de alternativas que cualquier evaluación de la eficiencia e la idoneidad del dinero en efectivo para satisfacer las necesidades básicas (Ali, 2012).

La introducción de normas mínimas

Buchanan-Smith considera la importancia de la evidencia de investigación en el desarrollo de la política humanitaria en un contexto muy diferente. Su evaluación "How the Sphere Project Came into Being" (Cómo se creó el proyecto Esfera) observa un cambio en particular – la decisión de introducir normas mínimas voluntarias para la acción humanitaria – y establece la compleja relación entre la Evaluación Conjunta de la Asistencia de Emergencia a Ruanda (JEEAR) y el desarrollo de las normas del Proyecto Esfera. Asimismo, concluye que, aunque la JEEAR tuvo un "impacto significativo" (Buchanan-Smith, 2005: 17), y ayudó a centrar la atención en la necesidad para establecer normas mínimas, de ninguna manera fue la única causa de estos cambios de política, los cuales tuvieron "menos que ver con la investigación, más por el contrario con la creciente preocupación(es) (en el sector humanitario)" (ibid.: 22). De hecho, si bien alguna investigación de la JEEAR fue influyente, muchas de las conclusiones más importantes "se eludieron y evadieron constantemente" (ibid.: 24). Más adelante consideramos una de las conclusiones que no condujeron a ningún cambio.

A partir de los ejemplos de los programas de dinero en efectivo y normas mínimas, parecería que la evidencia puede contribuir – y contribuye – al desarrollo de la política en el sector humanitario, pero que "el modelo sobre la elaboración de

¹⁹ Voir <http://www.cashlearning.org/what-we-do/research-focus>.

políticas como proceso racional que reúne evidencia y proporciona orientación para acciones adecuadas es bastante cuestionable" (Clarke y Ramalingam, 2008: 32). Por ejemplo, las evaluaciones son "importantes, pero solo es uno de los recursos e influencias para el cambio. Por lo general, se (les) da un rango medio en términos de...valor al responsable de toma de decisiones" (Sandison, 2006: 3).

El desarrollo de la política no se basa exclusivamente en la evidencia, y la evidencia no siempre se emplea para desarrollar la política. Nuestros dos ejemplos finales ilustran el fracaso del sistema humanitario para elaborar cambios estratégicos o políticos en base a la evidencia: la consideración de factores culturales en programas humanitarios y la intervención temprana en emergencias relacionadas a la sequía.

La consideración de factores culturales en programas humanitarios

La falta de consideración de los factores culturales es una de las conclusiones que surgen de la JEEAR y un tema que parece haber sido "constantemente evadido" según Buchanan-Smith (2005: 25).

En Ruanda, el equipo de evaluación señaló que "los organismos de ayuda tuvieron en general solo una comprensión bastante limitada sobre la estructura de la sociedad ruandesa; asimismo, no se tomaron en cuenta los puntos de vista de los beneficiarios...un gran número del personal de los organismos de ayuda no había trabajado antes en la región, pocos conocían la sociedad ruandesa y por lo tanto, no estaban al tanto de muchos temas de preocupación común para los ruandeses que hablaban kinyaruanda" (Borton et al., 1996a: 176).

Esta falta de conocimiento contextual conllevó a una serie de errores que redujeron la efectividad, eficiencia y *relevancia* de la respuesta, incluyendo: la distribución de productos inapropiados; la distribución de productos mediante mecanismos basados en la comunidad que no consideraron a personas vulnerables y permitió que los funcionarios construyeran una base de poder que contribuyó a la inseguridad; y apoyo a una política de repatriación temprana.

Teniendo en cuenta que muchos de estos errores no fueron cometidos por las ONG con experiencia en Ruanda, la evaluación concluyó que "las ONG que operaran en situaciones de emergencias complejas debían: contar con un personal profesional de campo calificado con experiencia previa en dichos trabajos y con una apreciación de la necesidad para ser sensible ante la cultura local; establecer relaciones con las organizaciones locales (e); incluir dentro de su personal al menos algunos asesores o personal con amplia experiencia en el país" (Borton et al., 1996b: 61).

Durante la siguiente década, estos hallazgos repercutieron en la investigación académica que apuntó hacia la importancia para comprender el contexto local en el cual se estaba desarrollando una respuesta de emergencia y tomando en cuenta las percepciones de las poblaciones locales (Dijkzeul, 2010). Las evaluaciones han a menudo retomado este tema (véase Ali et al., 2005; Oxfam, 2004; World Vision, 2011; Nicholson y Desta, 2010; Tache, 2010).

Sin embargo, durante la década posterior a la respuesta de Ruanda en el año 2006, la síntesis de la Coalición de Evaluación del Tsunami aún informaba ampliamente del "alejamiento...(de) las organizaciones locales;...el reemplazo del personal local apto



El desarrollo de la política no se basa exclusivamente en la evidencia, y la evidencia no siempre se emplea para desarrollar la política.



por personal internacional con una preparación deficiente; el dominio del inglés como una “lengua franca”;...aplicando condiciones más exigentes a los “socios” nacionales y locales que a aquellos aceptados por las organizaciones internacionales;...y una participación del beneficiario de baja calidad” que en conjunto conllevaron a “programas insensibles de conflictos y género, injusticias, humillaciones, ofensas e insultos culturales” (Telford y Cosgrave: Síntesis, 2006: 93–4).

La situación no parece haber mejorado mucho desde la respuesta ante el tsunami. Dieciséis años después de la JEEAR, y a pesar de la evidencia evaluadora y de investigación que sugiere una necesidad para el cambio, los beneficiarios aún sienten que no se les consulta de forma adecuada (ALNAP, 2012); muy pocos del personal “nacional” son promovidos a altos cargos operacionales (Buchanan-Smith y Scriven, 2011); las ONG locales y las organizaciones de sociedad civil no pueden por lo general llevar ayuda (ALNAP, 2012); y la rotación del personal internacional sigue siendo alta, impidiendo que los responsables de la toma de decisiones obtengan cualquier conocimiento profundo del contexto social, económico y político en el que trabajan (Bhattacharjee y Lossio, 2011; Currion, 2010; Darcy, 2012).

Intervención temprana en emergencias relacionadas con la sequía

Este no es el único caso en el que un fracaso puede crear respuestas de políticas contundentes ante las situaciones donde la evidencia sugiere que se requiere un cambio. Nuestro ejemplo final concierne a la forma en la que el sistema humanitario internacional responde a sequías en zonas de pastoreo.

La respuesta humanitaria “tradicional” ante la sequía ha sido una distribución de alimentos a gran escala, desencadenada generalmente por los altos niveles inaceptables de desnutrición. Sin embargo, durante las últimas dos décadas, se ha pedido el cambio hacia un modelo alternativo de “respuesta temprana”, donde los organismos responden a alertas tempranas de sequía con una serie de intervenciones de medios de subsistencia, asegurando el bienestar de los rebaños y aumentando el ingreso de ventas de ganado.

En su informe *Money to Burn (Dinero para gastar)* (2012), Catley y Cullis indican que estos enfoques se emplearon en el Sahel y en el Cuerno de África en la década de los 80 y 90. En el 2001, Aklilu y Wakesa, reflejan la respuesta ante sequías en Kenia de 1999 al 2001 y concluyeron que “el marco de políticas de respuesta ante sequías requiere de mucho análisis...que abarque mucho más que la ayuda alimentaria... para apoyar y mantener, no a las personas mismas, sino a su capacidad para comercializar y cuidar su ganado (2001: 33). Cuatro años después, Jaspars obtuvo conclusiones similares, luego de revisar la literatura y realizar estudios de casos (Jaspars, 2006).

Durante el resto de la década, una serie de otras evaluaciones y documentos de investigación apoyaron la idea de que se requería un cambio significativo en las políticas (véase Sadler et al., 2009; VSF, 2009; Burns et al., 2008; ODI, 2006). La evidencia sugirió que la respuesta temprana era más efectiva, más aceptable para las poblaciones locales, y significativamente más rentable (Abebe et al., 2008).

Estos estudios y los programas en que se basaron, conllevaron a una respuesta limitada. Un documento previo de ALNAP sobre estos temas indica que algunos

donantes han presentado mecanismos de financiamiento multianuales y flexibles, para permitir que las respuestas relevantes se realicen sin la necesidad de recurrir a nuevos fondos (Hedlund y Knox Clarke, 2011). Sin embargo, estas iniciativas son “a menudo de menor escala y no coinciden con las necesidades de las poblaciones afectadas” (ibid.: 6).

Durante la última década, las intervenciones de medios de subsistencia fueron en general pobremente financiadas (HPG, 2006); no fueron priorizadas por el Fondo Central de Respuesta a Emergencia (CERF) de la ONU (Pantuliano y Wakesa, 2008); y en Etiopía solo representó el 2.2% del financiamiento total para la ayuda ante la sequía del 2011 (Catley, 2012). En el 2012, 11 años después del llamado de Aklilu y Wakesa para analizar la respuesta ante la sequía, un informe financiado por el DFID recomendó que “la respuesta temprana y las medidas que refuerzan la resistencia debe ser la respuesta prioritaria contundente ante los desastres” (Venton et al., 2012). El mismo informe estimó que, si la comunidad internacional hubiera empleado la reducción de reservas como una opción predeterminada durante esa década, se habrían producido ahorros de casi \$11 mil millones. Lo que es más importante, “si una respuesta temprana hubiera salvado al menos una pequeña proporción de... vidas (perdidas como resultado de la sequía en los años 2010/2011) miles de niños, mujeres y hombres aún estarían vivos” (Save the Children y Oxfam, 2012: 13).

Afrontando las limitaciones ante el uso de evidencia

Existen, de hecho, muchas limitaciones para emplear la evidencia en el desarrollo de la política humanitaria. La JEEAR y el TEC observaron las verdaderas dificultades para generar cualquier evidencia, la cual sugiere que un organismo, una intervención o un enfoque pudo haber estado “errado”, y sin esto la creación de un conjunto sólido de evidencia se vuelve más difícil. En muchos casos, la evidencia está dispersa y no siempre está disponible en un formato único comparable (Borton et al., 1996a, 1996; Redmond, 2010). En el caso de las intervenciones tempranas, Levine y otros autores han observado que el número de programas, y por ende, la base de la evidencia, aún es bastante limitada (Levine et al., 2011).²⁰ Sin embargo, la falta de evidencia comparable no es una explicación suficiente ante la falta de atención prestada por los responsables de formular políticas al conocimiento contextual limitado o la respuesta tardía.

