



LA INTEGRACIÓN DE LA EVALUACIÓN EN EL CICLO DE LAS INTERVENCIONES DE LAS ONGD

¿SON ÚTILES PARA LAS ONGD LAS EVALUACIONES REALIZADAS EN EL MARCO DE LOS CONVENIOS DE FINANCIACIÓN CON AECID?

José María Larrú¹

María Méndez²

¹ Investigador del CECOD y profesor de la Universidad CEU San Pablo: larram@ceu.es

² Consultora y evaluadora independiente: m.mendezgq@gmail.com

INDICE

CAPÍTULO I: INTRODUCCIÓN. ACOTACIÓN DEL OBJETO DE EVALUACIÓN

CAPÍTULO II: MARCO TEÓRICO

CAPITULO III: LOGICA DE INTERVENCIÓN Y METODOLOGÍA DE EVALUACIÓN

CAPITULO IV: ANÁLISIS DE LA INFORMACIÓN

CAPITULO V: INTERPRETACIÓN Y ENJUICIAMIENTO

CAPITULO VI: RECOMENDACIONES

CAPITULO VII: PLAN DE COMUNICACIÓN

CAPITULO VIII: LÍMITES Y CONDICIONANTES DE LA EVALUACIÓN

BIBLIOGRAFÍA

ANEXOS

ANEXO I: MATRIZ DE EVALUACIÓN

ANEXO II: HERRAMIENTAS DE RECOGIDA DE INFORMACIÓN

ANEXO III: ESTADÍSTICOS DESCRIPTIVO

ANEXO IV: PRESENTACION PPT

ACRÓNIMOS

AECID: Agencia Española de Cooperación Internacional para el Desarrollo

CAD de la OCDE: Comité de Ayuda al Desarrollo de la Organización para la Cooperación y el Desarrollo Económico.

CECOD: Centro de Estudios de Cooperación al Desarrollo

CONGDE: Coordinadora de ONG para el Desarrollo de España.

CM: Comunidad de Madrid.

DGPOLDE: Dirección General de Planificación y Evaluación de Políticas de Desarrollo.

FONGDCAM: Federación de Organizaciones No Gubernamentales de Desarrollo de la Comunidad de Madrid.

ONGD: Organización No Gubernamental de Desarrollo.

OTC: Oficina Técnica de Cooperación

SL: Socio Local

TdR: Términos de Referencia

UE: Unión Europea

I- INTRODUCCIÓN

La finalidad de esta evaluación es conocer si las ONGD españolas integran la evaluación en el ciclo de sus intervenciones. Esta evaluación responde a un encargo realizado por el CECOD bajo el título: “**La integración de la evaluación en el ciclo del proyecto**”, que se enmarca dentro del programa cofinanciado por la Comunidad de Madrid: “Fortalecimiento y desarrollo de las capacidades de los agentes de la cooperación madrileña: fortalecimiento institucional y evaluación”, en la convocatoria de 2010.

El CECOD acotó la dimensión temporal e institucional del estudio en el propio encargo, ya que determinó que la evaluación debería hacerse sobre las ONGD. Para seleccionar la muestra a estudiar se tomó en cuenta que muchas ONGD que habían ejecutado intervenciones subvencionadas por la AECID a través de la modalidad de Convenios de Cooperación en las convocatorias habidas en los años 2006 y 2007, tenían la obligación de realizar evaluaciones externas intermedias. Este dato aseguraba tratar con Organizaciones que habían completado un ciclo entero de evaluación, desde los Términos de Referencia, hasta la retroalimentación y posible difusión de los informes y planes de mejora.

Por tanto, **el Objetivo de esta evaluación era** identificar si las ONGD que debían haber evaluado sus convenios con AECID hasta la fecha, habían realizado la obligada evaluación, cómo la habían realizado y, sobre todo si ésta había sido útil y ello se plasmaba en si habían incorporado las recomendaciones generadas en el plan de mejora dentro del ciclo de las intervenciones (proyectos, programas o nuevos convenios) de las ONGD.

Complementariamente al mencionado principal objetivo de la evaluación se pretendía comprender en profundidad la política de evaluación llevada a cabo por las ONGD, para diagnosticar cuáles habían sido los cuellos de botella y fortalezas en dicho proceso, con el fin de poder establecer algunas recomendaciones de mejora a la política de gestión de las evaluaciones de los Convenios.

El estudio, por tanto, se ha centrado en la muestra de las 37 ONGD que debían haber realizado las evaluaciones intermedias obligatorias en el marco de los Convenios con AECID

en las convocatorias de 2006 (25 ONGD con convenio) y 2007 (además de algunas del 2006, 12 nuevas tuvieron convenio en 2007).

II.- MARCO TEÓRICO

Propósito de la Evaluación. La evaluación surge, ante el interés del CECOD, por conocer en qué grado, la etapa de evaluación se integra en el ciclo del proyecto en las ONGD. Por lo tanto, con este trabajo se busca juzgar, poner en valor, la utilidad que dichas evaluaciones han reportado a las ONGD. Utilidad, tanto para la mejora del Convenio evaluado, como para la generación de cultura evaluativa en la propia ONGD. Complementariamente al primer objetivo de la evaluación, se pretende comprender en profundidad, la política de evaluación llevada a cabo por las ONGD, para diagnosticar, cuáles son los cuellos de botella y fortalezas en dicho proceso, con el fin de poder establecer, algunas recomendaciones de mejora a la política de gestión de las evaluaciones de los Convenios.

La **función principal de esta evaluación** sería: Producir conocimiento para la posible mejora del sistema. En este sentido, se evaluarán los principales cuellos de botella, dificultades, así como fortalezas del sistema, con el fin de iluminar posibles evaluaciones futuras. Igualmente, se tratará de concluir, con algunas recomendaciones o sugerencias, que de ser aplicadas, podrían hacer que el sistema funcionase mejor, destacando también, aquellas que se han identificado como fortalezas, en algunos de los casos estudiados, con el fin de iluminar otros casos (ONGD)

Aunque esta evaluación no pretende rendir cuentas, sino que su principal función, es **comprender** el proceso de evaluación, para **iluminar**, para poder facilitar conocimiento que pueda mejorar el sistema. La evaluación aportará información, sobre el grado de utilidad que ha resultado de las evaluaciones realizadas, lo que de alguna manera, es información que podría ser utilizada, a modo de **rendición de cuentas**, en cuanto a la utilidad de los recursos y esfuerzos invertidos, para realizar dichas evaluaciones (todas ellas realizadas con financiación pública proveniente de AECID).

Siguiendo a Stake, podríamos decir, que esta evaluación no pone su foco de atención en la medición de resultados del sistema evaluado, sino que se centra sobre todo, en la **comprensión profunda** de la naturaleza del evaluando, *en concreto en sus logros y puntos fuertes, así como en sus limitaciones y cuellos de botella*. Es decir, se busca la calidad en

sentido holístico (Stake, 2006, “Evaluación comprensiva y evaluación basada en estándares”. pág 85). Información, que pueda ser utilizada por quien la reciba, para la toma de decisiones o simplemente para la autorreflexión.

Por tanto, esta evaluación se tratará de hacer de manera “comprensiva-respondiente”, que más que un modelo, responde a una actitud, lo que supone basarse en la experiencia y estar orientada a la comprensión, entendimiento del sistema objeto de evaluación. Su enfoque principal, no será por tanto, una evaluación basada en estándares, orientada a la medición, el enfoque predominante será el “comprensivo”, pues se tratará de atender a los intereses y preocupaciones de los principales agentes implicados, las ONGD. Por ello, la evaluación se hará, siguiendo un **modelo sistémico, en base a la teoría del programa**, ya que como se recoge en Ligeró, 2011, “Dos Métodos de Evaluación: Criterios y Teoría del Programa” Documento de trabajo N° 15/2011 del CECOD (Pág. 44), “La evaluación de la teoría del programa, tiene como finalidad, comprender las cadenas o secuencias causales, que provocan los resultados. Su intención es saber, por qué han ocurrido las cosas. Cuando se sabe el porqué, es cuando se puede mejorar, proponiendo recomendaciones enfocadas y concretas, que atañen a la totalidad de la intervención”.

Por lo tanto, en base a la teoría del programa, se tratará de analizar, tanto el logro de los resultados, como el desarrollo de los procesos, para ver cómo han contribuido al logro de dichos resultados. Analizando los procesos en profundidad y cómo estos se han dado, podremos obtener mucha más información, para detectar los cuellos de botella y fortalezas del sistema de evaluación. Igualmente, analizando los principales elementos de la estructura, podremos determinar, si estos eran suficientes, adecuados, limitantes, etc. En cualquier caso, el énfasis de esta evaluación, radicará especialmente en los procesos, pues la principal finalidad del estudio, es saber las causas del grado de utilidad, que han reportado para las ONGD, las evaluaciones realizadas.

En cuanto a nuestra **perspectiva paradigmática de investigación**, nos basamos en un **enfoque constructivista**, (ni en el positivismo puro científico, ni en el participativo, co-creación, basado en la acción-reflexión). La evaluación comprensiva, no es considerada participativa, tal y como es entendida, en la definición de “evaluación participativa”, como un paso, que va más allá de la “evaluación democrática” (Stake). Sin embargo, sí se generarán, distintos momentos de **participación** del CECOD, como principal cliente, y de las propias

ONGD, enfocados a la determinación de necesidades informativas para el estudio, y a la revisión y discusión de los avances, en el proceso de evaluación, llevado a cabo por los evaluadores externos. De acuerdo, al encargo, se trataría de una **evaluación externa**, buscando en todo el proceso, la máxima colaboración de los agentes implicados, (CECOD, AECID, ONGD) y **expost ipso**, porque se realiza una vez finalizado el proceso de evaluación de los Convenios con AECID.

La evaluación de la calidad holística, precisa de una gran cantidad de datos, tanto cualitativos como cuantitativos. Son importantes, tanto las mediciones basadas en criterios, como las percepciones subjetivas (Stake, pág.87), por ello, se buscará la **triangulación de información**, mediante recogida de la misma, con **herramientas cuantitativas y cualitativas**. Por último, en cuanto a la interpretación de los resultados, es importante destacar, que en la evaluación comprensiva, el valor final suele establecerse, tras sucesivas depuraciones del juicio personal, para ello en este caso, sumaremos la idea de "consenso de interpretaciones", al disponer de dos investigadores, que deberán consensuar sobre las interpretaciones hechas al estudio. La triangulación citada, permitirá distinguir, entre lo que son interpretaciones fiables y simplemente opiniones particulares del evaluador.

III.- LÓGICA DE INTERVENCIÓN Y METODOLOGÍA DE LA EVALUACIÓN

En este capítulo, se describen las fases, que se siguieron para el diseño de la evaluación.

Como se ha explicado, el objetivo de la evaluación, era dar respuesta, al encargo del CECOD reflejado en la pregunta: ¿Las ONGD integran la evaluación dentro del ciclo del proyecto?

Para dar respuesta a esta pregunta, el equipo evaluador optó por reconstruir la teoría del programa, que subyacía a la política de evaluación seguida por las ONGD, (siempre en el marco de los Convenios con AECID), y analizar cómo sus diferentes componentes (estructura y procesos), así como las relaciones entre éstos, influían en mayor o menor grado, en la cadena de resultados, cuyo último fin, era precisamente, la pregunta hecha por el CECOD: ¿la política de evaluación, a través de todos sus componentes, hace que las ONGD integren la etapa de evaluación dentro del ciclo del proyecto?.

La construcción de la teoría del programa, se hizo según figura en el final de este apartado. Se identificaron los elementos estructurales, que podrían estar influyendo, en el grado de utilidad de las evaluaciones. Se identificaron los subprocesos, que constituían el proceso de evaluación, llevado a cabo por las ONGD. En la dimensión de resultados, se optó por fijar de antemano solo uno: el grado de utilidad, que había reportado para las ONGD, el proceso de evaluación, y tratar de indagar sobre todos los posibles efectos, que había producido el proceso de evaluación.

Una vez que el CECOD había expuesto su necesidad informativa, el equipo evaluador procedió a desagregar con detalle, todas las necesidades informativas subyacentes a la pregunta clave. Para ello, se basó en el modelo orientado por la teoría. Resultado de ello, son las preguntas que se reflejan en la matriz de evaluación, ordenada según las tres dimensiones de la teoría del programa: estructura, procesos y resultados. (Anexo I). En dicha matriz, se identificó quién(es) sería(n) el/los informante(s) clave, así como cuál sería, la técnica de recogida de información.

Para dar respuesta a estas preguntas, no se vio conveniente optar por una metodología experimental ,ni cuasi-experimental de la evaluación, pues no se pretendía la atribución de

resultados, sino contestar a dos preguntas, sobre la utilidad y la incorporación de la evaluación en el ciclo del proyecto, ambas íntimamente ligadas. Por ello, para contestar a estas dos preguntas, se optó por un diseño de un modelo no experimental (modelización estadística), en donde se analizó la correlación (vía análisis de regresión lineal), entre la utilidad de la evaluación, y algunas variables independientes, que podían influir en el grado de utilidad de las evaluaciones. Estas variables independientes serían: la calidad profesional de los evaluadores externos, la participación de las ONGD, y/o de sus socios locales, en todo el proceso, las condiciones determinadas por la propia normativa de AECID, para el desarrollo de estas evaluaciones (obligatoriamente externa e intermedia), y los recursos (tiempo, dinero y personal), empleados para la evaluación.

Como se ha explicado, no se trataba de establecer relaciones de atribución, pero sí al menos, de analizar la influencia, que estas variables tenían, en los diferentes procesos mencionados para el logro o no de los resultados.

Una vez analizado el grado de utilidad logrado, y la influencia que dichas variables podrían tener en ello, interesaba profundizar en el porqué de dicha valoración, conocer los cuellos de botella y aspectos clave que se habían dado en el proceso de evaluación. Para ello, se utilizaron herramientas de recogida de información, de carácter cualitativo, para profundizar en el conocimiento de las debilidades y fortalezas, que se habían dado en los procesos de evaluación realizados.