Otra explicación para la relativa inactividad de la comunidad humanitaria frente a la evidencia es que muchos temas son justamente demasiado difíciles de resolver. En otras palabras, no hay mucho que ganar del creciente conocimiento contextual donde otros factores dificultarían o hasta imposibilitarían el uso de este conocimiento. Estos factores incluyen: “la naturaleza de la inflexibilidad e impulsada por la oferta del sistema de ayuda internacional” (Borton et al., 1996a: 177); “restricciones estipuladas al donante sobre cómo (los organismos) emplean los fondos” (TGLLP Steering Committee: 11), como aspecto de la “trayectoria de dependencia” debatida anteriormente; y “la urgencia de gastar dinero visiblemente” (Telford y Cosgrave, 2006: 93).

20 Andy Catley sugiere otra razón interesante de por qué la base de evidencia formal para la intervención temprana puede no ser grande: que el enfoque se basa en una “causal lógica” bastante sólida, de modo que los profesionales no han considerado ningún requisito para probar la suposición formalmente, lo cual genera una pregunta interesante sobre el rol de los modelos lógicos en aportar evidencia.

Existe probablemente algo más de verdad en esto, aunque una vez más, solo es una parte de la historia. Los cambios de política que apuntan al fortalecimiento de las relaciones a largo plazo con actores locales de la sociedad civil, asegurando la toma de decisiones por parte del personal local contratado o disminuyendo las rotaciones de personal en emergencias, parecen factibles y ayudarían a garantizar los programas basados en conocimientos contextuales más contundentes. Los donantes podrían liberar más fondos antes de los desastres repentinos y los organismos se podrían preparar mejor para intervenir. Como muestra el ejemplo del programa de dinero en efectivo, la evidencia puede contribuir con las políticas que desafíen los elementos existentes del paradigma humanitario – en el caso del dinero en efectivo, la percepción de que “el dinero en efectivo no era factible porque no se podía confiar en los destinatarios para que lo gasten de forma efectiva” (Ramalingam et al., 2009a: 44). Entonces, ¿por qué se ignora frecuentemente la evidencia? y ¿cuánto demora obtener la evidencia empleada?

Los desafíos sobre el uso de evidencia para desarrollar la política no son exclusivamente prácticos: Sandison señala que es difícil tomar acciones (instrumentales o generales) en las evaluaciones que “desafían creencias y comportamientos fuertemente arraigados en la cultura de la organización” (Sandison, 2006: 111), y concluye que “emplear una evaluación es tanto un problema de las personas como un problema técnico” (ibid.: 132). En su estudio de cambio en las organizaciones humanitarias, Clarke y Ramalingam indican que “los entrevistados conversaron sobre “respuestas instintivas” que eran aparentemente, modificaciones técnicas bastante simples (2008: 45).

El cambio efectivo – que incluye el desarrollo y la introducción de una nueva política – requiere un proceso que aborde las necesidades racionales, políticas y emocionales de las partes interesadas dentro de la organización. Al reconocer esto, el Programa de Investigación y Política en el Desarrollo (RAPID) en ODI ha creado un marco para analizar la influencia de la investigación en la política (Young y Court, 2004). El marco observa no solo la credibilidad y comunicación de la información de evaluación, sino también las relaciones entre los evaluadores, elaboradores de políticas y otras redes; en el contexto político; y en la influencia del ambiente externo. Estos factores también parecen ser importantes para determinar el grado hasta el cual la evidencia influye en la política humanitaria.

El marco RAPID enfatiza la importancia de la evidencia comunicativa a los responsables de la toma de decisiones. En el caso del dinero en efectivo, el estudio de ALNAP encontró que “el uso de resultados (de las evaluaciones) en formas simples y de gran alcance...fue crucial” (Ramalingam et al., 2009a: 3). Como se mencionó, aunque los programas de dinero en efectivo habían ya aparecido hace algún tiempo, los elaboradores de políticas accedieron por primera vez a los resultados a partir del informe *Buying Power (El poder de compra)* de la Cruz Roja Británica en el 2000.

El HPG y recientemente el CaLP han sido influyentes en asegurar que la evidencia y el aprendizaje estén disponibles y sean recopilados. Del mismo modo, la JEEAR “dispuso claramente y analizó qué fue lo que la mayoría de organismos humanitarios ya sabían, de ser el caso” (Buchanan-Smith, 2005: 22) y se beneficiaron de un proceso de seguimiento financiado que permitió al equipo de evaluación “vender” el informe y los mensajes claves que contenían. Por el

contrario, la falta de atención a los temas de contexto puede ser el resultado, en parte, de la carencia de una síntesis clara de la evidencia: "si bien la atención a las percepciones locales de acción humanitaria han aumentado, no ha sido lo suficientemente sistemática...dichos estudios rara vez hacen referencia uno a otro" (Dijkzeul and Wakenge, 2010: 1146). Debido a la importancia de la accesibilidad a la evidencia, el trabajo de los organismos como Oxfam, Care y NORAD en la síntesis y publicación de investigaciones y evaluaciones (Hallam, 2011), además de grupos y redes como ALNAP pueden contribuir a un sistema que se base más en la evidencia (Dijkzeul et al., 2012).

Sin embargo, parecería que el acceso a la información no es suficiente para garantizar su uso. Varios estudios del mundo humanitario han sugerido que los responsables de toma de decisiones humanitarias en todos los niveles tienden a ser bastante influenciados por las actitudes y opiniones de sus pares (Clarke y Ramalingam, 2008; ALNAP, 2003; Sandison, 2007; Darcy, 2009; Maxwell et al., 2013). Esto sugiere que en el sector humanitario, el conocimiento se construye y se valida socialmente y para que la evidencia se emplee, debe primero convertirse en una parte del discurso humanitario.

La importancia de las redes y relaciones al impartir conocimiento "aceptable" se ha observado en otra parte. Latour ha demostrado "cómo el hacer ciencia es un esfuerzo social donde convencer a las personas de aceptar ciertas verdades depende más de las relaciones sociales que del uso de métodos científicos" (citado en Hilhorst, 2003) mientras Jones y Mendizabal sugieren que "las relaciones interpersonales directas entre el personal y tanto investigadores como evaluadores...son muy importantes" para hacer que la evidencia se emplee (Jones y Mendizabal, 2010).

Esta tendencia, que tiene implicaciones preocupantes sobre la habilidad del conocimiento "local" para influenciar la dirección de la acción humanitaria, ha sido influyente en la aceptación de dinero en efectivo como una herramienta de programación. Ramalingam et al. (2009a: 55) sugiere que "el surgimiento de un grupo disperso con experiencia basada en el terreno que empezó a explorar...las posibilidades para un programa de dinero en efectivo y abordar las preocupaciones particulares de los escépticos", y que conllevó a la iniciativa de aprendizaje basada en dinero en efectivo seguida al tsunami del año 2004 (y posteriormente CaLP), fue importante para generar la aceptación del enfoque. Este impulso social, combinado con la evidencia de las evaluaciones, sigue siendo importante en el desarrollo continuo del programa de dinero en efectivo.²¹

Del mismo modo, la JEEAR fue una iniciativa interinstitucional, con amplia participación a través del sistema humanitario: la investigación fue de naturaleza social desde el inicio. Por otro lado, Levine et al. (2011) sugiere que no existe una "plataforma" para el debate de respuestas tempranas, y que la mayoría de debates son bilaterales, y se relacionan con programas específicos: la red social en torno a respuestas tempranas parece no existir.

El marco de RAPID también resalta la importancia de la política (organizacional) y la presión externa para determinar hasta qué grado se emplea la evidencia. Los



En el sector humanitario, el conocimiento se construye y se valida socialmente.



21 Según el foro de debate de CaLP, que amablemente proporcionó sus puntos de vista sobre este tema.

elaboradores de políticas humanitarias son selectivos, "filtran" evidencia, y "por último, toman la decisión sobre qué recomendaciones de los investigadores para el cambio de política (están) dispuestos a aceptar" (Buchanan-Smith, 2005). Como resultado, "el sistema humanitario...es el más receptivo para el cambio bajo presión, cuando los factores de impulso son contundentes" (ibid.: 98).

Estos factores de impulso fueron particularmente contundentes tras la respuesta en Ruanda, por ejemplo, con los organismos involucrados en debates internos sobre cómo mejorar y donantes exigiendo acción, y no cabe duda que esto aseguró que por lo menos se prestara atención a la evidencia de los equipos de la JEEAR.

Algunas recomendaciones – como aquellas en torno a las normas – fueron impuestas a través del filtro. En el caso de los programas de dinero en efectivo, los acontecimientos se produjeron en un contexto de preocupación que existía desde hace tiempo por los efectos de la ayuda alimentaria, los que eran influenciados externamente por una variedad de factores. Estos incluyeron el crecimiento masivo en el financiamiento que se realizó luego del tsunami, el apoyo de los gobiernos en la región del Océano Índico a los programas de dinero en efectivo y el incremento mundial en los precios de los alimentos y el petróleo en 2008, que hizo que la entrega de ayuda alimentaria fuera más costosa.

Por otro lado, en el caso de respuesta temprana, los factores políticos y organizacionales parecen ir en contra del cambio de política. Muchos organismos evitan programas de medios de subsistencia porque carecen de habilidades y conocimiento contextual requerido (Aklilu and Wakesa, 2001), y porque el modelo de "cuerpo de bomberos" de establecer una presencia en una área donde ocurre un desastre no es efectiva para la respuesta temprana (HPG, 2006). Asimismo, algunas organizaciones reciben un financiamiento importante de monetización (grandes cantidades) de la ayuda alimentaria requerida para abordar condiciones críticas (Jaspars, 2006): las sumas relativamente limitadas requeridas para el apoyo de los medios de subsistencia no proporcionarían el mismo nivel de ingresos.

Mientras tanto, puede ser que los donantes no estén dispuestos a responder solo en base a predicciones, requiriendo de "datos concluyentes" antes de consignar el dinero de los contribuyentes (Save the Children y Oxfam, 2012; Levine et al., 2011) y con frecuencia son incentivados a gastar mayores sumas de dinero que las ONG solicitan para intervenciones tempranas. Levine et al. cita a un representante de los donantes diciendo que "las ONGs toman pequeñas cantidades de dinero...si giramos un cheque cuantioso a la ONU, podemos darlo por perdido inmediatamente" (2011: 7). Existen también limitaciones importantes para el cambio a nivel político: "Los gobiernos nacionales consideran a menudo una declaración de emergencia como signo de "debilidad" (Save the Children y Oxfam, 2012) y por eso retrasan el "llamado" de emergencia hasta cuando es demasiado tarde para que los enfoques de medios de subsistencia sean particularmente efectivos (Hedlund y Knox Clarke, 2011).

La trayectoria de la evidencia en la políticas es rara vez clara o directa. En algunos casos -y en particular aquellos cuando los desafíos de evidencia se basan en el sentido común y enfoques estándar, puede desaparecer por completo. No obstante, en muchos casos, la evidencia informa (si no sirve de referencia) a las políticas: este es el caso particular donde la evidencia se acepta de forma general, ya sea por el

apoyo de alianzas u otros programas, o bajo la influencia de una crisis que hace a las personas más abiertas a un cambio.

EN RESUMEN:

- Existe una diferencia entre el uso “instrumental” directo de las evaluaciones y su rol indirecto en la influencia a la política humanitaria a más largo plazo.
- Las evaluaciones se emplean con frecuencia instrumentalmente (para el rediseño inmediato de los programas o proyectos evaluados), aunque este uso es selectivo.
- Es poco probable que se empleen recomendaciones “difíciles de implementar”.
- ALNAP ha investigado los factores que conllevaron al uso instrumental: el más importante parece ser un compromiso cercano de los responsables de la toma de decisiones con el proceso de evaluación y actualmente muchas organizaciones humanitarias emplean métodos para respaldar este tipo de compromiso.
- El rol de las evaluaciones (y de la investigación evaluadora) en la influencia a la política se considera a través del uso de cuatro ejemplos: programa de dinero en efectivo; programas de diseño para la adecuación contextual; la evolución de las normas humanitarias; y respuesta temprana ante situaciones de sequía.
- Las evaluaciones son solo una fuente de evidencia para el desarrollo de la política, y su influencia en la política se limita con frecuencia por el hecho que las evaluaciones pueden ser difíciles de hallar (en parte un resultado de una renuencia por parte de los organismos para difundir evidencia que se refleja de forma negativa en sus programas), y por la resistencia natural al cambio demostrado por los individuos y organizaciones en el sector humanitario.
- Además, muchas evaluaciones apuntan a cambios que podrían requerir reajustes fundamentales de la arquitectura humanitaria y el “modelo comercial” humanitario.
- Las evaluaciones tienen más probabilidades de servir de base a la política en situaciones donde ya existe presión para el cambio.
- Las plataformas que reúnen una variedad de actores y “socializan” el conocimiento encontrado en las evaluaciones, también son importantes para respaldar el uso de las evaluaciones.

5

Conclusiones y recomendaciones: ¿Cómo se puede mejorar la calidad y el uso de la evidencia en la acción humanitaria?

5.1 Mejorando la calidad de la evidencia

La buena evidencia – y en particular la evidencia sobre la naturaleza de necesidades humanitarias y sobre el éxito o fracaso de respuestas – constituye un componente esencial del programa humanitario. Como resultado, los procesos para la recopilación de información y el análisis se encuentran en varias etapas en el “ciclo del programa” humanitario estándar.

Una evaluación de la calidad de evidencia generada por estos procesos (empleando los criterios para la “buena” evidencia descrita en la Sección 2) sugiere que aún hay mucho por mejorar. Esta revisión también sugiere que una mejora es posible: mientras que los contextos humanitarios presentan serias limitaciones para la generación de evidencia, muchas organizaciones dentro del sistema humanitario están encontrando formas para vencerlas.

En base a esta revisión, sugerimos que los esfuerzos para mejorar la calidad de la evidencia en el sector humanitario deben considerar cinco principios claves.

METODOLOGÍA CONTUNDENTE: Gran parte de la evidencia generada en el sector se obtiene mediante métodos cualitativos. Sin embargo, la comprensión e implementación de estos métodos es con frecuencia deficiente. Los organismos pueden – y deben – mejorar su rigor metodológico en esta área. Al mismo tiempo, algunos organismos están experimentando con el uso de métodos cuantitativos de recopilación y análisis de datos. Estas iniciativas han demostrado no solo las posibilidades de estos enfoques que se deben alentar, sino también las limitaciones. Se debe también aceptar la tentativa de trasladar los enfoques de “método combinado” hacia la generación de evidencia.

Las alertas tempranas, la evaluación, la valoración y los informes de investigación generados por los actores humanitarios deben ser claros sobre la solidez de la evidencia empleada, además deben incluir una descripción de la metodología empleada para la recopilación de información y el análisis, así como las limitaciones de esta metodología.

Las organizaciones deben compartir experiencias y aprendizajes sobre su uso de metodologías cualitativas y cuantitativas, redes bilaterales y directas, grupos inter-organizacionales como ALNAP, ACAPS y los Grupos de Trabajo para la Evaluación de Necesidades del IASC.

Estas redes deben sintetizar las mejores prácticas y deben respaldar la producción de materiales de capacitación y otras herramientas para la difusión de mejores prácticas.

Las organizaciones humanitarias deben considerar el establecimiento de organismos para revisar la calidad de sus informes.

Las organizaciones humanitarias deben evaluar si tienen la capacidad para llevar a cabo la recopilación de evidencia y el análisis relevante para su trabajo, en momentos que no lo hagan, deben llenar vacíos mediante asociaciones estratégicas, reclutamiento y capacitación.

Los actores humanitarios se deben asociar con los académicos y aprender de ellos para mejorar las normas de generación de evidencia.

EFICIENCIA: La evidencia es una aportación al programa humanitario y como tal se debe considerar el tiempo y dinero invertido para generar evidencia en el contexto de gasto total en los programas actuales y futuros.

Los donantes y organizaciones humanitarias deben considerar explícitamente los beneficios potenciales – tanto directos como a largo plazo – cuando invierten en la recopilación de evidencia y estén preparados para invertir sumas proporcionales a los ahorros o mejoras que se podrían realizar.

COLABORACIÓN: Una de las limitaciones claves para la generación efectiva (y uso) de evidencia de alta calidad en el sistema humanitario es una falta de colaboración dentro y entre las organizaciones. La evidencia se recopila con frecuencia en silos funcionales u organizacionales. Esto previene la realización de revisiones cruzadas y controles de calidad rigurosos, fomenta la repetición de actividades y rechaza efectivamente a la mayoría de actores humanitarios que acceden a grandes cantidades de evidencia. Las iniciativas que apuntan a la generación de evidencia humanitaria deben reconocer una responsabilidad para contribuir a un “conjunto de evidencia” en todo el sector.

Los donantes deben emplear las redes y programas interinstitucionales para identificar los “vacíos de evidencia” que existen dentro del sistema humanitario, y desarrollar enfoques comunes para llenar dicho vacío.

Todos los actores humanitarios, y en particular los donantes, deben seguir apoyando los procesos para identificar marcos globales estandarizados que se pueden emplear en las alertas tempranas conjuntas o coordinadas, evaluación y valoración.

Los HCT, Clústers y Mecanismos de Coordinación entre Clústers deben aclarar las necesidades de la evidencia a nivel nacional y deben fomentar las valoraciones y evaluaciones conjuntas o coordinadas.

Todas las actividades de recopilación de evidencia debe incluir presupuestos y planificaciones para asegurar la difusión de la evidencia y ampliar el uso de conjuntos de evidencia en todo el sistema (como ACAPS, ALNAP y CaLP) para apoyar la difusión

CONTINUIDAD: Los ciclos cortos de financiamiento y la rotación rápida de personal suelen con el tiempo a ir en contra de la recopilación de conjuntos de evidencia comparables. Como resultado, la evidencia se produce a menudo de forma sucesiva, en valoraciones y evaluaciones únicas. Esto reduce considerablemente nuestra habilidad para entender los efectos de la intervención humanitaria con el tiempo, y nuestra capacidad para decir con confianza, cuándo y cómo debemos intervenir en el futuro. Los actores humanitarios deben ver la generación de evidencia como un proceso y no como un caso, además deben con el tiempo apuntar a fortalecer el conjunto de la evidencia.

Los agentes humanitarios deben abstenerse del rediseño continuo de evaluación y sistemas de monitoreo. Los sistemas de evaluación y sistemas de monitoreo se deben diseñar para complementarse entre sí. Cuando los sistemas se rediseñan, se debe tener cuidado para seguir con la recopilación de datos en indicadores claves.

Los agentes humanitarios deben respaldar el trabajo nacional para crear sistemas de monitoreo y otros sistemas que puedan proporcionar datos de referencia y series con tiempos más prolongados de datos para indicadores claves, mediante la reducción de riesgo ante desastres y la planificación de contingencia.

INCLUSIÓN Y APROPIACIÓN: Todos aquellos profundamente preocupados por la acción humanitaria – y afectados de manera directa por una crisis – suelen ser dejados de lado por los enfoques actuales sobre la generación de evidencia en dos formas importantes. Primero, no logran formular preguntas: la evidencia se recopila por lo general para satisfacer las necesidades de las organizaciones internacionales, en vez de las necesidades de las personas afectadas. Por ejemplo, las evaluaciones tratan de obtener evidencia para cualquier tipo de ayuda necesaria, en vez de saber en primer lugar si la ayuda es necesaria. Segundo, a menudo se da poca importancia a sus respuestas: esta revisión ha sugerido que incluso cuando se reúne evidencia para entender si la situación es una crisis, o si “funcionó” una intervención particular, el conocimiento y las opiniones de la mayoría de afectados directamente suelen recibir solo una atención limitada. En el futuro, las organizaciones internacionales deben aclarar por qué recogen evidencia y para quién lo hacen; deben considerar hasta qué grado pueden recopilar evidencia que sea de uso para las organizaciones de sociedad civil en las áreas afectadas y esforzarse para incluir en su recopilación de evidencia las voces de las personas afectadas.