Una vez definidas las preguntas de evaluación y la metodología no experimental por la que se había optado, se procedió al diseño y elaboración de las técnicas de recogida de información. En este aspecto se optó por un modelo mixto, que combinara tanto herramientas cuantitativas como cualitativas con el fin de poder triangular la información obtenida por las diversas herramientas utilizadas y sobre todo para poder tener una comprensión holística de los procesos de evaluación llevados a cabo por las ONGD.

Igualmente se establecieron entrevistas con el personal experto en evaluación en la AECID y de la DGPOLDE para recoger información sobre su punto de vista en cuanto a la utilidad del sistema en cuestión.

Los pasos seguidos, por tanto para el diseño de las herramientas de recogida de información fueron. En primer lugar, se elaboró un cuestionario al que se invitó a responder a las 37 ONGD que se identificaron como universo poblacional con convenios 2006-2007. El cuestionario combinaba aspectos cuantitativos y cualitativos. Para obtener información más detallada, se seleccionaron de forma aleatoria 10 ONGD que fueron objeto de entrevistas personales en profundidad a partir del cuestionario. La selección estas ONGD se hizo mediante un proceso aleatorio. Ordenadas alfabéticamente las organizaciones, se generaron números aleatorios en un intervalo (0-1) y las diez organizaciones más cercanas a 0,5 fueron las seleccionadas. Este criterio aleatorio elimina sesgos de selección de entidades por tamaño, orientación ideológica, año de ejecución del convenio u otra característica elegida de forma discrecional por los investigadores.

La combinación de los datos cuantitativos (con preguntas cerradas en escalas dicotómicas y ordinales entre 1 y 4 evitando la tendencia central), así como las preguntas abiertas que explican algunas de las respuestas, permiten a las entrevistas ser una herramienta muy válida que evita el positivismo extremo y la pura opinión subjetiva recogida en una entrevista informal. Además, los cuestionarios fueron previamente validados con miembros de ONGD no pertenecientes al universo (sin convenio) lo que permitió depurar y aclarar mucho las preguntas de forma que se limitaran al máximo las preguntas de juicio de valor, de doble sentido o tendenciosas.

Una vez que la información obtenida del cuestionario y de las 10 entrevistas, fue objeto de un primer análisis, se procedió a convocar un grupo de discusión entre las ONGD que habían participado en la primera fase, con el fin de profundizar en algunos aspectos e incluso de poder contrastar alguna información contradictoria que se había obtenido con el cruce de las dos herramientas. El trabajo realizado en el taller bajo grupo focal entre los miembros de las ONGD participantes permitió una triangulación adicional de las valoraciones obtenidas a partir de los cuestionarios.

A partir de la información obtenida con las tres herramientas, los evaluadores procedieron a realizar un segundo y más completo análisis de la información obteniendo unas conclusiones preliminares.

Estas conclusiones fueron contrastadas en primer lugar con el personal responsable de la gestión de las evaluaciones por parte de la AECID (Consejera técnica y experta externa en evaluación) a través de una entrevista en profundidad, y en segundo lugar con las ONGD participantes en el estudio a través de un taller de devolución de conclusiones preliminares. Esta última fase sirvió para afinar más las conclusiones últimas obtenidas en el informe de evaluación así como para poder establecer las recomendaciones finales.

En conclusión, el diseño metodológico ha sido cuidado en extremo y la combinación de herramientas cuantitativas y cualitativas permite confiar razonablemente en las valoraciones descritas. (En anexo II se pueden consultar los guiones utilizados para cada una de las herramientas cualitativas utilizadas y en anexo III los principales estadísticos descriptivos).

El resto del estudio se organiza en las siguientes secciones. En la sección 4 se realiza el análisis cuantitativo y cualitativo de la información recogida. En las dos siguientes secciones, se describen las principales conclusiones y recomendaciones. Y por último, se exponen los condicionantes y límites de la evaluación

Análisis sistémico del objeto de evaluación.

DISEÑO/ESTRUCTURA

Normativa AECID, presupuesto, recursos humanos de las ONGD y sus SL, capacidades y experiencia en las empresas externas evaluadoras, cultura de evaluación en la ONGD y sus SL, tiempo para realizar la evaluación, tiempo para aplicar recomendaciones

Procesos clave:

1. Diseño de la evaluación: Elaboración de TdR (ONGD- AECID)
2. Externalización de la evaluación
3. Selección de empresa evaluadora (locales, conocidos, conocimiento den metodología de evaluación, conocimiento en el sector y/o país, etc.)
4. Ejecución de la evaluación (recogida de información de campo, análisis de la misma, comunicación con las ONGD y AECID, comunicación de resultados, etc.)
5. Discusión de conclusiones y recomendaciones (ONGD-AECID) para retroalimentar convenio
6. Aplicación de recomendaciones al ciclo del proyecto/convenio (ONGD y Sus SL)

OE y RESULTADOS

OE: Intervenciones de cooperación al desarrollo de las ONGD de mayor calidad.

R1. El proceso de evaluación intermedia de los Convenios con AECID genera utilidad/aprendizaje en las ONGD (cultura evaluativa, aprendizajes que van más allá de la aplicación en el propio convenio evaluado).

R.2. Las evaluaciones intermedias producen recomendaciones útiles y aplicables al ciclo del proyecto en el marco de los convenios AECID

IV. ANÁLISIS CUANTITATIVO Y CUALITATIVO.

En esta sección se presenta el análisis cuantitativo y cualitativo obtenido a partir de la información primaria creada por la investigación.

Para el análisis cuantitativo la herramienta de recogida de información fue un cuestionario (véase el Anexo 1). El cuestionario diseñado estuvo compuesto por 45 preguntas (o variables). Todas las variables son discretas y cesuradas. El cuestionario se organizó en torno a seis grupos temáticos incluyendo en cada grupo diversa cuantía de preguntas. Los grupos, nombre de las variables y recorrido de cada una se ofrecen en la Tabla 1.

En el análisis previo a la entrega del cuestionario, se identificó un universo de 37 ONGD que había obtenido convenio en la convocatoria de AECID en los años 2006 y 2007. De este universo se han logrado respuestas de 28 organizaciones lo que representa una muestra del 75,67%.

Para el análisis cualitativo se utilizaron dos herramientas de recogida de información. La entrevista en profundidad realizada a 10 de las ONGD de la muestra, cuyo guión fue la propia encuesta pero formulando las preguntas de la misma de manera abierta. La entrevista se aplicó paralelamente en el tiempo a la encuesta. Posteriormente, y tras un primer análisis de la información obtenida a través del cuestionario y las entrevistas, se desarrolló un grupo de discusión para profundizar en el análisis cualitativo. Para la realización de dicho grupo se invitó a la totalidad de las 28 ONGD que habían contestado al cuestionario, y se obtuvo una participación de 7 ONGD.

A continuación, se comentan los principales resultados obtenidos del análisis de los cuestionarios, entrevistas y grupo de discusión, bajo la perspectiva cuantitativa y cualitativa, agrupados en los seis bloques temáticos. Los principales estadísticos descriptivos de cada variable pueden consultarse en el Anexo 2.

Tabla 1. Grupos, variables y recorrido de las preguntas del cuestionario.

	Variable	Recorrido		Variable	Recorrido
organización y promoción	MOMENTO	1-3	equipo evaluador	Diseño	1-4
	MOTIVO	1-4		Informac	1-4
	3OFERTAS	1-4		Obj+Credib	1-4
	AECID_acepta	1-0		Comunic	1-4
participación	PO_TdR	1-0		Recomends	1-4
	PSL_TdR	1-0		completa	1-4
	PO_preg	1-0		intereses	1-4
	PSL_preg	1-0		equipo	1-4
	PO_crit	1-0		ajuste	1-4
	PSL_crit	1-0	valoración global	Valor	1-10
	PO_campo	1-0	utilidad y credibilidad	Util_proceso	1-4
	PSL_campo	1-0		Creib_concl	1-4
	PO_concl	1-0		Util_recom	1-4
	PSL_concl	1-0		P-mejora	1-0
	PO_recom	1-0		cambiosAECID	1-0
	PSL_recom	1-0		recuerda	1-0
	PO_taller	1-0		Concl-util	1-0
	PS_taller	1-0		Más-útil	1-5
	Ev_pregO	1-0		Limit-dific	1-0
	EV_pregSL	1-0		tiempo	1-4
	PO_informe	1-0		RRHH	1-4
	PSL_informe	1-0		dinero	1-4
			difusión	difundida	1-0

2.1. Organización y promoción de las evaluaciones.

El cuestionario/entrevista se iniciaba con cuatro preguntas relativas a la organización y promoción del proceso evaluativo. Al ser la primera vez que se emprendían evaluaciones intermedias de los convenios resultaba interesante conocer si **el momento**, que venía determinado por la propia normativa, de la evaluación fue el más oportuno (o hubiera sido mejor otro como al final del convenio o algún tiempo antes/después de la mitad del proceso cuatrienal). También se dio la opción de responder que no se hubiera hecho de haber sido obligado.

El promedio se situó en 1,85 (recorrido 1-3) con una desviación típica de 0,45. La mayoría de las organizaciones responde que se hizo en el momento oportuno (21 de las 27), cinco opina que hubiera sido más interesante en otro momento y sólo 1 revela que realizó la evaluación por obligación. La explicación cualitativa de esta opción reveló que el convenio comprendía 15 países y en muchos de ellos el proceso planteado es muy largo (incluye segunda fase) y hubiera sido mejor hacerla al final. Esta opción, de poder haber evaluado algo más tarde debido al proceso de desarrollo o al inicio real de las intervenciones más allá de la fecha formal, es la razón más abundante tras la selección de “hubiera sido más interesante realizarla en otro momento”. A modo de ejemplo, se cita una frase textual recogida de la información cualitativa:

“la realidad es que estaban pensados como convenios de mayor duración (todos ellos tienen segunda fase) por lo que hubiera sido más útil una evaluación final que en realidad hubiera sido intermedia del proceso tal y como estaba pensado”.

De la información cualitativa, también se desprende que el proceso de evaluación completo es bastante largo, mínimo cinco o seis meses desde que se comienza a elaborar los Términos de Referencia (TdR de aquí en adelante) hasta que das por concluido el informe, lo que hace que apenas quede un año de ejecución para aplicar las recomendaciones, muchas de las cuales son demasiado ambiciosas para aplicarlas en lo que resta del convenio.

El **motivo** predominante al realizar la evaluación ofrece mayor dispersión (desviación típica de 0,97 para un promedio de 2,7). Dado que se abría la posibilidad de respuesta múltiple hay que interpretar con cautela el promedio. La respuesta con frecuencia más alta es la evaluación se realiza para establecer un proceso de *mejora continua* (se señala en 16 ocasiones, el 42%

de las observaciones), seguida por analizar los *procesos* internos (9 veces, 24%) y conocer los *resultados* logrados (8 veces). Cumplir con la *obligación* marcada por el convenio se señaló 5 veces (el 13%). Resulta interesante advertir que, de la información facilitada por las entidades, no se desprende una dicotomía entre la evaluación de procesos y la de resultados, obteniendo ambas casi la misma selección.

De la información cualitativa se desprende que el modelo de TdR que es obligado a utilizar actúa como un corsé. Los siguientes testimonios lo muestran con claridad:

“Más allá de los cinco criterios del CAD nos interesaba evaluar los procesos, pero conociendo a la AECID no te vas a meter a añadir cosas. Nosotros directamente les dijimos a los evaluadores donde nos interesaba que se centraran más allá de lo que reflejaban los TdR”;

“Quisimos meter un criterio más y fue todo un lío”;

“Otra debilidad era que el objetivo de la evaluación era demasiado amplio, los 5 criterios, mas procesos, etc. Hubiera sido más interesante haberse centrado en alguna cuestión concreta porque por ejemplo, los procesos no fueron bien analizados y hubiera sido más interesante en este momento”;

“El informe no se ajustó a los 5 criterios y AECID nos hizo reordenar todo el informe de nuevo. Para nosotros fue tiempo perdido pues el informe nos parecía bueno ya que respondía a lo que a nosotros nos interesaba”.

Varias organizaciones afirman que trataron de no meter algún criterio de los 5 recomendados pero la AECID les obligó a incorporarlo.

Incluso más allá de los propios TdR, las ONGD opinan que depende del técnico responsable del Convenio por parte de AECID y de la OTC correspondiente para tener mayor o menor flexibilidad para centrar la evaluación en lo que realmente le interesa a la ONGD.

Esta participación de la AECI para reflejar sus intereses en el proceso de evaluación, ha implicado una gran participación en la fase de elaboración de los TdR. En este sentido, se recogen algunas citas de la opinión que las ONGD tienen sobre este proceso:

“El tiempo empleado por nuestro personal ha sido excesivo sobre todo en la negociación con AECID de los TdR que se meten a opinar hasta el último detalle de la logística (por ejemplo les parecía mucho tiempo de trabajo de campo y poco de revisión de documentación), este

tipo de observaciones, no aportan mucho, casi lo llamaríamos de “intromisión”, y hace perder muchísimo tiempo”;

“En todos los casos, el tiempo empleado en la elaboración de los TdR fue excesivo debido al nivel de detalle en que AECID hacía sus aportaciones y exigencias, la discusión de los TdR fue un proceso interminable y agotador”;

“Los TdR se manipulan mucho, de aquí para allá y se acaba perdiendo el enfoque de lo esencial, de lo importante”.

A pesar de todo lo expresado por las ONGD, de la entrevista realizada a las responsables de la evaluación en AECID, se recoge que la propia AECID también considera que los TdR eran demasiado ambiciosos y que incluso no era obligatorio evaluar de acuerdo a los 5 criterios del CAD. En palabras de la propia AECID: “la normativa no lo recoge así”. De hecho, la AECID consideraba imprescindible evaluar en todos los convenios tres de los criterios: diseños (debido a las debilidades de la mayoría de ellos), eficacia y pertinencia, considerando que sostenibilidad e impacto no eran necesarios.