Las organizaciones internacionales deben identificar la información y las necesidades de evidencia de las organizaciones de sociedad civil en las áreas afectadas por los desastres y brindar información para cumplir estas necesidades. De ser relevante, deben brindar también capacitaciones y recursos que permitan recopilar información a las organizaciones nacionales y locales.

La alerta temprana, valoración, evaluación e informes de investigación realizados por los actores humanitarios deben presentar el conocimiento local y las propuestas de manera clara. Si las evaluaciones no están conformes con la información local o las propuestas, los autores deben explicar el porqué del caso.

En todos los casos, la información se debe recopilar en base a un consentimiento informado. Además, las organizaciones que recopilan información deben identificar

las maneras de informar a las personas sobre cómo se empleó la información y sobre cualquier decisión que se tomó en base a esta información.

5.2 Mejorando el uso de la evidencia en la toma de decisiones humanitarias

La presente revisión también consideró hasta qué grado los responsables de la toma de decisiones dentro del sistema humanitario emplean la evidencia cuando se toman decisiones. Por lo general, se encontró que es posible realizar mejoras significativas en el uso de la evidencia por los agentes humanitarios: existen muchos casos donde los procesos de toma de decisiones han ignorado evidencia importante de los sistemas de alerta temprana, valoraciones o evaluaciones de necesidades con consecuencias negativas para las vidas y medios de subsistencia de las personas afectadas por el desastre. Este es el caso particular donde la evidencia desafía las interpretaciones preconcebidas de una situación o puntos ante una necesidad de cambio radical.

Al mismo tiempo, se debe dar mayor importancia a cómo la evidencia se puede o debe emplear en la toma de decisiones. La mayoría de decisiones en el sector humanitario requieren consideración de múltiples conjuntos de evidencia, con frecuencia recopilados mediante el uso de distintas metodologías y algunas veces indicando diferentes cursos de acción. En estas circunstancias, no es razonable ni conveniente esperar que solo la evidencia determine la decisión. Habrá siempre un elemento necesario de juicio en la toma de decisiones y se debe asegurar que en todos los casos, estos juicios sean informados mediante la mejor evidencia disponible.

Mejorando el uso de la evidencia – recopilación, análisis y difusión

En base a esta revisión, se sugiere que la evidencia tiende más a influir en la toma de decisiones de aquellos responsables orientados por los siguientes principios para generar evidencia:

ACCESIBILIDAD: Los responsables de la toma de decisiones rara vez tendrán tiempo (y en ocasiones pueden carecer de disposición) para buscar la información que necesitan en bases de datos, informes u hojas de cálculos y así tomar una decisión contundente. La evidencia se debe agrupar de forma que sea fácilmente accesible: informes cortos en un lenguaje sin argot, infográficos y “portales” con capacidad de búsqueda.

Las redes de aprendizaje humanitario (como ALNAP), iniciativas interinstitucionales (como los clústers mundiales) y las organizaciones especializadas se deben esforzar para asegurar que la información y la evidencia sean de acceso inmediato, a través de:

- Síntesis de las mejores prácticas y lecciones aprendidas de las evaluaciones e investigación.
- La creación y mantenimiento de plataformas de fácil acceso, abiertas a la búsqueda de documentales y otras formas de evidencia, con aquellas entidades que crean bases de datos, asegurando que colaboren con iniciativas existentes para prevenir la duplicación y confusión.

PUNTUALIDAD: Información que llega después de haber tomado las decisiones y que frecuentemente se ignorará, así como las decisiones más importantes – aquellas que establecen el amplio alcance de un programa – se tomarán a menudo oportunamente. Los grupos o individuos que generan evidencia deben apuntar hacia el entendimiento del proceso de toma de decisiones y poner a disponibilidad la evidencia relevante en puntos críticos en el proceso.

Las organizaciones humanitarias responsables de alerta temprana, valoración y evaluación deben evaluar hasta qué grado se generan estos productos en el tiempo correcto para influenciar las decisiones. En caso de ser necesario, deben intentar alterar sus procesos para asegurar que la información esté disponible de manera oportuna.

AMPLIA CIRCULACIÓN MEDIANTE UNA VARIEDAD DE MEDIOS: El conocimiento en el sector humanitario parece que, en gran parte, se construye socialmente. Los responsables de las tomas de decisiones están influenciados por el conocimiento y pensamiento de sus pares. Como resultado, la información que ha circulado ampliamente, que es conocida y debatida ampliamente tiene mayor probabilidad de ser empleada como evidencia para respaldar o ir en contra de un cierto curso de acción.

Las organizaciones que generan evidencia deben asegurarse de que se difunda activamente mediante varios métodos que incluyan seminarios web, comunidades de práctica, reuniones, videos y cursos de capacitación.

Deben comprometer una variedad de grupos, que incluyan a los elaboradores de políticas, profesionales y líderes organizacionales mediante estos diversos medios.

Los donantes que respaldan la generación de evidencia deben asegurarse que los presupuestos incluyan planificaciones para la difusión de resultados.

Mejorando el uso de la evidencia – toma de decisiones

Si bien los agentes humanitarios tienen una responsabilidad de asegurar que su evidencia sea accesible, oportuna y ampliamente circulada, estas actividades no serán por sí solas suficientes para incrementar el compromiso de los responsables de la toma de decisiones con la evidencia. Las organizaciones miembros de ALNAP deben también tomar acción para mejorar su enfoque ante la toma de decisiones.

ESTABLECER PROCESOS CLAROS PARA LA TOMA DE DECISIONES: La evidencia tiene mayor probabilidad de ser empleada donde las organizaciones instauren procesos de toma de decisiones que incluyan etapas claras que requieran evidencia para ser consideradas.

Las organizaciones humanitarias deben describir claramente el proceso por el cual se diseñan (y rediseñan) los programas, indicando las etapas claves donde se requiera información y, de ser posible, la naturaleza de la información requerida para tomar decisiones.

Los donantes deben solicitar que las organizaciones humanitarias que les brindan propuestas de financiamiento proporcionen la mejor evidencia disponible (considerando el tiempo e implicaciones de recursos) para demostrar que:

- Se requiere ayuda externa.
- La respuesta propuesta es la mejor opción disponible.
- La organización encargada de la propuesta tiene la capacidad para implementar la respuesta efectivamente.

Los organismos humanitarios deben considerar la introducción de matrices para rastrear la respuesta/decisión (donde no se emplean todavía) para permitir a las partes interesadas analizar hasta qué grado la organización ha actuado en las recomendaciones de evaluación.

En el campo de alerta temprana/respuesta temprana, el sistema humanitario debe establecer una plataforma común para considerar el proceso completo, desde alerta temprana hasta acción temprana, enfocándose en el rol de los sistemas de alerta temprana para informar decisiones; la ubicación de la autoridad de toma de decisiones en los organismos de respuesta; y el aprendizaje conjunto en torno a las actividades de respuesta temprana.

ASEGURAR QUE EXISTAN INCENTIVOS PARA EL USO DE EVIDENCIA: El uso de evidencia en la toma de decisiones con frecuencia incurrirá un costo de tiempo (al encontrar y procesar la evidencia) y puede llevar a decisiones que amenacen las formas de trabajo anteriormente establecidas, o por el contrario “dificulten la implementación”. Puede ser más sencillo para los responsables de toma de decisiones evitar la búsqueda o el uso de evidencia. Para abordar esto, las organizaciones deben considerar incentivos para el uso de la evidencia.

Cualquier norma futura y procesos de certificación asociados usados en el sector humanitario deben exigir a los organismos cumplir con las normas acordadas sobre la calidad de recopilación de evidencia.

Las organizaciones humanitarias deben aclarar los roles en la toma de decisiones y las expectativas de los responsables de toma de decisiones. Se debe aclarar quién es responsable para tomar las decisiones: aquellos responsables deben ser conscientes de las expectativas de las organizaciones sobre el uso de la evidencia.

Recuadro 6. ¿Qué hará ALNAP?

Se espera que todas las organizaciones miembros de ALNAP consideren e implementen estas recomendaciones. Mientras tanto, en representación de la Red, el Secretario General de ALNAP realizará las siguientes acciones:

- Trabajar con la Comunidad de Evaluación de Práctica para desarrollar y compartir guías prácticas sobre formas para mejorar la calidad probatoria de las evaluaciones humanitarias.
- Desarrollar materiales de capacitación para apoyar a los Miembros de ALNAP en la mejora de la calidad probatoria de evaluaciones, basadas en la guía *Evaluación de la Acción Humanitaria* (EHA) (Buchanan-Smith y Cosgrave, 2013).
- Seguir actualizando y perfeccionando la Biblioteca de Recursos: asegurando que contenga el mayor número posible de evaluaciones, documentos de lecciones e informes en acción humanitaria, además de que sea accesible y fácil de buscar.
- Continuar experimentando con una variedad de metodologías de investigación (que incluye: revisiones de literatura estructurada; enfoques de concordancia de patrones para el análisis del estudio del caso; métodos de encuesta cuantitativa) para asegurar que los informes de ALNAP se basen en las evidencias más contundentes posibles: publicar y difundir nuestras experiencias con el uso de estas metodologías.
- Trabajar con los Miembros de ALNAP (mediante nuestro trabajo en el liderazgo operacional) para entender mejor los procesos de toma de decisiones y rechazo al riesgo/incertidumbre, así como publicar recomendaciones concretas para mejorar la toma de decisiones basadas en una investigación contundente.

Anexo 1: Metodología

La versión inicial de este informe, Evidence and Knowledge in Humanitarian Action (Evidencia y Conocimiento en la Acción Humanitaria), se basó en una revisión de la literatura y se reforzó con algunas entrevistas. Se circuló como el documento antecedente para la 28a Reunión Anual de ALNAP y se presentó ante la reunión el 5 de marzo de 2013.