Por otro lado, reconocen que fue necesario meterse muy a fondo en la supervisión de los TdR porque la mayoría de ellos eran muy ambiciosos y las preguntas de evaluación eran demasiado generales, sin concretar lo que verdaderamente se quería evaluar.

La pregunta sobre **cómo fueron seleccionadas las empresas evaluadoras** también permitió una selección múltiple. Se han obtenido 44 observaciones. La más frecuente es el contacto con empresas de las que *ya se tiene buenas referencias* (15 selecciones que representan el 34%). Prácticamente al mismo nivel se sitúa el *concurso público* (14 observaciones, el 32%). La posibilidad de recibir ofertas de empresas con las que ya se había trabajado anteriormente obtuvo 8 observaciones y 7 la opción de recibir ofertas de las que no se tenían referencias previas. Las elecciones combinadas se reparten del siguiente modo: 3 entidades optan por la combinación 2-3 (empresas con contacto previo y de las que se tenían buenas referencias por otras organizaciones); dos veces la combinación 2-3-4 que añade a la anterior empresas sin referencias previas; una vez la 1-3-4 y una vez 1-2-4 y una entidad que marcó todas las respuestas posibles. Como conclusión, puede señalarse que las ONGD, dado que les es obligado hacer la evaluación externa, optan por hacer un concurso pero con cierta “restricción” a empresas con las que ya se ha trabajado o se han obtenido buenas referencias.

Desde un punto de vista de la información cualitativa recogida, al responder a esta pregunta, inevitablemente surgió la opinión de las ONGD sobre la pertinencia de la “obligatoriedad de evaluaciones, lo cual se refleja a través de las siguientes citas:

“De igual modo, pienso que se es muy estricto con el carácter externo de las evaluaciones, ya que en algunos casos si las evaluaciones son mixtas pueden conseguirse mejores resultados para todos”;

“Un equipo mixto agregaría valor añadido”;

“Si lo que se está buscando es aprendizaje entonces sería mejor una evaluación mixta, pero si lo que se busca es una rendición de cuentas entonces creo que es mejor que sea externa e incluso debería ser contratada por la propia AECID para garantizar mayor transparencia”;

“En una evaluación intermedia lo principal es aprender para corregir y por ello sería mucho mejor una evaluación mixta”.

La AECID considera que todas las evaluaciones fueron mixtas en su fase inicial, en el momento de elaboración de los TdR. También opina que si la evaluación es financiada al 100% por fondos públicos, ésta no puede ser mixta, ya que para ello, las ONGD deberían aportar de sus propios fondos.

Una mención aparte ha sido la introducción en la evaluación de los convenios de 2007 de que una de las ofertas tenía que ser de una **empresa local**.

Algunos testimonios sobre la opinión de las ONGD con respecto a este punto son:

“Los evaluadores locales tenían menos experiencia y no llegaban a comprender la metodología impuesta por AECID, los criterios, etc. Se les quedó grande el trabajo, además de no respetar los tiempos marcados, no se respetaron los TdR”;

“En todos los casos nos obligaron a presentar alguna oferta de empresa evaluadora local. Sobre este tema no le encontramos mucho el significado porque en convenios multi-país, el hecho de que el evaluador sea boliviano, por ejemplo, a veces es peor que si fuera español, para evaluar en Perú, por la rencilla cultural que hay entre estos dos países, y por otro lado, ¿por qué el boliviano va a conocer mejor la cultura de Perú?”;

“Se contrató una empresa española que tenía personal local en cada país del convenio pero al final el resultado fue que cada evaluador evaluó su país y no se dio una visión regional”.

Con respecto a la pregunta sobre si **AECID aceptó** la propuesta realizada en la elección de la empresa, todas las ONGD responden que sí excepto una. El motivo fue que al ser de 2007, se solicitó por parte de AECID la repetición del concurso público a nivel local (el convenio se ejecutaba en Colombia). La experiencia no fue buena ya que no se presentó nadie y se volvió a la propuesta original, retrasando mucho el proceso.

2.2. La participación a lo largo del proceso evaluativo.

Esta sección del cuestionario se compuso de 9 ítems que preguntaban de forma dicotómica (por tanto se da el valor 1 si se contestó con un “sí” y valor 0 si la selección fue un “no”) sobre nueve momentos en los que se dividió el ciclo de la evaluación. Desde los términos de referencia a participación en los momento anteriores a la recepción del informe de evaluación. Para aligerar el análisis, hemos procedido a agregar las respuestas obtenidas en las que sí se participa por parte de la ONGD y –en otro grupo- la participación de los socios locales. Dos entidades sólo respondieron a la participación de la ONGD, con lo que éstas tienen más respuestas que la de los socios locales. Los resultados de forma resumida se ofrecen en la tabla 2.

Tabla 2. Participación a lo largo del proceso evaluativo: frecuencias acumuladas.

	TdR	preguntas	criterios	trabajo de campo	conclusiones	recomendaciones	taller en terreno	¿pidió el evaluador?	antes de recibir el informe
ONGD	27	27	12	16	18	15	22	26	27
Socio Local	14	15	*2	15	14	8	18	20	19

Nota: el número de respuestas de ONGD es 28, mientras que el de Socio Local son 26, excepto en el ítem “criterios” que es 25.

Como se desprende de la tabla, prácticamente todas las ONGD han participado en la elaboración de los términos de referencia, la formulación de las preguntas evaluativas y en la fase previa a la recepción del informe final de evaluación. La frecuencia más baja de participación es en la selección de los criterios de evaluación (12 de 28). No contar con la libertad por parte de AECID para cambiar los cinco criterios canónicos del CAD (aunque esto no fuera explícito en la normativa) está tras esta menor participación. También la participación descende en el momento del trabajo de campo, elaboración de las conclusiones

y de las recomendaciones. La interpretación que se hace de esta menor frecuencia es que se respeta la independencia de los evaluadores al ser una evaluación externa.

Por lo que se refiere a la participación de los socios locales en las evaluaciones, vemos que es menor que la de las ONGD. Mientras las ONGD ofrecen una participación global de 190 respuestas positivas (“sí”) el valor del socio local es de 123 síes. La *menor participación de los socios* se produce en la selección de criterios (2), seguida de la elaboración de las recomendaciones (8). La participación es de casi el doble cuando se hace referencia a las fases finales del ciclo evaluativo: taller en terreno, solicitudes del evaluador o antes de recibir el informe final (fase de borrador del informe).

Es llamativo que en la mitad de los casos posibles, los socios no participan en la elaboración de los términos de referencia ni en la formulación de las preguntas evaluativas. De hecho, las mayores desviaciones típicas dentro del grupo de variables relativas a la participación se producen con los socios locales (términos de referencia, preguntas, campo y conclusiones superan el 0.50 de desviación típica).

La información recogida de manera cualitativa muestra que las ONGD consideran fundamental participar en el proceso de evaluación:

“Lo fundamental es participar desde el principio porque si no le dedicas el tiempo suficiente desde el primer momento, luego te encuentras con los problemas”;

Cuando participas todo fluye mucho mejor y tienes la capacidad de negociar aunque estés a mitad del proceso”;

“Si hay participación y un buen equipo de evaluación, todo lo demás viene dado”;

De hecho, las 7 ONGD participantes en el grupo de discusión afirmaron que la participación es uno de los elementos clave para el desarrollo de buenas evaluaciones.

Del análisis de la información recogida de manera cualitativa se desprende que para las ONGD con mayor cultura evaluativa (las que ya cuentan con personal específico formado en la materia o cuentan con un plan de evaluaciones, por ejemplo), el aspecto de la participación es fundamental e incluso consideran que el aspecto más clave en todo el proceso es saber formular bien las preguntas de evaluación, transmitir lo que se quiere conocer de manera adecuada al equipo de evaluación.

Para las ONGD con menor cultura evaluativa, incluso para aquellas que estas evaluaciones eran las primeras que se realizaban en la organización, también se reconoció el aspecto de la participación como clave:

“La evaluación ha sido muy útil, en primer lugar por la posibilidad de acompañar el proceso de evaluación, siendo el aprendizaje muy interesante para la organización”;

“Se han sacado muchos aprendizajes sobre todo de cómo hacer un proceso de evaluación, ha sido compartido con toda la organización”.

Incluso, como veremos más adelante cuando analicemos los principales cuellos de botella, se verá como algunos de ellos se podrían haber superado con una mayor participación por parte de las ONGD.

A pesar de esta alta participación, especialmente en las fases iniciales del proceso de evaluación, la AECID considera que el trabajo de elaboración de las preguntas evaluativas podría mejorarse por parte de las ONGD, haciendo que éstas fueran mucho más concretas y que realmente respondieran a las necesidades informativas de las organizaciones.

Sin embargo, con respecto a la participación de los Socios Locales (SL a partir de ahora), corroborando los datos del análisis cuantitativo, se aprecia una menor participación en las fases iniciales del proceso: TdR, definición de preguntas de evaluación y de los criterios. Las ONGD, a través de la información cualitativa recogida, expresan que los socios locales tienen menos cultura de evaluación y que ésta es considerada como un mero trámite a cumplir, sin que exista un mayor interés por el aprendizaje o la mejora. Según las ONGD éste es el motivo principal que explicaría la menor participación de los SL en todo el proceso de evaluación: *“Los socios locales lo siguen viendo como una imposición, nosotros ya lo vemos como una oportunidad”.*

En opinión de la AECID, la experiencia demuestra que cuanto más grande es el SL, mayor capacidad para implicarse y participar en la evaluación tenía, y que al igual que ocurre con las ONGD se debería garantizar una mayor participación de los SL en la elaboración de las preguntas de evaluación para conocer realmente cuáles son sus necesidades informativas.

2.3. Los equipos de evaluadores externos y su trabajo.

El tercer grupo de preguntas tuvieron la finalidad de conocer la percepción que las ONGD tenían sobre el desempeño de los equipos de evaluación (o evaluador individual). Se diseñaron nueve ítems y se ofrecía una escala ordinal de dimensión cuatro (para evitar la tendencia central). Las opciones de respuesta eran malo-regular-bueno-muy bueno. La codificación se ha realizado en números naturales positivos, dando valor 1 a malo y 4 a muy bueno.

La organización del análisis de este grupo es en dos etapas. Primero procedemos a una visión de conjunto de las 9 variables y posteriormente analizamos una por una.

Todas, excepto la de diseño, tienen 28 observaciones (respuesta completa). El diseño, la recogida de la información, la valoración del equipo y el ajuste entre trabajo diseñado y realmente efectuado tienen un valor mínimo de 1 (malo). El mínimo del resto de variables es 2. El máximo es 4 en todas las variables.

La frecuencia del valor 1 (“malo”) es 4 (el 2% de las 239 observaciones que componen el total del grupo sobre el trabajo de los evaluadores). La de 2 (“regular”) es 62 (26% de las respuestas). El valor 3 (“bueno”) se eligió en 151 ocasiones, representando el 63%. La opción 4 (“muy bueno”) obtuvo 22 observaciones (el 9%). En resumen, la opinión predominante en las ONGD con convenio ha sido que el trabajo realizado por los evaluadores ha sido “bueno”, aunque la cuarta parte se decanta por regular (26%).

La mayor satisfacción se concentra en la *comunicación* mantenida entre evaluadores y ONGD (7 frecuencias de “muy bueno” que son el 26%, muy por encima del 9% que tiene ese valor en el global). Por encima del valor promedio de la calificación “bueno”, están las variables “*diseño*” (frecuencia 20 que es un 74%); “*ajuste*” (frecuencia 19 que es un 73%); “*equipo*” (con frecuencia 18 que es un 69%) y “*recomendaciones*” (con 18 observaciones sobre 26 – una menos que “equipo”, lo que supone un 67%).

Por el lado de la valoración “regular”, sobresalen “*completa*” (11 observaciones que es un 42%) y la *información recogida* (10 que supone un 38%).

En resumen, las ONGD expresan satisfacción con la comunicación mantenida por los equipos evaluadores y, en menor medida, con el diseño, ajuste entre el trabajo finalmente realizado y el diseñado, el equipo y las recomendaciones recibidas. En la valoración cualitativa se detecta que hay varias quejas sobre las recomendaciones en el sentido de que son muy generales, ya sabidas o muy difíciles de implementar en la práctica:

“Las recomendaciones emitidas por los evaluadores son, con frecuencia, poco aplicables. Al basarse en una observación puntual y un conocimiento del contexto específico con limitaciones, en ocasiones las recomendaciones resultan un tanto teóricas y poco adaptadas a la realidad. Y en la mayor parte de los casos no aportan el valor añadido de asistencia técnica que se le supone a una evaluación, ya que no se aportan los “cómos” a los “qués”. En general las evaluaciones son muy útiles para poner de relieve “qué” se debe hacer o corregir, pero raramente aportan “cómo” hacerlo o corregirlo en el contexto real concreto, por lo que la puesta en marcha de las propuestas de mejora es difícil”;

“Las propuestas de mejora se quedaron escasas y el documento era demasiado extenso (372 páginas), lo que hace que sea poco práctico”.

Donde menos satisfacción parece haber (o mayor diferencia de opiniones) es la cantidad de información recogida (relacionada directamente con el tiempo pasado en el terreno), tal y como se recogen en los siguientes comentarios:

“La principal limitación es no tener tiempo suficiente para el levantamiento de la información y eso de lugar a evaluaciones de poco rigor”;

“Es poco el tiempo que se pasa en terreno para evaluar un proyecto y entender la ejecución del mismo, conocer actores importantes, etc....”.

También hay menor satisfacción en que la evaluación haya resultado completa, identificando fortalezas y debilidades. La información recogida de manera cualitativa lo corrobora a través de comentarios como:

“Se ensañaron en fallos que incluso no representaban ni el 1% del presupuesto”;

“Hay una tendencia a destacar sobre todo las debilidades”;

“El informe fue duro, solo se centró en las debilidades”.

Pasando ahora al análisis variable por variable, los estadísticos descriptivos revelan que, en promedio, la mayor satisfacción se obtiene en la *comunicación* (media de 3,05 y desviación típica de 0,67 la mayor de entre los nueve ítems). A continuación, el *diseño* obtiene un

promedio de 2,85 puntos (y dispersión de 0,60). Le siguen *objetividad/credibilidad* de la interpretación de la información recogida, las *recomendaciones*, y que se hayan tenido en cuenta los *intereses* de la ONGD. Las tres tienen un promedio de 2,80 y desviación típica de 0,61. Un cuarto grupo lo constituyen la valoración promedio del equipo y el ajuste a lo propuesto, con una media de 2,78 y desviación típica de 0,61 y 0,64 respectivamente. Por último, obtienen un promedio de 2,57 la información recogida (dispersión de 0,62) y la completitud de la evaluación (dispersión de 0,48).

Para una visualización más amable de este “ranking” la tabla 3 contiene el promedio y desviación típica de los ítems ordenados de mayor a menor media.

Tabla 3. Ranking del promedio de los ítems relacionados con la calidad del equipo evaluador.

variable	promedio	desviación típica
Comunicación	3,0536	0,6713
Diseño	2,8519	0,6015
Objetividad-Credibilidad	2,8036	0,6137
Recomendaciones	2,8036	0,5500
Intereses	2,8036	0,6137
Equipo	2,7857	0,6150
Ajuste	2,7857	0,6444
Información	2,5714	0,6193
Completa	2,5714	0,4852

2.4. La valoración cuantitativa global.

La variable con mayor recorrido dentro del cuestionario fue la de ofrecer una valoración global del proceso de evaluación entre 0 (peor) y 10 (mejor valoración). El mínimo ha resultado ser un 3 y el máximo un 9, con un promedio de 6.58 y desviación típica de 1,36.

Como se puede observar en la tabla de frecuencias (Tabla 4), la valoración más elegida ha sido el 7 (frecuencia 12) agrupando casi el 43% de las respuestas.

Tabla 4. Frecuencias de la valoración global.

Valor

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	3,00	1	3,6	3,6	3,6
	4,00	2	7,1	7,1	10,7
	5,00	2	7,1	7,1	17,9
	6,00	4	14,3	14,3	32,1
	6,50	1	3,6	3,6	35,7
	7,00	12	42,9	42,9	78,6
	8,00	5	17,9	17,9	96,4
	9,00	1	3,6	3,6	100,0
	Total	28	100,0	100,0	

Puede concluirse que la satisfacción promedio con el proceso evaluativo es de “notable”, lo que puede considerarse alta si se tiene en cuenta que es una experiencia novedosa que se realiza por primera vez.

De la información cualitativa, se recoge alguna información que complementa la anterior y que surgió de manera espontánea al no ser preguntas directamente formuladas en el cuestionario (por ejemplo, cuando se preguntó sobre los principales cuellos de botella en el proceso). Como se podrá observar a continuación todos los comentarios se centraron en el propio equipo de evaluación, por lo que parece ser el tema que más preocupa a las ONGD, y que como veremos más adelante en el análisis cuantitativo, es la variable que más influye en el grado de utilidad de las evaluaciones realizadas. A continuación se detallan los comentarios más relevantes.

En general, las ONGD opinan que es mejor contar con equipos evaluadores compuestos de un mínimo de 2 personas (el máximo de personas en un equipo fue de 5) que con un solo evaluador.

También existe una queja generalizada cuando en el proceso se contrata a una empresa que libremente cambia al equipo evaluador que inicialmente se había comprometido. Las ONGD consideran que la clave de la calidad del equipo evaluador está en las personas y no en la empresa porque incluso de la misma empresa hay personas que trabajan mucho mejor que otras:

“En algún caso la persona que ha realizado la evaluación no tenía buena preparación, en ese caso se contrató con empresa consultora unos TdR y luego ellos buscaron a la persona evaluadora y la propia empresa consultora reconoció el desastre de la evaluación”;

“La explicación de los malos resultados se debe a que en los TdR la empresa ganadora ofrecía un determinado personal que luego no se respetó”;

“Al final, los resultados dependen mucho de la aptitud y actitud de los evaluadores”.

También surge repetidamente un dilema entre la contratación de expertos en evaluación y/o la contratación de expertos en el sector y/o país en el que se desarrolla la intervención. En este sentido:

“Se dio un choque de teorías entre el evaluador, que además era experto en su tema (educación, o agrario) y el propio responsable del SL que era experto en esa misma materia. Al final, el informe parece más un informe de experto en la materia (con recomendaciones muy útiles) pero no tanto un informe de evaluación, pierden un poco el rol de evaluador para decantarse por el rol de experto en la materia”;

“Es muy difícil encontrar gente que sea buena evaluadora y a la vez expertos en el tema, por eso la importancia de los equipos mixtos porque la ONGD puede aportar la experticia en la materia”.

Según los diversos comentarios y tal y como se expresó en el grupo de discusión, parece que la mejor opción es decantarse por expertos en evaluación y complementar sus limitaciones en conocimientos concretos en el tema a través de la propia ONGD o mediante la creación de equipos.

Por último, otro tema que salió recurrentemente en la información cualitativa sobre el equipo evaluador es la dificultad de encontrar verdaderos expertos en evaluación en muchos de los equipos que se contrataron. Esta fue apuntada como la principal debilidad de los equipos de evaluación. Las ONGD echan en falta experiencia tanto en el diseño de la propia metodología de evaluación como en las técnicas de recogida de información. Algunos comentarios lo muestran:

“Muchas veces los datos no son fiables porque van allí y apuntan lo que les dice el primero que se encuentran y no hacen una adecuada interpretación”;

“Es difícil saber cuál es la mejor propuesta técnica porque todas las metodologías eran super-pobres”;

“...Y hablar de técnicas cualitativas y cuantitativas, eso ya es ciencia ficción, ja, ja...”.

En opinión expresada por la AECID con respecto al trabajo realizado por los equipos evaluadores, la nota media que se daría al conjunto de procesos de evaluación realizados es menor que el “notable” que se obtenía de la percepción de las ONGD, situándose su percepción en un “suficiente” ó 5, especialmente si se puntuaba cada proceso de acuerdo a la ficha que utiliza la UE con 9 criterios para realizar la meta-evaluación de las evaluaciones³. Una cuestión clave que la AECID califica como “deficiencia” es que fueron poquísimas las situaciones en que los equipos evaluadores decidieron entrevistar al donante, la propia AECID.

Algunas otras debilidades señaladas por AECID de los procesos de evaluación son:

- En ocasiones eran más informes de seguimiento que de evaluación. Utilizaban los indicadores de la matriz de intervención como indicadores de evaluación en vez de crear el sistema de indicadores de evaluación;
- En ocasiones no había ni matriz de evaluación;
- Casi en ninguna evaluación había una justificación de la metodología ni del porqué de utilizar una u otra herramienta de recogida de información y los criterios para haberla desarrollado;

³ Para conocer en detalle los nueve criterios de calidad que utiliza EuropeAid y cómo operativizarlos véase http://ec.europa.eu/europeaid/evaluation/methodology/guidelines/gui_qal_flr_es.htm#03

- En recomendaciones en ocasiones se metían “peticiones o reclamos” del propio SL.

2.5. Utilidad y credibilidad.

Este grupo se estructuró en torno a cuatro bloques de conocimiento. Primero se interrogó sobre la utilidad general del proceso, luego sobre las consecuencias concretas que ha tenido la evaluación, en un tercer momento para quién ha resultado más útil y en cuarto lugar cómo se percibía el tiempo, recursos humanos y dinero empleados en la evaluación.

Sobre la **utilidad** general de la evaluación se obtienen los siguientes resultados. Ninguna ONGD ha valorado el proceso con la puntuación mínima (“nada útil”), predominando con la misma frecuencia (13 observaciones) la valoración “algo” y “mucho” o muy útil (las dos máximas puntuaciones que había en la escala de 4). Ha habido un caso de “poco” y una indecisión entre “algo-mucho” (codificada con un puntaje de 3,5). El promedio es 3,44 con desviación típica de 1,36. Es decir, el 92% de las entidades valora el proceso como útil o muy útil. Este dato cuantitativo se corrobora con la información recogida de manera cualitativa, tal y como lo muestran comentarios del tipo:

“La evaluación ha sido muy útil, en primer lugar por la posibilidad de acompañar el proceso de evaluación, siendo el aprendizaje muy interesante”;

“Nos ha aportado mucho para mejorar los convenios que estamos ejecutando en la actualidad”.

Sobre la credibilidad de *las conclusiones y juicios*, los resultados son muy parecidos. Con un promedio de 3,41 ($\sigma=0,62$) hay 13 valoraciones de “mucho” y 12 de “algo” (es decir, 25 valoraciones máximas en la escala de 4).

Un poco menos útiles han sido valoradas las recomendaciones, pero con un promedio de 3,34 ($\sigma=0,61$), la opción “algo” es elegida en 14 ocasiones y “mucho” en 11, es decir, 25 ONGD se muestran satisfechas con las recomendaciones obtenidas a pesar de que en la información cualitativa se recogieron algunos comentarios críticos hacia las recomendaciones en el sentido de que eran muy generales, ya sabidas o muy difíciles de implementar e incluso hubo alguna ONGD que mencionó la poca utilidad de las recomendaciones como principal cuello de botella para integrar la evaluación en el ciclo del proyecto.

Es decir, la percepción revelada sobre la utilidad del proceso, las conclusiones y recomendaciones parece ser alta y muy alta en la mayoría de las ONGD.

Esta alta valoración conduce a querer conocer **las consecuencias** concretas que ha tenido la evaluación sobre las intervenciones de las ONGD.

Una primera consecuencia inmediata podía ser elaborar un *plan de mejora* por escrito. Aunque esta acción es obligada según el procedimiento establecido por AECID, 5 organizaciones (el 18%) responden de forma negativa. En las ONGD que fueron entrevistadas se aludió que el motivo era precisamente el momento en el que se encontraba actualmente el proceso de evaluación, es decir, que recién se había recibido el informe y todavía se estaba analizando o que se estaba en pleno proceso de elaboración. Para el resto de organizaciones, las encuestadas, no se aluden los motivos para no contar con este plan de mejora. Sin embargo, resulta interesante mencionar que en la información cualitativa recogida son muchas las ONGD que afirman que su plan de mejora va más allá de la propia adenda firmada con AECID, tal y como lo muestra comentarios del tipo:

“Nuestro plan de mejora incluye una hoja de ruta mucho más exigente que lo que nos pedía la AECID, fuimos recomendación por recomendación aplicando mejoras”;

“El convenio estaba formado por 6 SL y se han elaborado un plan de mejora para ser ejecutado por cada uno de ellos. El plan de mejora va mucho más allá de lo exigido por la AECID”.

Otra consecuencia de la evaluación, al ser intermedia, es que se introduzcan *modificaciones* y se soliciten cambios al financiador (AECID). La distribución de respuestas es que el 46% (13 organizaciones) sí solicitaron cambios y el 54% no (15 entidades). En este punto es importante aclarar que por exigencias de la propia normativa, sí se pueden estar dando modificaciones y plan de mejora a un convenio evaluado pero que no requiera de solicitar formalmente un cambio porque no supone modificación sustancial (para conocer en mayor

detalle las características de las modificaciones sustanciales puede remitirse a la normativa AECID que regula a los convenios⁴).

Para verificar de forma más segura las consecuencias de la evaluación, se preguntó si se recordaba alguna recomendación surgida del informe de evaluación y se dijera cuál. El 89% responde de forma afirmativa, pero 3 ONGD (10,7%) confiesan no recordar en ese momento ninguna recomendación.

Por último, dentro del bloque de las consecuencias, se preguntó si se habían extraído conclusiones útiles para ser aplicados en programas similares de la ONGD (*replicabilidad*). Eligieron la respuesta afirmativa 26 entidades (el 93%).

De la información cualitativa se recogen algunas de estas utilidades. A modo de ejemplo, podemos citar:

“Una clara recomendación ha sido centrarse en un tipo de derechos a trabajar con las mujeres y no de manera tan integral como hicimos pues esto provocó una dispersión en el efecto”.

Otra organización comentó que se ha realizado una sistematización interna de todas las recomendaciones de las evaluaciones y éstas han sido difundidas dentro de la organización. También se recogen comentarios sobre recomendaciones útiles para el diseño de nuevas intervenciones en relación a su estructura de gestión (integrar menos proyectos, trabajar en menos países en un mismo convenio, reformulación de indicadores para que sean más SMART, mejorar las fuentes de verificación, acotar los sectores de intervención), incluso se han obtenido recomendaciones a aplicar sobre el propio proceso de evaluación para ser aplicadas en futuras evaluaciones (cerrar muy bien desde los TdR los CV de las personas evaluadoras, concretar mucho más las preguntas de evaluación y definir en los TdR de forma muy precisa el producto final que se quiere obtener).

⁴ Véase http://www.aecid.es/galerias/ongds/descargas/Guia_aplicacion_Normas_ONGD_2010.pdf pág 19.

Tras la valoración de la utilidad y las consecuencias, se preguntó **a qué agente** se pensaba que había beneficiado más la evaluación y se ofrecían, AECID, la ONGD, el socio local, los beneficiarios u otros. La distribución de frecuencias (se permitió la respuesta múltiple) ha sido la siguiente:

Tabla 5. Frecuencias de la pregunta “¿Para quién cree que ha sido más útil?”

Opción	frecuencia	%
AECID	7	14%
ONGD	23	47%
Socio Local	13	27%
Beneficiarios	6	12%
Suma	49	100%

El ranking de elecciones es ilustrativo. La evaluación parece haber sido más útil a la ONGD (sujeto que responde) seguido de las contrapartes locales. AECID y beneficiarios son escogidos casi en la misma cuantía. Las razones cualitativas de esta elección son amplias. Las ONGD pueden considerar que los beneficiarios obtienen fruto de la evaluación de manera indirecta, en la medida en que se mejoran las intervenciones, pero no de manera directa porque consideran que todavía siguen siendo sujetos pasivos a los que se les hace un feedback del proceso.