El informe actual se basa en el documento antecedente que se redactó nuevamente para incluir:

- Comentarios realizados por el panel y los participantes de la reunión en el mismo documento antecedente.
- Recomendaciones dadas por los participantes de la reunión en la sesión final de la reunión el 6 de marzo de 2013.
- Datos obtenidos de las 26 presentaciones realizadas en la reunión.

Las presentaciones incluidas en este informe se seleccionaron en base a un llamado abierto a los profesionales humanitarios y académicos, mediante la Membresía de ALNAP. Se solicitó a los presentadores debatir sobre “nuevos aprendizajes y el surgimiento de mejores prácticas en el entendimiento de la evidencia y en la recopilación, análisis y uso de la evidencia en acción humanitaria” (ALNAP, 2013: 1). En total, se seleccionaron 26 presentaciones de 110 propuestas. Todas las presentaciones se realizaron en inglés.

El presente documento incluye también resultados de una revisión adicional de la literatura, basada en temas que surgieron en la Reunión Anual. El documento fue revisado por un panel de expertos.

En este enfoque, las limitaciones claves están relacionadas al sesgo de selección. Las revisiones de literatura no se basaron en una búsqueda estructurada de bases de datos académicos, ni emplearon criterios formales de inclusión o exclusión (Hagen-zanker y Mallett, 2013).

Del mismo modo, la elección de las presentaciones incluidas en la Reunión Anual y referidas en el presente documento, si bien están menos influenciadas por un único autor, se sesgan por el canal de selección (mediante la Red de ALNAP) y por el hecho que todas las presentaciones se impartieron en inglés.

Anexo 2: Lista de entrevistados y colaboradores

Agradecemos la disposición de las siguientes personas para debatir temas en el presente informe, ya sea personalmente, vía telefónica o respondiendo preguntas mediante correo electrónico:

Miriam Ait Aissa, ACH

Ben Allen, ACH

Haley Bowcock, CaLP

Jeff Crisp, ACNUR

Andrew Catley, Universidad de Tufts

Mary Dixon-Woods, Universidad de Leicester

Merry Fitzpatrick, World Concern

Richard Garfield, CDC

Francois Grünewald, Grupo URD

Joyce Luma, PMA

John Mitchell, ALNAP

Lili Mohiddin, CaLP

Lars Peter Nissen, ACAPS

Jock Paul, UNOCHA

Chloe Puett, Universidad de Tufts

Anthony Redmond, Universidad de Manchester

Cecile Salpeteur, ACH

Louise Shaxson, ODI

Sharon Truelove, Cruz Roja Británica

Vivien Margaret Walden, Oxfam

Bibliografía

- Abebe, D. et al. (2008) 'Impact of a commercial destocking relief intervention in Moyale district, southern Ethiopia', in *Disasters* 32(2): 167–189. www.alnap.org/resource/8200.aspx
- ACAPS (2010) 'Rapid Initial Needs Assessment for Haiti (RINAH)'. Port-au-Prince: ACAPS. www.alnap.org/resource/9811.aspx
- ACAPS (2011) 'Joint Rapid Assessment of the Northern Governorates of Yemen'. Sanaa: ACAPS. www.alnap.org/resource/9812.aspx
- ACAPS (2012) 'Technical Brief: Estimation of affected population figures'. Geneva: ACAPS. www.alnap.org/resource/9813.aspx
- ACAPS (2012) 'Coordinated Assessments in emergencies. What we know now : Key lessons from field experience'. Geneva: ACAPS. www.alnap.org/resource/7912.aspx
- ACF (2011) 'Learning review'. London: ACF. www.alnap.org/resource/7001.aspx
- Aklilu, Y. and Wekesa, M. (2001) 'Livestock and Livelihoods in Emergencies: Lessons Learnt from the 1999–2001 Emergency Response in the Pastoral Sector in Kenya'. Nairobi: OUA IBAR. www.alnap.org/resource/9815.aspx
- Alexander, J. et al. (2013) '2013 Humanitarian accountability report'. Geneva: HAP International. www.alnap.org/resource/8758.aspx
- Ali et al. (2005) 'Cash Relief in a Contested Area: Lessons from Somalia: HPN Network paper no. 50'. London: HPN/ODI. www.alnap.org/resource/7585.aspx
- Ali, D. (2012) 'A Deadly Delay: Risk Aversion and Cash in the 2011 Somalia Famine,' in Bailey, S. and Ridsdel, B. (eds), *New learning in cash transfer programming [special feature] Humanitarian Exchange* 54. London: HPN/ODI. 10–15. www.alnap.org/resource/9817.aspx
- Allen, B. (2012) 'Collecting and Reporting best practices from the field'. *Msg 1, thread 2 under Capacity Area 2: Evaluation policy and purpose*, in *ALNAP Evaluation capacities – Community of Practice* (internal forum) London: ALNAP. (Accessed: 9 September 2012) www.alnap.org/resource/9818.aspx
- ALNAP (2003) 'Learning by Field Level Workers,' in *ALNAP review of Humanitarian Action in 2003*. London: ALNAP. www.alnap.org/resource/5208.aspx

- ALNAP (2010) 'The State of the Humanitarian System – Assessing Performance and Progress: a Pilot Study'. London: ALNAP/ODI. www.alnap.org/resource/8746.aspx
- ALNAP (2012) 'The State of the Humanitarian System 2012 edition'. London: ALNAP/ODI. www.alnap.org/sohsreport
- ALNAP (2013) 'Evidence and Knowledge in Humanitarian Action – Call for Presentations'. London: ALNAP/ODI. www.alnap.org/resource/7162.aspx
- Anderson et al. (2012) 'Time to listen: Hearing people on the receiving end of international aid'. Cambridge, Mass: CDA Collaborative Learning Projects. www.alnap.org/resource/8530.aspx
- Argren, R. (2013) 'What goes up must come down: the challenges of getting evidence back to the ground'. Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8015.aspx
- Ashdown, P. (2011) 'Humanitarian Emergency Response Review'. London: HERR. www.alnap.org/resource/6355.aspx
- Assessment Working Group for Northern Syria (2013) 'Joint Rapid Assessment of Northern Syria – Aleppo City Assessment'. Syria: Assessment Working Group for Northern Syria. www.alnap.org/resource/8755.aspx
- Avila et al. (2010) 'Global mapping of technology for transparency and accountability'. London: Open Society Foundation. www.alnap.org/resource/6421.aspx
- Banatvala, N. and Zwi, A. (2000) 'Public health and humanitarian interventions: developing the evidence base', in *The British Medical Journal* 321:101–5. www.alnap.org/resource/9820.aspx
- Beck, T. (2006) 'Evaluating humanitarian action using the OECD-DAC criteria – An ALNAP guide for humanitarian agencies'. London: ALNAP/ODI. www.alnap.org/resource/5253.aspx
- Bekele, G. and Abera, T. (2008) 'Livelihoods-based Drought Response in Ethiopia: Impact Assessment of Livestock Feed Supplementation', Pastoralist Livelihoods Initiative. www.alnap.org/resource/9936.aspx
- Bekele, G. (2010) 'Review of Save the Children US Livestock Marketing Initiative in Southern Ethiopia'. Westport: Save the Children USA. www.alnap.org/resource/9937.aspx
- Beynon et al. (2012) 'What difference does a policy brief make?'. Brighton: IDS and 3iE. www.alnap.org/resource/9938.aspx
- Bhattacharjee, A. and Lossio, R. (2011) 'Evaluation of OCHA Response to the Haiti

- Earthquake Final Report'. New York: UN OCHA.
www.alnap.org/resource/6002.aspx
- Bickman, L. and Reich, S. (2009) 'Randomised Controlled trials: A gold standard with feet of clay?', in Donaldson et al. (2009) *What counts as credible evidence in applied research and evaluation practice*. Thousand Oaks: Sage.
www.alnap.org/resource/10028.aspx
- Bond. (n.d.). 'The NGO Evidence Principles'. London: Bond.
www.alnap.org/resource/9941.aspx
- Bonino, F. with Jean, I. and Knox Clarke, P. (forthcoming) *Effective feedback mechanisms in humanitarian contexts - Guidance for practitioners and policy-makers based on findings from ALNAP-CDA case studies in Sudan, Pakistan and Haiti*. ALNAP Study. London: ALNAP/ODI. www.alnap.org/resource/9944.aspx
- Borton et al. (1996a) 'The International Response to Conflict and Genocide: Lessons from the Rwanda Experience Study 3: Humanitarian Aid and Effects'. London: ODI.
www.alnap.org/resource/9976.aspx
- Borton, J. et al. (1996b) 'The International Response to Conflict and Genocide: Lessons from the Rwanda Experience: Synthesis Report'. London: ODI.
www.alnap.org/resource/2517.aspx
- Bradt, D. A. (2009) 'Evidence-based decision-making in humanitarian assistance'. *HPN Network Paper 67*. London: HPN/ODI. www.alnap.org/resource/7503.aspx
- Broadbent, E. (2012) 'Politics of research-based policy in African policy debates: Synthesis of case study findings'. London: ebpdn.
www.alnap.org/resource/9942.aspx
- Buchanan-Smith, M. and Davies, S. (1995) 'Famine Early Warning Systems and Response: the Missing Link'. London: IT Publications.
www.alnap.org/resource/9943.aspx
- Buchanan-Smith, M. (2000) 'Role of Early Warning Systems in Decision-Making Processes'. London: HPN/ODI. www.alnap.org/resource/9946.aspx
- Buchanan-Smith, M. and Cosgrave, J. (2013) 'Evaluation of humanitarian action guide: Pilot guide'. London: ALNAP/ODI.
www.alnap.org/www.alnap.org/eha
- Buchanan-Smith, M. and Beck, T. (2008) 'Joint evaluations coming of age? The quality and future scope of joint evaluations', in *ALNAP Review of humanitarian action: Chapter three*. London: ALNAP/ODI. www.alnap.org/resource/5232.aspx
- Buchanan-Smith, M. and Scriven, K. (2011) 'Leadership in Action: Leading Effectively in Humanitarian Operations'. London: ALNAP/ODI.
www.alnap.org/resource/6118.aspx
- Buchanan-Smith, M. (2005) 'How the Sphere Project Came into Being: A Case Study

of Policy Making in the Humanitarian-aid Sector and the Relative Influence of Research' in Court et al. (eds) *Bridging Research and Policy in Development: Evidence and the Change Process*. London: ODI. www.alnap.org/resource/9947.aspx