Con respecto a la AECID, las ONGD consideran que ésta no tiene capacidad para asimilar todo el volumen de evaluaciones realizado.

Este ranking es corroborado por la percepción de la AECID sobre el grado de utilidad para cada agente. La propia AECID considera que el hecho de haber externalizado en una experta la gestión de las evaluaciones, resta utilidad para el aprendizaje del personal funcionario, quedando ésta en la motivación personal de cada trabajador.

Consideran también que las principales recomendaciones van en la línea de comprender mejor la propia intervención, mejorar los procesos y mejorar los sistemas de recogida de información y en ese sentido son las ONGD y los SL los responsables de esta tarea.

Respecto a **las limitaciones** señaladas, tras haber respondido de forma afirmativa a la pregunta de si hubo *condicionantes* importantes para alcanzar los resultados previstos en la evaluación se analizaron las posibles causas. De las 28 ONGD que responden, 20 perciben que sí hubo limitaciones (el 71,4%) mientras que 8 responden negativamente. La mayoría de los condicionantes se ha ido recogiendo a lo largo del texto cuando se han analizado los diversos apartados, sin embargo, a modo de ejemplo, mencionamos a continuación algunos de

ellos: falta de dominio del idioma local por parte del equipo evaluador (quechua, español, etc.), el momento era el marcado por la normativa pero no era el adecuado, la metodología no es la adecuada para un convenio regional (se dan recomendaciones por país pero no por región), el hecho de que se obligue (aunque esto no sea recogido por normativa) a evaluar bajo los criterios del CAD no permite dar tanta importancia a los procesos que en una evaluación intermedia debería ser el elemento central, el tiempo que se pasa en terreno por parte del equipo evaluador es muy escaso, etc.

El último bloque en torno a la utilidad-credibilidad de la evaluación gira en torno a la percepción de **los recursos empleados**. Bajo una escala ordinal de cuatro elementos (insuficiente-suficiente-proporcional-excesivo) codificada entre 0 (insuficiente) y 4 (excesivo) se preguntó sobre el tiempo empleado, los recursos humanos (RR.HH.) involucrados y el dinero invertido.

Los estadísticos de tendencia central ofrecen valoraciones muy similares: el promedio de tiempo es 2,46 (con $\sigma=0,92$); de RR.HH. es 2,60 ($\sigma=0,73$); y de dinero 2,55 ($\sigma=0,97$). Las medias de dispersión son, quizá, más informativas que el promedio. Hay más diversidad de opinión en valorar el dinero, seguida del tiempo. Más concentrado aparece el juicio sobre los RR.HH.

Respecto al *tiempo*, la distribución de frecuencias contiene 4 casos en los que se optó por los valores más extremos. Cuatro ONGD consideran que el tiempo fue excesivo y cuatro que fue insuficiente. La mayoría (39,3%) lo valora “suficiente”, seguido muy de cerca por “proporcional” (32,1%). La información cualitativa recogió algunas quejas sobre la limitación del tiempo que habitualmente los evaluadores emplean en el trabajo de campo, opinión corroborada por la percepción de AECID, lo que limita mucho la aplicación de diversas técnicas para la recogida de información y una organización también subrayó la falta de tiempo dedicado al trabajo de gabinete.

Con respecto al tiempo también existe la percepción generalizada de que es escaso no para realizar la evaluación pero sí para procesar el aprendizaje obtenido y para implementar mejoras al quedar poco tiempo de vida del convenio una vez obtenido el informe de evaluación.

Por último, también se manifiesta de forma generalizada que el tiempo empleado por el personal ha sido excesivo en la negociación de los TdR con AECID.

Tabla 6. Tiempo.

tiempo

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1,00 (insuficiente)	4	14,3	14,3	14,3
	2,00 (suficiente)	11	39,3	39,3	53,6
	3,00 (proporcional)	9	32,1	32,1	85,7
	4,00 (excesivo)	4	14,3	14,3	100,0
	Total	28	100,0	100,0	

Con respecto a los *RR.HH.* se mantiene la proporción de las organizaciones que valoran el esfuerzo como “excesivo”, pero hay una dominancia notable en torno a “suficiente” (el 53,6%).

Tabla 7. Recursos Humanos.

RRHH

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	2,00 (suficiente)	15	53,6	53,6	53,6
	3,00 (proporcional)	9	32,1	32,1	85,7
	4,00 (Excesivo)	4	14,3	14,3	100,0
	Total	28	100,0	100,0	

Por último, respecto a los *recursos monetarios*, se mantiene el número de valoraciones de insuficiente en cuatro casos, excesivo acumula casi el 18% (5 casos), y la dominancia se centra entre suficiente 28,6% y proporcional (35,7%). No parece que la percepción del coste de las evaluaciones haya sido excesiva, aunque de forma cualitativa sí hay testimonios de que el presupuesto para llevar a cabo las evaluaciones fue reducido. De forma cualitativa hubo quejas de escasez de recursos económicos, pero relacionadas con la cantidad de preguntas evaluativas formuladas, fruto de no poder seleccionar (reducir) los criterios de evaluación (los cinco del CAD), de hecho algunas ONGD optaron por realizar evaluaciones finales con recursos propios y éstas fueron dotadas de un mayor presupuesto que las intermedias.

En opinión de la AECID, se considera que los recursos económicos para hacer las evaluaciones, especialmente en los convenios de la primera convocatoria, la del 2006, fueron insuficientes, lo cual condicionó el alcance de las evaluaciones. También, al igual que las ONGD, considera que los recursos son insuficientes en relación a la cantidad de preguntas y dimensiones a evaluar en los convenios, que resulta demasiado ambicioso.

Tabla 8. Dinero invertido.

Dinero

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1,00 (insuficiente)	4	14,3	14,3	14,3
	1,50*	1	3,6	3,6	17,9
	2,00 (suficiente)	8	28,6	28,6	46,4
	3,00 (proporcional)	10	35,7	35,7	82,1
	4,00 (excesivo)	5	17,9	17,9	100,0
	Total	28	100,0	100,0	

Nota: la valoración 1,5 puntos fue asignada al cuestionario que optó por responder insuficiente-suficiente sin discriminar hacia una de ellas.

En resumen, no se deduce de las respuestas una falta de recursos que correlacione con la utilidad-credibilidad de la evaluación, sin embargo, recogemos un testimonio que agrupa la percepción de los tres tipos de recursos utilizados:

“El dinero para hacer la evaluación, determina la composición del equipo evaluador, y el tiempo que se emplea para recoger información en terreno, son dos cuestiones básicas para hacer un trabajo de calidad”.

2.6. Difusión.

La última sección del cuestionario preguntó sobre la **difusión** del informe de evaluación. El medio que se proponía era la página web de la ONGD y se daba libertad para revelar otros. De las 28 ONGD, sólo 6 (21%) han respondido que han difundido de alguna forma su evaluación. Este tema fue ampliamente debatido en el grupo de discusión y se obtuvieron las siguientes explicaciones al porcentaje tan bajo de difusión recogido en el análisis cuantitativo. Parece que hay un común acuerdo que cultura evaluativa organizacional y difusión van de la mano. En opinión de la mayoría de las ONGD, esta cultura evaluativa se está generando (y en muchos casos a raíz del mecanismo de evaluación inducido por AECID a través de la normativa de obligatoriedad para los Convenios) y consecuencia de ello es que hay varias

ONGD (precisamente las que parecen tener mayor cultura evaluativa: personas contratadas exclusivamente para labores de evaluación, personal formado específicamente en el tema, mayor número de evaluaciones realizadas, etc.) que ya están integrando la difusión de las evaluaciones en sus Planes organizativos y como tema de reflexión institucional. Algunos comentarios reflejan este hecho:

“Antes ni nos lo planteábamos, ahora con la cultura evaluativa que se está implantando ya sí se ha recogido en el Plan Estratégico”;

“Era algo nuestro, hoy en día nos hemos propuesto tener un plan de difusión para 2012);

“Si los convenios se ejecutan con fondos públicos y las evaluaciones también, éstas deberían ser un bien público”;

“El departamento de evaluación es nuevo, de reciente creación, y todavía estamos elaborando los procedimientos y precisamente uno de ellos es la difusión y el uso de las evaluaciones”.

En línea con este último comentario, también hay ONGD que muestran que la falta de cultura evaluativa en sus SL les limita para poder publicar las evaluaciones ya que estos últimos no autorizan a ello:

“Se quiso publicar una evaluación en la red de los propios SL porque había sido muy buena, pero el propio SL se opuso a ello”.

Sin embargo, también se recogieron comentarios que reflejan el temor a publicar las evaluaciones:

“Es un problema estructural del propio sistema. En un sistema que dependes de la subvención no te puedes arriesgar a publicar una nota negativa”.

Y, también hay opiniones sobre que la propia falta de calidad de los informes limita el hecho de publicarlos:

“Para que una evaluación sea útil para otras organizaciones ajenas a la intervención evaluada ha de ser bueno, conciso, constructivo, original, y está claro que normalmente no es así”.

2.7. Análisis por correlaciones.

Para un análisis de mayor profundidad pero sin pretensión de abordar causalidades, elaboramos cinco nuevas variables ficticias a partir de la suma simple de las frecuencias obtenidas en los cuestionarios. Las nuevas variables y su recorrido son:

- Participación de la ONGD: recorrido [0-9] por la suma de los nueve ítems en los que se señaló la participación a lo largo del ciclo de la evaluación;
- Participación del socio local: análoga a la anterior mediante la suma de frecuencias del socio local durante el proceso evaluativo;
- Suma de Utilidad: la suma de los tres ítems sobre la utilidad (proceso, conclusiones y recomendaciones); recorrido [3-12].
- Suma Equipo: suma de los nueve ítems que calificaban la acción del equipo evaluador; recorrido [9-36].
- Suma recursos: los tres ítems que preguntaron por tiempo, recursos humanos y dinero; recorrido [3-14].

Añadiendo a estas variables la del “Valor” global otorgada por la ONGD entre cero y diez podemos obtener información relevante para intentar explicar de qué ha sido función la utilidad de la evaluación de los convenios y qué no la explica.

Un primer acercamiento de las interrelaciones entre estas variables lo ofrece la matriz de correlaciones que se detalla en la Tabla 9.

Tabla 9. Matriz de correlaciones en torno a la utilidad del proceso evaluativo.

	Valor	PART_ONGD	PART_SL	SUMA_equipo	SUMA_Util	SUMA_Recursos
Valor	1					
PART_ONGD	-0,030	1				
PART_SL	0,289	-0,005	1			
SUMA_equipo	0,923	0,098	0,228	1		
SUMA_Util	0,689	-0,078	0,243	0,708	1	
SUMA_Recursos	-0,197	0,078	-0,383	-0,115	-0,178	1

La asociación estadística más fuerte se produce entre el *valor* que subjetivamente otorgan al proceso las ONGD y la percepción que tienen del trabajo de los evaluadores (0.92) y de la utilidad (0.69). También es alta la correlación entre la utilidad y el equipo (0,71). El resto de variables ofrecen niveles bajos de correlación. Esta correlación mostrada por el análisis

cuantitativo, se fortalece con la información obtenida del análisis cualitativo en el que las ONGD muestran reiteradamente como una de las debilidades más importantes del proceso, (los cuellos de botella) a los propios equipos evaluadores, mencionando aspectos como: la falta de dominio de la metodología de evaluación y el uso de diversas técnicas de recogida de información, la escasez de tiempo dedicado en terreno a la recogida de información, falta de entendimiento entre los miembros de un mismo equipo, evaluadores no expertos en la temática o sector de evaluación, entre los principales.

Quizá pueda resultar de interés la correlación negativa entre la participación de los socios locales y los recursos invertidos (-0.38) pues más bien podría ser al revés: una mayor participación de los socios locales habría podido aumentar el tiempo, los RR.HH. y el coste monetario de la evaluación.

El *valor* atribuido al proceso de evaluación de los convenios no se explica, del análisis cuantitativo, por la participación de la propia ONGD en el proceso (-0.03) ni por los recursos invertidos que incluso tienen una correlación levemente negativa (-0.19). Sin embargo, de la información cualitativa sí es importante resaltar la importancia que las ONGD dieron a todo el proceso participativo para que la evaluación fuera realmente útil. De hecho, en el grupo de discusión hubo unanimidad en cuanto a la que la participación de la ONGD desde el principio del proceso era clave para evitar cuellos de botella. Hubo un comentario que resumiría esta percepción: *“Si hay participación y un buen equipo de evaluación, todo lo demás viene dado”*.

Si, a modo de intento de modelización, regresamos la **utilidad** sobre el resto de las variables ficticias obtenemos los siguientes resultados (Tabla 10) a partir de la ecuación:

$$\text{Suma_Util}_i = \alpha + \beta \text{Valor}_i + \chi \text{Part_ONGD}_i + \delta \text{Part_SL}_i + \varepsilon \text{Suma_equipo}_i + \phi \text{Suma_recursos}_i + \gamma_i$$

Ni el *valor* subjetivamente asignado, ni la *participación de los socios locales* aparecen estadísticamente asociadas a la utilidad. El único factor explicativo significativo al 99% de confianza es el *equipo* evaluador. Para conocer de forma más detallada cuál de las nueve variables que componen la “Suma_equipo” se procedió a hacer la regresión de ésta sobre los nueve componentes. La única que resultó estadísticamente significativa fue “*equipo*”.

Lo que de nuevo claramente se sustenta con todo el análisis cualitativo de la información, en donde las ONGD han puesto el peso de la utilidad de la evaluación en la composición de buenos equipos evaluadores.

Tabla 10. Regresiones sobre la utilidad del proceso evaluativo.

	1	2	3
Constante	4,07	4,02	4,18
	<i>0,067</i>	<i>0,061</i>	<i>0,043</i>
Valor	0,08		
	<i>0,873</i>		
Part_ONGD	-0,13	-0,13	-0,137
	<i>0,404</i>	<i>0,342</i>	<i>0,334</i>
Part_SL	0,03	0,03	
	<i>0,749</i>	<i>0,723</i>	
Suma_equipo	0,26	0,291***	0,295***
	<i>0,14</i>	<i>0,000</i>	<i>0,000</i>
Suma_recursos	-0,04	-0,05	-0,06
	<i>0,720</i>	<i>0,691</i>	<i>0,563</i>
R2	0,5339	0,5333	0,530
F	4,811**	6,286**	8,666**
Observaciones	28	28	28

Nota: Var dependiente: SUMA_Util; en cursiva los p-values. *** significativa al 99% y ** al 95%.

Resultados muy similares se obtienen considerando el **valor** subjetivo del ejercicio evaluativo como variable dependiente. Como muestra la Tabla 11, únicamente la suma de las frecuencias otorgadas a los ítems que medían al equipo evaluador aparece como estadísticamente explicativa del valor percibido.

Tabla 11. Regresiones sobre el valor de las evaluaciones realizadas.

	1	2	3
Constante	-0,74	-0,68	-0,55
	<i>0,457</i>	<i>0,449</i>	<i>0,526</i>
Suma utilidad	0,015		
	<i>0,873</i>		
Part_ONGD	-0,09	-0,094	-0,09
	<i>0,169</i>	<i>0,142</i>	<i>0,138</i>
Part_SL	0,02	0,029	
	<i>0,519</i>	<i>0,5</i>	
Suma_equipo	0,317***	0,322***	0,326***
	<i>0,000</i>	<i>0,000</i>	<i>0,000</i>
Suma_recursos	-0,04	-0,04	-0,05
	<i>0,48</i>	<i>0,459</i>	<i>0,285</i>
R2	0,8762	0,8761	0,8734
F	29,75**	38,9**	52,93**
Observaciones	28	28	28

Nota: Var dependiente: Valor; en cursiva los p-values. *** significativa al 99% y ** al 95%.

Análogamente a como se hizo para la utilidad, se calculó la regresión del valor sobre los 9 ítems de “suma_equipo”. Esta vez los componentes significativos fueron “ajuste” entre el trabajo finalmente hecho por el evaluador y la oferta inicial, el equipo y la comunicación con la ONGD.

Tabla 12. Regresión del “valor” sobre los ítems del trabajo del equipo evaluador.

		Parte del R ²
constante	0,21	0,0062
	<i>0,703</i>	
comunicación	0,33*	0,1162
	<i>0,088</i>	
equipo	0,84**	0,3267
	<i>0,002</i>	
ajuste	1,07***	0,4424
	<i>0,000</i>	
R2	0,8618	
F	49,91**	
Observaciones	28	

Nota: Var dependiente: Valor; en cursiva los p-values. *** significativa al 99%, ** al 95% y * al 90%.

En resumen, según la experiencia concreta tenida en la evaluación de los convenios 2006-2007, y de acuerdo a las valoraciones otorgadas por 27 ONGD, el proceso ha resultado útil gracias al desempeño de los equipos evaluadores. Esta utilidad no está estadísticamente asociada al valor percibido por las ONGD, ni a la participación de los socios locales, ni de la ONGD, ni a los recursos invertidos.

Por último, llama la atención la falta de significatividad de la participación del socio local. Las ONGD son defensoras de un modelo de cooperación inclusivo y participativo, pero en las evaluaciones aún no se ha traducido en una práctica común. Por ejemplo, dos entidades afirmaron que los socios locales no habían participado en ninguna de las fases de la evaluación y otra que participaron sólo en los TdR.

V. INTERPRETACIÓN DE LOS HALLAZGOS Y ENJUICIAMIENTO

En esta sección se recogen de forma sumaria las principales conclusiones obtenidas en base a los juicios de valor del equipo de evaluación una vez analizada la información recogida de forma primaria.

Esta interpretación de los hallazgos trata de exponer las principales relaciones causales entre los elementos estructurales y los subprocesos con los resultados obtenidos. La interpretación se ha centrado especialmente en identificar las fortalezas y debilidades, tanto de los elementos estructurales como de los subprocesos, y analizar cómo éstas influyen en los resultados obtenidos

Elementos estructurales:

La propia normativa y su carácter de obligatoriedad: La mayoría de las ONGD considera muy positivo que haya habido un impulso para realizar evaluaciones intermedias. Una externalidad positiva fruto de esta iniciativa es que algunas de éstas organizaciones están ahora llevando a cabo evaluaciones finales por propia iniciativa y con autofinanciación. Es decir, se podría afirmar, que la obligatoriedad de la normativa ha servido para generar mayor cultura evaluativa entre las ONGD (de hecho, hay organizaciones que empiezan a contar con personal especializado y/o departamento de evaluación al interior de sus organizaciones, a raíz de la puesta en práctica de esta normativa). Luego, la normativa influyó positivamente en un resultado que no había sido previsto.

Por otro lado, la normativa establecía que la evaluación debía hacerse con carácter intermedio. Esta obligatoriedad de hacer evaluaciones intermedias ha podido restar utilidad para la propia intervención y la ONGD. Hay convenios que están pensados como una primera fase de un proceso de más largo plazo. En estos casos una evaluación al final del primer convenio (en vez de la intermedia) hubiera sido de mayor utilidad.

Sobre los recursos utilizados. No se deduce del análisis cuantitativo una falta de recursos (económicos, humanos y de tiempo) que correlacione significativamente con la utilidad-credibilidad de la evaluación. Sin embargo, desde un punto de vista cualitativo surgen reflexiones interesantes, tanto por parte de las ONGD como por parte de la AECID, sobre los condicionantes de los recursos para obtener trabajos de calidad. Un testimonio, que podría representar esta reflexión, afirmó que “el dinero para hacer la evaluación parece determinar la composición del equipo evaluador, y el tiempo que se

emplea para recoger información en terreno. Estas son dos cuestiones básicas para hacer un trabajo de calidad”.

También existe la percepción generalizada de que el tiempo es escaso no para realizar la evaluación en sí, sino para procesar el aprendizaje obtenido y para implementar mejoras al quedar poco tiempo de vida del convenio una vez obtenido el informe de evaluación. Por último, también se manifiesta de forma generalizada que el tiempo empleado por el personal de la ONGD ha llegado a ser excesivo sobre todo durante la negociación de los TdR con AECID.

En relación a la profesionalidad de los evaluadores externos. La principal debilidad que apuntaron las ONGD en cuanto a la profesionalidad del equipo evaluador era la dificultad de encontrar verdaderos expertos en evaluación con amplios conocimientos metodológicos.

Enjuiciamiento: La creación y puesta en marcha de la norma ha influido positivamente en la cadena de resultados, generando cultura evaluativa entre otros. Sin embargo, la norma condiciona el grado de utilidad de las evaluaciones en el momento en que se determina el carácter obligatorio intermedio de las mismas.

De las variables estudiadas en esta investigación como endógenas a la utilidad de la evaluación, tanto el análisis cuantitativo como cualitativo revela que el propio equipo de evaluadores es el mayor condicionante para que la utilidad sea alta para las ONGD, siendo un aspecto muy importante el ajuste entre el trabajo finalmente hecho y la oferta inicial realizada por los evaluadores.

Igualmente, los recursos económicos condicionan la composición de dicho equipo evaluador que finalmente es el que determina en mayor medida el grado de utilidad de las evaluaciones.

Es decir, los recursos (norma y sus condicionantes, recursos humanos, dinero) son determinantes en cuanto a los resultados obtenidos (mayor o menor utilidad de las evaluaciones realizadas) en cuanto que condicionan los subprocesos, especialmente el de externalización de la evaluación.

Procesos

Tal y como se puede observar en la graficación de la teoría del Programa expuesta más arriba, el equipo evaluador determinó la existencia de cuatro principales subprocesos del único proceso estudiado, la evaluación de los Convenios entre ONGD y AECID. A

continuación se trata de interpretar, en base a la información recogida y analizada, la relación de estos con los resultados.

1.- Diseño de la Evaluación

La mayoría de las ONGD ha visto en las evaluaciones una posibilidad de aprendizaje y de mejora más allá de la obligatoriedad de la norma. Sin embargo, existe una disconformidad entre las ONGD y AECID en la fase de elaboración de los TdR ya que los intereses evaluativos de la ONGD no siempre se correspondían con los de la AECID. Algunas ONGD han vivido los intereses de la AECID como “imposición” lo que ha restado utilidad para las Organizaciones. La apreciación de las ONGD es que la participación de la AECID en esta fase ha generado procesos demasiado largos, generado poco valor añadido y produciendo cierto “desgaste” en la mayoría de las Organizaciones, debido al tiempo que ha habido que emplear en esta fase.

La AECID, por su parte, ha tratado de que sus intereses se vieran reflejados en las evaluaciones a través de su participación en la definición de los TdR y reconoce que, en ocasiones, han tenido que ser puntillosos especialmente cuando la marcha del convenio no era buena.

Enjuiciamiento: En general el tiempo empleado en el diseño de la evaluación ha sido muy largo y poco eficaz debido a las largas negociaciones entre ONGD y AECID para reflejar los intereses de unos y otros. Finalmente, los TdR han resultado ser demasiado ambiciosos, poco realistas e incluso “oficiosos”. Esto ha restado utilidad a las evaluaciones al no centrarse en los aspectos clave a ser evaluados (los más importantes o sobre los que había mayor incertidumbre). Por otro lado, aunque los documentos de evaluación no han sido objeto de esta evaluación, parece existir un claro consenso entre ONGD y AECID en que es necesario un trabajo más reflexivo y profundo por parte de las ONGD y sus SL para la concreción de las preguntas de evaluación reflejadas en los TdR. Es decir, el grado de utilidad obtenido de las evaluaciones podría incrementarse considerablemente si los TdR fueran más realistas y concretos especialmente a la hora de formular las preguntas de evaluación a las que se quiere dar respuesta.

2.- Externalización de la evaluación

¿Evaluación externa o mixta?: Como primera conclusión, puede señalarse que las ONGD, optan por hacer un concurso público para la búsqueda de la empresa evaluadora pero con cierta “restricción-invitación expresa” a empresas con las que ya se ha trabajado o se han obtenido buenas referencias, por lo tanto, en el proceso de elección se combina el factor de transparencia con el factor de confianza debido a los antecedentes.

Como segunda conclusión de este punto, cabe mencionar que para las ONGD, la función principal de la evaluación del convenio era el aprendizaje y no la rendición de cuentas. Es por ello que consideran que una evaluación mixta hubiera permitido un mayor aprendizaje y nivel de apropiación. Los equipos mixtos pueden generar sinergias ya que la ONGD y su SL pueden aportar de forma más eficiente a los evaluadores externos el conocimiento sobre el país, sobre el contexto de ejecución y del sector de la intervención. Por su parte, los evaluadores externos pueden aportar mayor conocimiento evaluativo (metodología y herramientas) y contraste con otras experiencias comparables que quizá la ONGD y el SL no tengan.

De la entrevista realizada a la AECID se deduce que la principal restricción para no permitir evaluaciones mixtas es un motivo de procedencia de los fondos con los que es financiada. Se percibe una apertura de la AECID a permitir evaluaciones mixtas siempre que las ONGD las cofinancien con sus fondos.

Exigencia de evaluadores locales. Sobre este punto se obtienen dos conclusiones.

En primer lugar, mencionar que en convenios multi-país, este requisito no ha generado los frutos esperados ya que generalmente se ha perdido la mirada regional del convenio. El supuesto de que los locales pueden tener valor añadido en cuanto a conocer mejor el contexto regional, no se ajusta a la realidad ya que un evaluador de origen de uno de los países del convenio no siempre posee un conocimiento en profundidad del resto de los países del convenio. Esta opinión parece ser compartida tanto por las ONGD como por la AECID que admite que para este tipo de convenio se mostró mucho más flexible con respecto a este criterio.

En segundo lugar, otra dificultad encontrada ha sido la falta de conocimiento del evaluador local del sistema de la Cooperación Española, dando origen a recomendaciones poco prácticas o contrarias a la normativa española vigente. Respecto a este punto, en opinión de las ONGD este factor ha determinado la menor utilidad de la evaluación. Sin embargo, para la AECID, aunque no esté explícito en normativa, prima la promoción de los recursos locales aunque ésta vaya algo en detrimento de la utilidad de la evaluación.

Otro tema relevante de este proceso de externalización es que la contratación de una empresa determinada no es garantía de buen trabajo ya que a veces éstas suelen subcontratar estos trabajos. En definitiva, la calidad del equipo evaluador lo aportan las

propias personas que lo componen y todavía no se identifican empresas con suficiente marca que las garantice un buen producto evaluativo.

Enjuiciamiento: En opinión de este equipo evaluador, quien finalmente pone en práctica la mayoría de las recomendaciones es el SL con el apoyo de la ONGD, con lo que si su participación se traduce en mayor motivación y conocimiento fundado de las conclusiones que originan dichas recomendaciones, su ejecución real y transformativa puede ser mayor, y la evaluación puede ser más influyente y útil. Es decir, una mayor participación de SL y ONGD en todos los subprocesos es determinante en el logro de mayor utilidad, es decir, en el logro de los resultados.

No existen buenos y malos evaluadores por ser locales o nacionales. También existe el dilema de priorizar expertos en evaluación o expertos en el sector que se evalúa. La experiencia tenida en estos convenios apunta a la priorización de expertos en evaluación tratando de compensar las limitaciones en el conocimiento específico de la temática por otras vías.

La casuística de cada intervención es la que debe determinar cuál la mejor composición del equipo evaluador (recursos humanos) lo cual deberá reflejarse en el proceso de elección de la empresa-equipo de evaluación (procesos) para el logro de una mayor utilidad de la evaluación (resultados).