Burns et al. (2008) 'Impact Assessment of the Pastoralist Survival and Recovery Project, Dakoro'. Niger: Lutheran World Federation and Feinstein International Center. www.alnap.org/resource/5722.aspx

CaLP (2013) 'Protecting Beneficiary Privacy: Principles and Operational Standards for the Secure Use of Personal Data in Cash and E-Transfer Programmes' Oxford: CaLP. www.alnap.org/resource/9750.aspx

Carr, H. (2013) 'Tracking Beneficiaries and Services Delivered'. Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. Washington: Catholic Relief Services. www.alnap.org/resource/8023.aspx

Catley, A. (2013) 'Developing international standards and guidelines with limited hard evidence'. Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. Washington: Tufts University. www.alnap.org/resource/8017.aspx

Catley et al. (n.d.) 'Participatory impact assessment: A guide for Practitioners'. Boston: Feinstein International Centre and Tufts University. www.alnap.org/resource/8094.aspx

Catley, A. and Cullis, A. (2012) 'Money to burn? Comparing the costs and benefits of drought responses in pastoralist areas of Ethiopia', in *The Journal of Humanitarian Assistance*. Ethiopia: FAO. www.alnap.org/resource/9948.aspx

Clarke, P. and Ramalingam, B. (2008) 'Organisational change in the humanitarian sector' in *ALNAP 7th Review of Humanitarian Action*. London: ALNAP/ODI. www.alnap.org/resource/5235

Cooley, A. and Ron, J. (2002) 'The NGO scramble: organizational insecurity and the political economy of transnational action', in *International Security* 27(1): 5–39. www.alnap.org/resource/9949.aspx

Cosgrave, J. (2007) 'Tsunami Evaluation Coalition, Synthesis Report: Expanded Summary, Joint evaluation of the international response to the Indian Ocean tsunami'. London: ALNAP and TEC/ODI. www.alnap.org/resource/5536.aspx

Coyle, D. and Meier, P. (2009) 'New Technologies in Emergencies and Conflicts: The Role of Information and Social Networks'. Washington: United Nations Foundation. www.alnap.org/resource/6572.aspx

Currion, P. (2010) 'Coordination at the Crossroads: NGO coordination in Southern Sudan 2007–2011'. London: ALNAP and TEC/ODI. www.alnap.org/resource/5781.aspx

- Damerell, J. (2013) 'The role of standards for evidence based humanitarian response', Presentation at *Evidence in Humanitarian Action, ALNAP 28th Annual Meeting*. Washington: ALNAP. www.alnap.org/resource/8016.aspx
- DARA (2011) 'The Humanitarian Response Index 2011'. Madrid: DARA. www.alnap.org/resource/6243.aspx
- Darcy, J. and Hofmann, C. (2003) 'According to need? Needs assessment and decision making in the humanitarian sector'. London: HPG/ODI. www.alnap.org/resource/9384.aspx
- Darcy, J. (2009) 'Humanitarian diagnostics: the use of information and analysis in crisis response decisions'. London: HPG/ODI. www.alnap.org/resource/9485.aspx
- Darcy, J. and Garfield, R. (2011) 'Review of the Rapid Initial Needs Assessment for Haiti', paper prepared for ACAPS. Geneva: Assessment Capacities Project. www.alnap.org/resource/9951.aspx
- Darcy et al. (2013) 'The Use of Evidence in Humanitarian Decision Making ACAPS Operational Learning Paper'. Boston: Feinstein International Center and Tufts University. www.alnap.org/resource/8003.aspx
- Development Initiatives (2012) 'Global Humanitarian Assistance: Tracking spending on cash transfer programming in a humanitarian context'. Wells: Development Initiatives. www.alnap.org/resource/9952.aspx
- de Ville de Goyet, C. and Moriniere, L. C. (2006) 'The Role of Needs of Assessment in the Tsunami Response'. London: ALNAP and TEC/ODI. www.alnap.org/resource/3531.aspx
- DFID (2012) 'Promoting innovation and evidence-based approaches to building resilience and responding to humanitarian crises: a DFID strategy paper'. London: DFID. www.alnap.org/resource/9823.aspx
- DFID (2013) 'How to note – Assessing the Strength of Evidence'. London: DFID. www.alnap.org/resource/8402.aspx
- DG ECHO (2009) 'The Use of Cash and Vouchers in Humanitarian Crises: DG ECHO funding guidelines'. Brussels: DG ECHO. www.alnap.org/resource/9824.aspx
- Dijkzeul, D. and Wakenge, C. (2010) 'Doing good, but looking bad? Local perceptions of two humanitarian organisations in eastern Democratic Republic of the Congo', in *Disasters* 34(4): 1139–70. www.alnap.org/resource/9953.aspx
- Dijkzeul et al. (2013) 'Introduction to Evidence based action in humanitarian crises'. [Manuscript submitted for publication], in *Disasters*, 37: S1–19. www.alnap.org/resource/9954.aspx
- ECOSOC (2007) 'UNICEF evaluation policy'. New York: UNICEF. www.alnap.org/resource/10037.aspx

- Epstein et al. (2004) 'Communicating evidence for participatory decision making' in *The Journal of the American Medical Association*, 291(19): 2359–66.
www.alnap.org/resource/9955.aspx
- Evans et al. (2011) 'Testing treatments: better research for better health care'. London: Pinter and Martin. www.alnap.org/resource/9956.aspx
- FAO and WFP (2013) 'Joint Evaluation of Food Security Cluster Coordination'. Rome: FAO/WFP. www.alnap.org/resource/9957.aspx
- Fearon et al. (2008) 'Community Driven Reconstruction in Lofa County: Impact assessment'. New York: International Rescue Committee.
www.alnap.org/resource/8192.aspx
- Featherstone, A. (2011) 'Strength in numbers: a global mapping review of NGO engagement in coordinated assessments'. ECB Project.
www.alnap.org/resource/9958.aspx
- Featherstone, A. (2013) 'Improving Impact: Do Accountability Mechanisms Deliver Results?'. London: Christian Aid, Save the Children for the Humanitarian Accountability Partnership. www.alnap.org/resource/8388.aspx
- Feinstein International Center (2007) 'Impact Assessments of Livelihoods-based Drought Interventions' in Moyale and Dire Woredas, Ethiopia: A Pastoralist Livelihoods Initiative report'. Boston: Feinstein International Center, Tufts University www.alnap.org/resource/3490.aspx
- Fenton et al. (eds) (2011) Humanitarian Accountability [special feature] *Humanitarian Exchange* 52. London: HPN/ODI. www.alnap.org/resource/6256.aspx
- Fetouh et al. (2013) 'Data quality in remote monitoring – a comparative analysis'. Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8027.aspx
- Few, R. et al. (2013) 'Contribution to Change: An approach to evaluating the role of intervention in disaster recovery'. Oxford: Oxfam.
www.alnap.org/resource/10305.aspx
- Foresti, M. (2007) 'A Comparative Study of Evaluation Policies and Practices in Development Agencies'. London: ODI. www.alnap.org/resource/7773.aspx
- Garfield, R. (2013) 'Evidence and needs assessments', Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/9959.aspx
- Garfield et al. (2011) 'Common Needs Assessments and Humanitarian Action'. *Humanitarian Network Paper* Number 69. London: HPN/ODI.
www.alnap.org/resource/6262.aspx
- Gerring, J. (2011) *Social Science Methodology: A Unified Framework*. 2nd ed. Cambridge: Cambridge University Press. www.alnap.org/resource/9015.aspx

- Gilman, D and Noyes, A. (2013) 'Humanitarianism in the Network Age'. New York: OCHA. www.alnap.org/resource/9157.aspx
- Goldacre, B. (2012) *Bad pharma: how drug companies mislead doctors and harm patients*. London: Fourth Estate. www.alnap.org/resource/9960.aspx
- Gottwald, M. (2010) 'Competing in the humanitarian marketplace: UNHCR's organizational culture and decision-making process'. Geneva: UNHCR Policy Development and Evaluation Service. www.alnap.org/resource/9961.aspx
- Grünewald, F. (2012) 'Sustaining learning from evaluation: stakeholder engagement and 'customer care'?'. Msg 1, thread 2 under Capacity Area 3: Evaluation processes and systems. Message posted in *ALNAP Evaluation capacities – Community of Practice* (internal forum). London: ALNAP. www.alnap.org/resource/9818.aspx
- Grünewald et al. (2010) 'Inter-agency real-time evaluation in Haiti: 3 months after the earthquake'. Plaisans: Groupe URD. www.alnap.org/resource/9962.aspx
- Guenther et al. (2010) 'The politics of evaluation: evidence-based policy or policy-based evidence?' Paper presented to *the NARU Public Seminar Series*, 30 November 2010, Darwin. www.alnap.org/resource/9963.aspx
- Guerrero, S. (2012) 'Our vision of linking evaluations and learning: ACF Learning Review 2011' *ALNAP blog*. 18th May. London: ALNAP. www.alnap.org/resource/9964.aspx
- Guerrero et al. (2013) '2012 Learning review'. London: ACF. www.alnap.org/resource/8219.aspx
- Hagen-zanker, J. and Mallett, R. (2013) 'How to do a rigorous, evidence-focused literature review in international development'. London: ODI. www.alnap.org/resource/8675.aspx
- Hallam, A. (2011) 'Harnessing the power of evaluation in humanitarian action'. London: ALNAP/ODI. www.alnap.org/resource/6123.aspx
- Hallam, A. and Bonino, F. (2013) *Using Evaluation for a Change: Insights from humanitarian practitioners*. ALNAP Study. London: ALNAP/ODI. www.alnap.org/resource/8980.aspx
- Hammersley, M. (1990) *Reading ethnographic research: A critical guide*. London: Longman. www.alnap.org/resource/9967.aspx
- Hammersley, M. (2005) 'Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice'. *Evidence Policy* 1(1): 85–100. www.alnap.org/resource/9968.aspx
- HAP (2010) 'The 2010 HAP Standard in Accountability and Quality Management'. Geneva: HAP International. www.alnap.org/resource/8125.aspx