3.- Ejecución de las Evaluaciones

Participación de las ONGD. Tanto entre las organizaciones con mayor cultura evaluativa como entre las ONGD que empiezan a interesarse por la evaluación, la participación aparece como un elemento clave en su discurso, que es corroborado con los datos del análisis cuantitativo. Destacan sobre todo la necesidad de un alto nivel de participación en las primeras fases del proceso a la hora de definir las preguntas de evaluación y los TdR. Debido a que eran evaluaciones externas, el nivel de participación es menor en las fases de recogida e interpretación de la información así como en la elaboración de conclusiones y recomendaciones. Sin embargo, en esta última fase, las ONGD reconocen que depende de la flexibilidad del propio equipo evaluador el que se tengan en mayor o menor consideración las valoraciones y opiniones de las ONGD, convirtiéndose en algunos casos en una ardua negociación ante la diversidad de interpretaciones.

La participación de los socios locales es relativamente baja. En opinión de las ONGD, esta menor participación se debe a la falta de interés porque la evaluación no se concibe

como una oportunidad de mejora sino como un trámite más, una imposición, debido a la falta de cultura evaluativa en sus organizaciones.

La menor satisfacción de las ONGD sobre el trabajo de los evaluadores se muestra en la fase de recogida de información, tanto en el tiempo empleado como en las técnicas empleadas para ello, y en que la evaluación haya sido completa. Se aprecia una tendencia a recoger y resaltar lo negativo y no identificar o señalar también las fortalezas de las intervenciones.

Para la AECID la menor satisfacción se muestra en la justificación de la metodología elegida y la justificación de cada una de las técnicas de recogida de información aplicadas, además de falta de profesionalidad en la construcción de matrices de evaluación, elaboración de adecuados criterios de evaluación y redacción de recomendaciones adecuadas.

La mayor satisfacción se muestra en la comunicación mantenida entre los equipos evaluadores y las ONGD.

Enjuiciamiento: El análisis cualitativo, refleja como significativa en la utilidad de la evaluación, la amplitud y profundidad de la participación de la propia ONGD en todo el proceso de evaluación. No se refleja, por el contrario, la participación de los Socios Locales, como elemento clave para reportar mayor utilidad al proceso de evaluación lo que demuestra que el modelo inclusivo y participativo por el que abogan las ONGD no se pone en práctica en el momento de realizar las evaluaciones, ya sea por la falta de interés y cultura evaluativa de algunos de los socios locales, ya sea por la falta de tiempo e interés por generar un proceso de reflexión y formación en materia de evaluación por parte de las ONGD entre sus SL.

El trabajo realizado por los evaluadores, en términos generales, es calificado de bueno por las ONGD (aunque la cuarta parte de las ONGD califican este trabajo como regular). En general, la valoración global sobre el trabajo realizado se calificaría de notable (promedio de 6,5 y moda de 7) en una escala del 1-10. Mientras que la AECID la calificaría más bajo, en un “suficiente” ó 5. En opinión del presente equipo evaluador, esta diferencia de opinión se puede deber al diferente criterio utilizado por unos y otros para calificar el trabajo hecho por los equipos evaluadores, mientras que las ONGD se basan en el aprendizaje obtenido y la utilidad del proceso para calificarlo, la AECID se estaría basando en los criterios establecidos por la UE para la meta evaluación de evaluaciones.

4.- Difusión

Los datos muestran que sólo 6 ONGD (21%) han respondido que han difundido de alguna forma su evaluación. De hecho muchas ONGD contestaron que éste era un tema que aún no se había debatido o reflexionado suficientemente al interior de su organización.

Enjuiciamiento: El equipo evaluador considera que el principal motivo de la falta de publicación parece ser la falta de cultura evaluativa (a mayor cultura evaluativa, mayor difusión de los resultados de las evaluaciones o de las propias evaluaciones). Aunque se recogen otras opiniones como la falta de calidad de los informes de evaluación y el temor de la publicación de malos resultados, estas opiniones son minoritarias.

Resultados

Como se ha expresado, en la construcción de la teoría del programa que fue la base para la realización de la evaluación, la dimensión de resultados se centró tan solo en la utilidad que las evaluaciones habían tenido, dejando esta dimensión abierta para obtener más información de la misma a través de las técnicas cualitativas de recogida de información.

La percepción revelada sobre la utilidad tanto del proceso de evaluación como de las conclusiones obtenidas y las recomendaciones hechas a cada informe de evaluación, parece ser alta o muy alta en la mayoría de las ONGD.

Los argumentos más repetidos entre las ONGD sobre los motivos que han hecho útiles las evaluaciones son los siguientes:

El proceso ha servido para conocer algo más sobre la metodología de evaluación.

Las evaluaciones han servido para conocer en profundidad las intervenciones y darse cuenta que éstas muchas veces se plantean de manera ambiciosa (muchos sectores, muchos países, muchos socios locales, etc.)

También han servido para mejorar mucho los sistemas de seguimiento. Indicadores poco SMART y fuentes de verificación que no aportaban la información necesaria para la medición del seguimiento.

Por último, pero no por ello menos importante, han servido para mejorar las propias intervenciones aplicando las recomendaciones obtenidas.

La opinión expresada por la AECID corrobora estas causas de utilidad, de hecho hace hincapié en la falta de calidad de los sistemas de seguimiento establecidos y la gran mejora que ha supuesto el desarrollo de la evaluación para ello. Igualmente, la AECID señala que las evaluaciones han sido útiles a modo de rendición de cuentas por parte de las ONGD sobre la gestión de los fondos públicos. En este sentido se expresa que

aproximadamente el 25% del presupuesto de la AECID es gestionado por las ONGD y no existe capacidad ni recursos humanos en la AECID para visitar todas estas intervenciones, por lo que la evaluación ha resultado de utilidad a modo de transparencia y rendición de cuentas desde las ONGD hacia la propia AECID.

Esta valoración sobre la utilidad conduce a querer conocer *las consecuencias* concretas que ha tenido la evaluación sobre las intervenciones de las ONGD. La primera consecuencia es contar con un plan de mejora. La mayoría de las ONGD afirma que ya cuenta con él o están en proceso de elaboración. Muchas de las ONGD afirman que su plan de mejora es más completo y de mayor alcance que la adenda firmada con la AECID. Además, el 93% de las ONGD afirman que se han obtenido recomendaciones útiles para ser aplicadas a otras intervenciones de la organización (replicabilidad) además de a la continuidad de los convenios que tuvieron una segunda fase.

En cuanto *al agente para el que la evaluación ha sido más útil*, las ONGD perciben (al igual que la AECID) que ellas mismas son el principal beneficiario de la evaluación, seguidos de sus socios locales, quedando en último término la propia AECID y los beneficiarios, aunque la mayoría reconoce que de manera indirecta, en la medida en que se mejoran los procesos, la evaluación también resulta útil para los beneficiarios últimos. Las ONGD perciben que la AECID no tiene capacidad para asimilar todas las evaluaciones realizadas y por tanto esto restaría utilidad de las mismas para este agente.

Enjuiciamiento: Se podría concluir que el conjunto de las evaluaciones en su mayoría han sido útiles dando cumplimiento a las tres funciones naturales de la evaluación: la comprensión de las intervenciones y la iluminación para la mejora de las mismas y futuras intervenciones y la rendición de cuentas.

Podríamos, ahora, dibujar la cadena causal de los objetivos logrados con el desarrollo de la política de evaluación en las ONGD.

En un nivel más bajo de la relación causal, podríamos afirmar que las evaluaciones han sido útiles porque han generado tres tipos de productos: En primer lugar han supuesto un **aprendizaje** para las ONGD. Para muchas de ellas, ésta era la primera experiencia de evaluación y solo el hecho de evaluar ya ha supuesto para ellas un importante aprendizaje del propio proceso. Por la misma razón se podría decir que se ha producido un despertar de la **cultura evaluativa** (toma de conciencia de la necesidad de evaluar por la utilidad que puede llegar a reportar), y por supuesto los **planes de mejora** con los que cuentan la mayoría de las ONGD es el principal y más importante producto que

demuestra la utilidad de las intervenciones aunque el alcance de esta evaluación no permite conocer la aplicación de dichos planes de mejora.

En un segundo nivel, podríamos decir que los resultados esperados nos dicen que las evaluaciones han sido **útiles**. Útiles para las ONGD por todo lo expresado, y útiles para el financiador, la AECID a modo de rendición de cuentas.

Por último, y como objetivo más elevado en la relación causal de estos, no podemos dar respuesta a si las evaluaciones al incorporarse en el ciclo del proyecto a través de los planes de mejora, han logrado **el incremento de la calidad de las intervenciones**. Todo a punto a pensar que sí es así pero no ha sido objeto de este estudio demostrar empíricamente este último escalón de la relación causal de la cadena de objetivos.

Interpretación holística del evaluando

A modo de resumen de este capítulo, se ofrece una interpretación de las relaciones existentes entre los diferentes elementos (estructura, procesos y resultados) del evaluando como un todo. Para ello, se analiza la manera en que los elementos estructurales influyen en mayor o menor grado en los procesos y a su vez cómo estos son determinantes de los resultados obtenidos.

Del análisis de los **elementos estructurales**, se desprende que estos han sido suficientes y determinantes para el buen desarrollo de los procesos, aunque en ocasiones también han sido limitados y/o limitantes para los procesos. El elemento estructural más relevante en cuanto a haber condicionado los procesos y que estos hayan determinado los resultados, ha sido la propia normativa que regulaba la forma en que debían hacerse las evaluaciones (externas, intermedias y preferentemente realizadas por personal local). Dependiendo de la casuística de cada convenio, estas obligaciones marcadas por la normativa han condicionado los procesos (diseño de los TdR, contratación de la evaluación y la propia ejecución) en cuanto la elección del equipo evaluador, la mayor o menor participación de las ONGD y sus SL en los procesos..., repercutiendo claramente en los resultados en cuanto a mayor o menor utilidad obtenida en los procesos de evaluación. Como se ha enjuiciado en páginas más arriba, hay elementos estructurales que podrían mejorarse para obtener mejores resultados (principalmente se señala la revisión de algunos aspectos de la normativa y la existencia de recursos humanos-evaluadores y comanditarios de evaluaciones-más formados en metodología evaluativa), pero a grosso modo se puede afirmar que la normativa ha supuesto una obligación que se ha convertido en oportunidad en cuanto a los resultados logrados: las evaluaciones

han sido útiles pero además se ha despertado una cultura evaluativa entre las ONGD que antes existía en muy pocas de ellas.

Con respecto a los **procesos**, el aspecto positivo más determinante de los resultados obtenidos, ha sido el papel de liderazgo que han asumido las ONGD (impulsado desde la normativa) en el diseño de los TdR, la contratación (externalización) de la evaluación y el seguimiento de la ejecución de la misma. Este liderazgo ha influido claramente en el nivel de satisfacción-grado de utilidad-obtenido de las evaluaciones, y en el despertar de una cultura evaluativa para muchas de ellas. Sin embargo, se echa en falta un papel más protagonista de los SL en cuanto a su participación en los diferentes procesos analizados y, dado que son estos los que deben implementar las mejoras de las intervenciones realizadas, este aspecto debería considerarse prioritariamente para el mayor logro de utilidad de las evaluaciones realizadas.

Existen otros aspectos mejorables que influirían también en un mayor logro de los resultados. Por ejemplo, el proceso de elaboración de TdR es sustancialmente mejorable. Se requieren TdR claros y concretos en cuanto a las preguntas de evaluación a plantear, así como en cuanto a las dimensiones a ser evaluadas, para generar mayor utilidad de las mismas. Las necesidades informativas que se requieran deben guardar coherencia con los tiempos y los presupuestos que se establezcan en los TdR. Por último, se deberían buscar canales de comunicación más ágiles que permitan incorporar los intereses evaluativos de los diferentes actores en un menor tiempo de negociación de los TdR. Las mejoras que aquí se proponen permitirían alcanzar mayores niveles de utilidad de las evaluaciones.

Del proceso de la externalización-contratación de la evaluación, lo más significativo en cuanto a su influencia en el grado de utilidad, sería la composición del equipo evaluador lo que dependerá en último caso de la casuística de cada convenio, pero en todo caso se ha demostrado que es importante contar en el equipo con expertos en metodología de evaluación ya que este aspecto ha sido uno de los peores valorados tanto por las ONGD como por la AECID en cuanto a la composición de los equipos evaluadores, y que ha repercutido claramente en el grado de credibilidad de los resultados obtenidos en las evaluaciones y por tanto en la utilidad de las mismas.

El proceso de difusión pública de las evaluaciones es necesariamente mejorable ya que es prácticamente inexistente. Pareciera lógico pensar que este proceso está íntimamente ligado a la cultura existente al interior de las ONGD sobre la importancia que tiene la evaluación. Muchas de ellas, como se ha señalado, están “despertando” a este interés.

En cuanto a **los resultados**, tal y como se refleja en las respuestas de los cuestionarios y entrevistas, la realización de las evaluaciones para las ONGD ha supuesto, más allá del cumplimiento de una normativa de AECID, una oportunidad de aprendizaje en la propia materia de evaluación. Muchos de los técnicos de las ONGD que han participado en las evaluaciones era la primera vez que se enfrentaban a la labor de comanditario de las mismas e incluso para algunas ONGD estas evaluaciones eran las primeras que se realizaban en su institución. Al añadido del aprendizaje se suma el despertar de una cultura evaluativa en muchas de las ONGD (se ha despertado la necesidad de formarse en evaluación, se ha contratado personal experto en la materia y/o se han creado departamentos de evaluación en las instituciones. En definitiva, se ha generado el interés por la evaluación y se ha tomado conciencia de la importancia de incorporar esta fase del ciclo del proyecto a las intervenciones de una manera más sistemática). Por último, en todos los casos se ha generado un plan de mejora de la intervención evaluada (no se ha analizado en este estudio la aplicación de dicho plan de mejora) que en muchos casos tiene un alcance mucho mayor que la exigencia que imponía la propia normativa.