- Harragin, S., and Chol, C. (1998) 'The southern Sudan vulnerability study'. Nairobi: Save the Children and South Sudan Programme.
www.alnap.org/resource/9969.aspx
- Hedlund, K. and Knox-Clarke, P. (2011) 'Humanitarian action in drought-related emergencies'. ALNAP Lessons Paper. London: ALNAP/ODI.
www.alnap.org/resource/6156.aspx
- Henry, G. T. and Mark, M. (2003) 'Beyond use: understanding evaluation's influence on attitudes and actions' in *American Journal of Evaluation* 24(3): 293–314.
www.alnap.org/resource/9970.aspx
- Hilhorst, D. (2003) 'Responding to disasters: diversity of bureaucrats, technocrats and local people' in *International Journal of Mass Emergencies and Disasters* 21(3): 37–55.
www.alnap.org/resource/9971.aspx
- Hillier, D. and B. Dempsey (2012) 'A Dangerous Delay: The cost of late response to early warnings in the 2011 drought in the Horn of Africa'. Oxford: Oxfam.
www.alnap.org/resource/6291.aspx
- House, S. (2012) 'Evaluation of Pakistan flood response 2011/12: using Oxfam GB's global humanitarian indicator tool'. Oxford: Oxfam.
www.alnap.org/resource/8366.aspx
- HPG (2006) 'Saving lives through livelihoods: critical gaps in the response to the drought in the Greater Horn of Africa: HPG Briefing Paper'. London: HPG/ODI.
www.alnap.org/resource/8656.aspx
- Humphreys, M. and Weinstein, J. M. (2009) 'Field experiments and the political economy of development' in *Annual Review of Political Science* 12(1): 367–78.
www.alnap.org/resource/9972.aspx
- Huybregts et al (2012) 'The effect of adding ready-to-use supplementary food to a general food distribution on child nutritional status and morbidity: a cluster-randomized controlled trial'. *PLoS medicine* 9(9).
www.alnap.org/resource/9973.aspx
- IASC (2011) 'Humanitarian System-Wide Emergency Activation: definition and procedures'. Geneva: IASC. www.alnap.org/resource/8502.aspx
- IASC (2012a) 'Multi Cluster/Sector Initial Rapid Assessment (MIRA) Provisional version'. Geneva: IASC. www.alnap.org/resource/9974.aspx
- IASC (2012b) 'Operational Guidance for Coordinated Assessments in Humanitarian Crises – Provisional Version'. Geneva: IASC. www.alnap.org/resource/9975.aspx
- ICRC and IFRC (2008) 'Guidelines for assessment in emergencies'. Geneva: ICRC/IFRC. www.alnap.org/resource/6404.aspx
- IFRC (2011) 'Project/programme monitoring and evaluation (M&E) guide'. Geneva: IFRC. www.alnap.org/resource/8542.aspx

- IFRC (2013) 'World disasters report.' Geneva: IFRC.
www.alnap.org/resource/10054.aspx
- Jacobsen, K. and Furst Nichols, S. (2011) 'Developing a Profiling Methodology for Displaced People in Urban Areas'. Boston: Feinstein International Center and Tufts University. www.alnap.org/resource/6306.aspx
- Jaspars, S. (2006) 'From Food Crisis to Fair Trade: Livelihoods analysis, protection and support in Emergencies'. Series no 3. Oxford: Oxfam.
www.alnap.org/resource/9842.aspx
- Johnson et al. (2009) 'Research on evaluation use: a review of the empirical literature from 1986 to 2005' in *American Journal of Evaluation* 30(3): 377–410.
www.alnap.org/resource/9977.aspx
- Jones, H. (2012) 'Promoting evidence-based decision-making in development agencies'. London: ODI. www.alnap.org/resource/9978.aspx
- Jones, H. and Mendizabal, E. (2010) 'Strengthening learning from research and evaluation: going with the grain'. London: ODI.
www.alnap.org/resource/8429.aspx
- Kitson, A. (2002) 'Recognising relationships: reflections on evidence-based practice' in *Nursing Inquiry* 9(3): 179–86. www.alnap.org/resource/9979.aspx
- Knox Clarke, P. (2013a) 'Measuring the effectiveness of humanitarian action'. Presentation at *OCHA roundtable on Humanitarian Effectiveness*, March 2013.
www.alnap.org/resource/9980.aspx
- Knox Clarke, P. (2013b) 'Who's in Charge Here?' A literature review of approaches to leadership in humanitarian operations'. London: ALNAP/ODI.
www.alnap.org/resource/8640.aspx
- Laybourne, C. and Obrecht, A. (2013) 'Show me the money: generating evidence from and for accountability in humanitarian work', Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8018.aspx
- LeCompte, M. D. and Goetz, J. P. (1982) 'Problems of Reliability and Validity in Ethnographic Research'. *Review of Educational Research* 52(1): 31–60.
www.alnap.org/resource/9981.aspx
- Leeuw, Frans L. (2012) 'Theory Based Evaluation'.
www.alnap.org/resource/9982.aspx
- Levine et al. (eds) (2011) 'System failure ? Revisiting the problems of timely response to crises in the Horn of Africa: HPN Network Paper'. London: HPN/ODI.
www.alnap.org/resource/6260.aspx
- Macrae, J. (2013) 'Towards an evidence agenda in humanitarian action: reflections

- on the role of donors in investing in research and innovation in the sector', Presented at, *Evidence in Humanitarian Response: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/9843.aspx
- Mahmood et al. (2010) 'Guidelines on Data Issues in Humanitarian Crisis Situations'. New York: UNFPA. www.alnap.org/resource/9983.aspx
- Majid, N. (2011) 'Early Warning – Delayed Response?'. Geneva: International Federation of the Red Cross. www.alnap.org/resource/6275.aspx
- Maxwell, D. and Watkins, B. (2003) 'Humanitarian Information Systems and Emergencies in the Greater Horn of Africa: Logical Components and Logical Linkages' in *Disasters* 27(1): 72–90. www.alnap.org/resource/9984.aspx
- Maxwell et al. (2012) 'What Drives Program Choice in Food Security Crises? Examining the Response Analysis Question: Special Edition on Impacts of Innovative Food Assistance Instruments. Boston: Feinstein international. www.alnap.org/resource/9985.aspx
- Maxwell et al. (2013) 'HPN Network Paper: Response analysis and response choice in food security crises: a roadmap', 44(73). London: HPN/ODI. www.alnap.org/resource/9986.aspx
- Mays et al. (2005) 'Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field' in *Journal of Health Services Research & Policy* 10 (Suppl 1): 6–20. www.alnap.org/resource/9987.aspx
- Mazurana et al. (2011) 'Sex and age matter: improving humanitarian response in emergencies'. Boston: Feinstein International Center, Tufts University. www.alnap.org/resource/8144.aspx
- Mills, E. J. (2005) 'Sharing evidence on humanitarian relief' in *British Medical Journal* 331: 1485–86. www.alnap.org/resource/9988.aspx
- Minear, L. (2002) *The Humanitarian Enterprise: Dilemmas and Discoveries*. Bloomfield: Kumarian Press. www.alnap.org/resource/9994.aspx
- Mitchell, G. J. (1999) 'Evidence-based practice: critique and alternative view' in *Nursing Science Quarterly* 12(1): 30–35. www.alnap.org/resource/10000.aspx
- Morra Imas, L. and Rist, R. (2009) 'The Road to Results: designing and conducting effective development evaluations'. Washington: The International Bank for Reconstruction and Development and World Bank. www.alnap.org/resource/8470.aspx
- MSF Vienna Evaluation Unit (2012) 'Evaluation Manual: A handbook for Initiating, Managing and Conducting Evaluations in MSF'. Vienna : MSF. www.alnap.org/resource/10004.aspx
- Nicholson, N. and Desta, S. (2010) 'Evaluation of the Enhanced Livelihoods in Mendera Triangle and Southern Ethiopia 2007–2009'. Washington: USAID and

- ELMT/ELSE. www.alnap.org/resource/10009.aspx
- OCHA (2007) 'Management response matrix to the Intermediate Review of the Central Emergency Response Fund'. New York: OCHA/CERF. www.alnap.org/resource/10012.aspx
- OCHA (2012) 'Evaluation Initiale Rapide Multi-Cluster sur les inondations au Moyen Chari, Tandjilé, Mayo Kebbi Est – Tchad'. N'Djema: OCHA. www.alnap.org/resource/10013.aspx
- OCHA (2013) 'Meeting Summary: Consultative Workshop Humanitarian Effectiveness'. New York: OCHA. www.alnap.org/resource/10014.aspx
- OECD-DAC (2002) 'Glossary of key terms in evaluation and results based management'. Paris: OECD/DAC. www.alnap.org/resource/8489.aspx
- Oxfam (n.d.-a) 'How are effectiveness reviews carried out?'. Oxford: Oxfam. www.alnap.org/resource/8473.aspx
- Oxfam (n.d.-b) 'Monitoring, evaluation, accountability and learning'. Oxford: Oxfam. www.alnap.org/resource/10015.aspx
- Oxfam (n.d.-c) 'Oxfam GB evaluation guidelines'. Oxford: Oxfam. www.alnap.org/resource/10016.aspx
- Oxfam (n.d.-d) 'Rough guide to monitoring and evaluation in Oxfam GB'. Oxford: Oxfam. www.alnap.org/resource/10018.aspx
- Oxfam (2004) 'Evaluation of Oxfam GB's Food Aid and Food Security Emergency Intervention in Mauritania'. Oxford: Oxfam and Acacia Consultants Ltd. www.alnap.org/resource/10017.aspx
- Pantuliano, S. and Wekesa, M. (2008) 'Improving Drought Response in Pastoral Areas of Ethiopia, Somali and Afar Regions and Borena Zone of Oromiya Region'. London: HPG/ODI. www.alnap.org/resource/7510.aspx
- Parker, J. and Sanderson, D. (2013) 'Reviewing the quality of evidence in humanitarian evaluations', Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8019.aspx
- Peppiatt et al. (2000) *Buying Power: The Use of Cash Transfers in Emergencies*. London: The British Red Cross. www.alnap.org/resource/10019.aspx
- Piccio, R. (2012) 'Probing the paradox of the RCT craze in international development', *NGO Performance*, 24 May. www.alnap.org/resource/10020.aspx
- Poole, L. and Primrose, J. (2010) 'Southern Sudan: Funding according to need'. Wells: Development Initiatives. www.alnap.org/resource/10021.aspx
- Proudlock et al. (2009) 'Improving humanitarian impact assessment: Bridging theory and practice' in *ALNAP Review of humanitarian action in 2009*. London:

- ALNAP/ODI. www.alnap.org/resource/5663.aspx
- Puett, C. and Salpeteur, C. (2013) 'Cost-effectiveness of preventing child morbidity with ready-to-use food in urban Chad'. Presented at, *Evidence and Knowledge in Humanitarian Response: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8009.aspx
- Ramalingam et al. (2009a) 'Innovations in international humanitarian action' in *ALNAP Review of humanitarian action in 2009*. London: ALNAP/ODI. www.alnap.org/resource/5664.aspx
- Ramalingam et al. (2009b) 'Counting what counts : performance and effectiveness in the humanitarian sector'. in *ALNAP Review of humanitarian action in 2009*. www.alnap.org/resource/5666.aspx
- Ravallion, M. (2009) 'Evaluation in the practice of development,' in *The World Bank Research Observer* 24(1): 29–53. www.alnap.org/resource/10022.aspx
- Ravallion, M. (2011) 'Knowledgeable bankers? The demand for research in World Bank operations'. World Bank Research Working Paper No.5892. Washington: World Bank. www.alnap.org/resource/10023.aspx
- Redmond et al. (2010) 'A Qualitative and Quantitative Study of the Surgical and Rehabilitation Response to the Earthquake in Haiti' in *Prehospital and Disaster Medicine* 26(06): 449–456 www.alnap.org/resource/7446.aspx
- Robson et al. (2001) 'Guide to evaluating the effectiveness of strategies for preventing work injuries: how to show whether a safety intervention really works'. Cincinnati: CDC and NIOSH. www.alnap.org/resource/10024.aspx
- Rogers, P. J. (2009) 'Learning from the Evidence about Evidence-Based Policy. Strengthening Evidence-Based Policy in the Australian Federation', in *Vol. 1: Roundtable Proceedings*. Canberra: Australian Government Productivity Commission. www.alnap.org/resource/10025.aspx
- Sackett et al. (1996) 'Evidence based Medicine: What it is and What it isn't' in *British Medical Journal* 312(5): 71–2. www.alnap.org/resource/10026.aspx
- Sadler, K. and Catley, A. (eds). (2009) 'Milk Matters: A literature review of pastoralist nutrition and programming responses'. Boston: Feinstein International Center, Tufts University and Save the Children. www.alnap.org/resource/10027.aspx
- Sanderson et al. (2012) 'Responding to urban disasters'. ALNAP Lessons Paper. London: ALNAP/ODI. www.alnap.org/resource/7772.aspx
- Sandison, P. (2006) 'The utilisation of evaluations' in *ALNAP Review of humanitarian action 2005*. London: ALNAP/ODI. www.alnap.org/resource/5225.aspx
- Schwandt, T. (2009) 'Towards a practical theory of evidence for evaluation', in Donaldson et al. (2009) *What counts as credible evidence in applied research and*

- evaluation practice*. Thousand Oaks: Sage. www.alnap.org/resource/10028.aspx
- Scriven, M. (2009) 'Demythologising Causation and Evidence', in Donaldson et al. (2009) *What counts as credible evidence in applied research and evaluation practice*. Thousand Oaks: Sage. www.alnap.org/resource/10028.aspx
- Seaman et al. (2000) 'The Household Economy Approach: a resource manual for practitioners'. London: Save the Children. www.alnap.org/resource/10029.aspx
- Segone, M. (2009) 'Enhancing evidence-based policy-making through country-led monitoring and evaluation systems' in Segone, M. (ed.) *Country-led monitoring and evaluation systems*. Geneva: UNICEF. www.alnap.org/resource/8567.aspx
- Shaxson, L. (2005) 'Is your evidence robust enough? Questions for policy makers and practitioners' in *Evidence & Policy: A Journal of Research, Debate and Practice* 1(1): 101–12. www.alnap.org/resource/10030.aspx
- Shaxson, L. (2012) 'Expanding our understanding of K*(KT, KE, KTT, KMb, KB, KM, etc.)'. A concept paper emerging from the K* (Kstar) Initiative conference, 24–27 April, Hamilton, Canada. www.alnap.org/resource/10031.aspx
- Silverman, W.A. and Sackett, D. L. (1999) *Where's the evidence? Debates in modern science*. Oxford: Oxford University Press. www.alnap.org/resource/10032.aspx
- Slim, H. (2012) 'IASC Real-Time Evaluation of the Humanitarian Response to the Horn of Africa Drought Crisis Synthesis Report'. New York: IASC. www.alnap.org/resource/7761.aspx
- Smith et al. (2011) 'New Technologies in Cash Transfer Programming and Humanitarian Assistance' Oxford: CaLP. www.alnap.org/resource/9634.aspx
- Spence et al. (2003) 'Quality in qualitative evaluation: a framework for assessing research evidence'. London: Government Chief Social Researcher's Office. www.alnap.org/resource/10033.aspx
- Steen et al. (2013) 'Synthesis of Mixed Method Impact Evaluations of the Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations'. Rome: WFP. www.alnap.org/resource/7997.aspx
- Stern et al. (2012) 'Broadening the range of designs and methods for evaluationst'. DFID Working Paper 38. London: DFID. www.alnap.org/resource/8196.aspx
- Tache, B. (2010) 'Participatory Impacts Assessment of Drought Reserve Areas in Guji and Borana Zones', Oromia Region: Save the Children USA. www.alnap.org/resource/9940.aspx
- Telford, J. and J. Cosgrave (2006) 'Joint evaluation of the international response to the Indian Ocean tsunami: Synthesis report'. ALNAP and TEC/ODI. www.alnap.org/resource/3535.aspx

- Tessitore, S. (2013) 'Like a good trip to town without selling your animals: A study of FAO Somalia's Cash for Work programme'. Rome: FAO.
www.alnap.org/resource/8821.aspx
- UNHCR (2010) 'UNHCR's evaluation policy'. Geneva: UNHCR.
www.alnap.org/resource/10034.aspx
- UNICEF (2004) 'UNICEF evaluation report standards'. New York: UNICEF.
www.alnap.org/resource/10035.aspx
- UNICEF (2007) 'Programme policy and procedure manual: programme operations'. New York: UNICEF. www.alnap.org/resource/10036.aspx
- UNOCHA and Pakistan NDMA (2012) 'Multi-sector Initial Rapid Assessment (MIRA) Report Pakistan Floods 2012'. Islamabad: Assessment Working Group Pakistan.
www.alnap.org/resource/10038.aspx
- Valid International (2012) 'IASC Real Time Evaluation of the response to the Horn of Africa Drought: Somalia'. Oxford: Valid International.
www.alnap.org/resource/7506.aspx
- Van de Putte, B. (2000) 'The Utilisation of Evaluation Recommendations in Medecins Sans Frontieres – Holland: A Study of 10 Evaluation Reports (1997–1999)'. Amsterdam: MSF-H. www.alnap.org/resource/10040.aspx
- Van de Putte, B. (2001) 'Follow-up to Evaluations of Humanitarian Programmes'. London: ALNAP/ODI. www.alnap.org/resource/5011.aspx
- Venton et al. (2012) 'The Economics of Early Response and Disaster Resilience: Lessons from Kenya and Ethiopia: Economics of Resilience Final Report'. London: DFID and UKaid. www.alnap.org/resource/7758.aspx
- VSF (2009) 'Meat and Milk Voucher project (IMPACT I and II)', Clarke and Fison. Bahr al Gazal, South Sudan: VFS. www.alnap.org/resource/10042.aspx
- Walden, V. (2013) 'Improving quality of humanitarian programmes through the use of a scoring system: the Humanitarian Indicator Tool', Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. Oxford: Oxfam. www.alnap.org/resource/8024.aspx
- Walker, P. (2013) 'Cracks in the machine: is the humanitarian system fit for purpose?', Presented at, *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8014.aspx
- Willet, B. (2013) 'Mission metrics: An agency wide measurement approach for understanding emergency response', Presented at *Evidence and Knowledge in Humanitarian Action: ALNAP 28th Annual Meeting*, Washington, 5–7 March. www.alnap.org/resource/8025.aspx
- WFP (n.d.) 'Monitoring and Evaluation Guidelines. Module: How to design a Results-Oriented M&E Strategy for EMOPs and PRROs'. Rome: WFP.

www.alnap.org/resource/10045.aspx

WFP (2005) 'Summary report on WFP follow-up to recommendations'. Rome: WFP.
www.alnap.org/resource/10048.aspx

World Vision Pakistan (2011) 'End of Programme Evaluation Report for DEC-funded WV Relief Program in Sindh (July 2011 – Phase 1)'. Islamabad: Sustainable Solutions Ltd. www.alnap.org/resource/6148.aspx

Young, J. and Court, J. (2004) 'Bridging Research and Policy in International Development: An Analytical and Practical Framework'. London: ODI.
www.alnap.org/resource/8430.aspx

Zhang et al. (2002) 'A knowledge management framework for the support of decision making in humanitarian assistance/disaster relief' in *Knowledge and Information Systems* 4(3): 370–85. www.alnap.org/resource/10052.aspx



ALNAP

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom

Tél : +44 (0)20 7922 0388
Fax : +44 (0)20 7922 0399
Site Web : www.alnap.org
Email : alnap@alnap.org

ALNAP would like to acknowledge the financial support of USAID in carrying out this initiative..