Estos logros nos conducen directamente a los resultados intermedios de la cadena casual de objetivos, en donde los datos cuantitativos nos muestran que las evaluaciones (supuestamente por todas las razones expuestas en el nivel más bajo de la cadena causal de objetivos) han sido útiles tanto para las ONGD (porque ha supuesto un aprendizaje, se ha despertado una cultura evaluativa y se cuenta con un programa de mejora de las intervenciones e incluso de enseñanzas extrapolables a otras intervenciones) como para la AECID (porque la evaluación es claramente un medio de rendición de cuentas).

Con todo este análisis podemos dar respuesta a una de las preguntas que nos hacíamos en el diseño de esta evaluación: “Sí, las evaluaciones han sido útiles”, aunque podrían haberlo sido más si cambiaran algunos elementos estructurales que están claramente influyendo en los procesos de la evaluación. Sin embargo, el alcance de esta evaluación no permite responder a la pregunta de si la evaluación es una fase que se incorpora en “el ciclo del proyecto” para la mejora de la calidad de las intervenciones. Todo apunta a pensar que sí es así pues ya se cuenta con un plan de mejora concreto, pero en el momento en que se ejecutó la presente evaluación eran muy pocas las ONGD que ya estaban aplicando en la práctica dicho plan de mejora pues no había transcurrido el tiempo suficiente para ello.

Como foto final explicativa de los hallazgos, interpretaciones y juicios, volvemos a dibujar el programa mediante la teoría del programa. Allá donde se expresa “+” se entiende que el efecto es positivo (cuantas más “+”, mayor grado de efecto positivo), en donde se dibuja un “-“, se entiende que el efecto es mejorable.



VI.- RECOMENDACIONES

En base a los hallazgos, interpretaciones y enjuiciamientos hechos en los apartados anteriores, se recoge en este apartado algunas sugerencias de posibles líneas de acción a seguir, explicitando el agente al que va dirigido.

Recomendaciones para la AECID

1.- Aunque el borrador de la nueva normativa de AECID que regula las evaluaciones de Convenios (actualmente en fase de elaboración) parece apuntar a que las evaluaciones ya no serán obligatorias para todos los Convenios, ni serán encargadas por las ONGD, se recomienda que para aquellas que se realicen de manera obligatoria, de acuerdo al plan de evaluaciones que se elaborará anualmente, se constituya un comité mixto (AECID-ONGD) comanditario de la evaluación en donde ambas partes asuman funciones al mismo nivel y se convierta en un medio de comunicación ágil y eficaz. La existencia de un comité mixto permitiría una negociación más ágil de los intereses de ambos en la evaluación.

2.- Dado que la normativa está en discusión en estos momentos, también se recomienda la posibilidad de reconsiderar la obligatoriedad de hacer evaluaciones intermedias. En opinión de este equipo de evaluación, aportaría mayor utilidad a los resultados de la evaluación estudiar cada casuística para determinar cuál es el mejor momento. El criterio para determinar ese mejor momento podría estar en función de la duración total de la intervención prevista inicialmente. Hay convenios que están pensados como una primera fase de un proceso de más largo plazo. En estos casos sería recomendable una evaluación al final del primer convenio (en vez de la intermedia).

3.- Aunque los resultados muestran que el criterio de transparencia es importante para las ONGD en el proceso de elección de empresa evaluadora, este criterio podría tomar más fuerza si la AECID, a través de su normativa, regulara la obligatoriedad de sacar una oferta pública para todo proceso de selección de empresa evaluadora, o si se constituyese el mencionado comité de evaluación en donde la AECID compartiera con la ONGD un papel más protagonista en esta fase de la evaluación. Se debería ser cuidado para que la oferta fuera pública tanto en España como en el país en donde tiene lugar la ejecución del Convenio para promover el uso de los recursos locales, no obstante, con el único fin de lograr mayor utilidad a los procesos de evaluación, se recomienda que se elija al mejor equipo evaluador en función de los criterios

establecidos previamente por el comité evaluador, independientemente de la nacionalidad de los mismos.

4.- Con el único fin de lograr mayor utilidad de los procesos de evaluación, se recomienda a AECID abrir la posibilidad, a través de la normativa, de la realización de evaluaciones mixtas con mayor participación tanto de las ONGD como de los SL.

5.- Se recomienda que la AECID ponga en marcha los mecanismos necesarios para garantizar la publicación y difusión de las mismas con el fin de generar una mayor transparencia a la ciudadanía, facilitar el aprendizaje de otros actores de modo que las evaluaciones de conviertan en bienes públicos globales (sin rivalidad ni apropiación privativa).

6.- Si fuera de interés para la AECID conocer el grado de utilidad generado por las evaluaciones a más largo plazo, sería conveniente realizar un estudio sobre el seguimiento de la aplicación del plan de mejora obtenido y/o sobre los efectos a más largo plazo que las evaluaciones hayan podido tener al interior de las organizaciones.

Recomendaciones para las ONGD

1.- Esta investigación, se ha basado en las percepciones de las ONGD y de la AECID, pero no ha tenido en cuenta las percepciones de los Socios Locales. Dado que éstos son los últimos responsables de la aplicación del plan de recomendaciones, sería interesante conocer en profundidad el grado de utilidad que las evaluaciones les ha reportado a ellos, lo que tendría que ser objeto de una nueva investigación. No obstante, se recomienda a las ONGD liderar un proceso de desarrollo de cultura evaluativa entre sus SL generando espacios de reflexión y formación para asegurar la mayor participación e implicación de estos en los procesos de evaluación.

2.- Potenciar una reflexión-discusión al interior de sus organizaciones y con sus SL sobre las necesidades informativas propias de cada proceso de evaluación, y plasmar éstas en las preguntas de evaluación reflejadas en los TdR. Iniciar este proceso reflexivo con tiempo suficiente antes de la elaboración de los TdR ya que una buena explicitación de las necesidades informativas es el primer condicionante para obtener buenas evaluaciones.

3.-Dado que los Convenios se ejecutan con dinero público, los resultados de las evaluaciones (al menos los principales hallazgos, como una cuestión de mínimos) deberían ser publicados en aras de transparencia a la ciudadanía. Se recomienda a las ONGD generar al interior de sus organizaciones un proceso de reflexión sobre esta cuestión que conduzca a una mayor publicación y difusión de sus evaluaciones.

Recomendaciones para el comité/organismos responsables del encargo de la evaluación:

- 1.- Prioridad en la contratación de equipos de evaluación, un mínimo de dos evaluadores, donde se combinen la experticia (conocimientos y experiencia demostrable) en evaluación con la experticia en el sector y/o país en donde se ejecuta el Convenio. Si se contrata a una empresa evaluadora, se recomienda cerrar en los mismos TdR los CV de las personas que compondrán el equipo de evaluación.
- 2.- Primar en la selección de ofertas la explicación clara de la metodología a utilizar y su justificación así como de las principales herramientas de recogida de información (técnicas cualitativas, cuantitativas o mixtas).
- 3.- Acotar las dimensiones, preguntas y ámbito sectorial y geográfico de las evaluaciones a realizar, y dotar a éstas del presupuesto adecuado en función de esta acotación así como de la composición del equipo evaluador adecuado para realizar la evaluación tal y como se diseña.
- 4.- Explicitar en los TdR y en las ofertas tiempos empleados (estimados) por el equipo de evaluador para cada una de las fases de la evaluación.

Recomendaciones de utilidad para la Comunidad de Madrid.

La Comunidad de Madrid también tiene experiencia en el seguimiento y ciclo de evaluaciones externas gestionadas por ONGD. Desde la convocatoria de Programas en 2002, ha sido necesaria la presentación de una evaluación externa final. Este proceso ha durado hasta 2009, última convocatoria de programas hasta la fecha. La Comunidad de Madrid, en entrevista mantenida con su actual subdirector, también advierte de la gran diversidad en la calidad de las evaluaciones que le presentan las ONGD y en su uso final. Tiene la percepción de que una minoría de Organizaciones, probablemente las más grandes en recursos y sobre todo las que tienen personal más formado en evaluaciones, tienden a presentar informes de mejor calidad. No obstante, no se advierte aún una sistematización de los aprendizajes generados por las evaluaciones, aunque existen casos puntuales en que se han presentado modificaciones a intervenciones presentadas o nuevas fases, que han surgido de las recomendaciones recibidas por una evaluación anterior. Los posibles factores que explican esta buena práctica son, entre otros: (i) la cultura de evaluación que tenga la ONGD, entendiendo por ella las ganas de aprender y mejorar, yendo más allá de la simple obligatoriedad de evaluar; (ii) la formación específica que tenga el personal promotor de la evaluación dentro de la

ONGD. En esta línea, se están diseñando actividades de formación conjunta entre la CM y la FONGDCAM relativa a gestión por resultados y evaluación, que pueden tener una positiva influencia para el futuro.

En línea con lo argumentado en este trabajo, también se detecta que cuando los informes superan la mera obligatoriedad de evaluar y la ONGD se involucra sinceramente en el aprendizaje, los informes obtenidos son de mayor calidad y utilidad.

Asimismo, las Organizaciones que disponen de alguna Unidad de Evaluación o que el responsable de la evaluación tiene una competencia específica, más allá de ser la persona que hace el seguimiento de una intervención, la utilidad también se acrecienta.

Por último, se adivina una escasa difusión interna de las evaluaciones dentro de las ONGD en modo de talleres o seminarios que reúnan a personal directivo, cooperantes, staff del departamento de comunicación o de estudios, campañas e incidencia.

Por su parte, la propia CM reconoce que debe fortalecer sus capacidades internas de seguimiento, difusión y retroalimentación de las evaluaciones recibidas. Próximamente, espera poder contar con dos personas formadas y destinadas de forma prioritaria a esta misión, tras el largo periodo de ajuste de personal y procedimientos administrativos atrasados que se venía arrastrando desde 2007. Las recomendaciones contenidas en esta investigación le podrán ser muy útiles como punto de partida para el fortalecimiento propio y de forma conjunta (cooperativa) con las ONGD con las que se asocia.

VII.- PLAN DE COMUNICACIÓN

El Plan de comunicación tiene como fin poner en común los hallazgos y recomendaciones con los distintos stakeholders. Para ello se han planificado una serie de actividades que permitirán que la evaluación sea pública y promoverá su utilidad. El plan de difusión ha sido el siguiente:

- 1.- Reunión con CECOD, en donde se revisa el borrador de informe y se hacen comentarios y sugerencias.
- 2.- Elaboración de algunos productos derivados para la difusión. Presentación PPT con los principales contenidos
- 3.- Difusión entre las ONGD, miembros del CECOD y autoridades públicas (DGPOLDE, CAM, Ayuntamiento de Madrid...), a través de un taller en donde se exponen los resultados de la evaluación y se abre un debate sobre las mejores prácticas identificadas en temas de evaluación entre las ONGD, celebrado en diciembre de 2011 en la sede del CECOD.
- 4.- Presentación de la metodología y resultados de la evaluación en la VII Conferencia bienal de la Sociedad Española de evaluación. Celebrada en Sevilla en el mes de octubre de 2011.
- 5.- Publicación de dicho informe en un documento de trabajo de la SERIE CECOD.

VIII.- LÍMITES Y CONDICIONANTES DE LA EVALUACIÓN

De manera sintética se mencionan los condicionantes que han afectado al desarrollo del proceso de esta investigación-evaluación.

1.- De las 37 ONGD que habían obtenido subvención mediante Convenios con AECID, y que fueron contactadas para ser encuestadas o entrevistadas, hubo 2 que habían aplazado la fecha de la evaluación y por tanto todavía no tenían resultados (ni los iban a tener hasta el mes de junio, mes en que ya estaría analizada la información recogida), hubo 3 que no quisieron participar y así lo expusieron, y hubo 4 que aunque no lo dijeron no participaron a pesar de los numerosos contactos (en algunos casos hubo hasta 5 intentos contactos vía mail y telefónico). Por lo tanto, aunque hubiera sido deseable haber contado con la respuesta de toda la muestra, se obtuvo un nivel de representación en las respuestas del 75,6%, es decir, 28 ONGD.

2.- Otro limitante fue sobre el diseño del cuestionario enviado a toda la muestra. Después de largas discusiones entre los miembros del equipo evaluador se optó por enviar un cuestionario por ONGD y no uno para cada intervención realizada. En el caso de las ONGD que tenían varias intervenciones (convenios), se trató de que contestaran con opiniones medias que representaran al conjunto de sus intervenciones. Sin embargo en algunas ocasiones esto no fue posible dado que su opinión era extremadamente diferente de una intervención a otra. Finalmente, fue el equipo evaluador el que tuvo que optar por dar una valoración media en cada una de las preguntas del cuestionario por cada ONGD para poder homogeneizar la información de manera que se pudiera analizar estadísticamente.

3.- Aunque hubiese sido deseable realizar las entrevistas en profundidad una vez recogida y analizada la información a través del cuestionario, por razón de tiempo, se tuvo que compaginar la recogida de información a través de las dos herramientas – cuestionario y entrevista-, porque hubo ONGD que se demoraron mucho en dar respuesta al cuestionario.

4.- Por falta de recursos económicos, la información recogida de manera cualitativa (entrevistas y grupo focal) no ha sido transcrita.

5.- Igualmente, por limitación de recursos, el estudio se ha centrado en recoger la percepción de las ONGD y de la propia AECID, la DGPOLDE y la Comunidad de Madrid pero no se ha podido abordar la percepción de la utilidad de las evaluaciones realizadas desde el punto de vista de los socios locales de las ONGD españolas.

6.- Por último, señalar que esta evaluación, debido a una limitación del tiempo transcurrido desde que se habían hecho las evaluaciones (algunas de ellas todavía en fase de finalización), ha considerado su alcance hasta el momento de elaboración de un plan de mejora por parte de las ONGD, sin haber sido posible profundizar en la aplicación de dicho plan de mejora o en las consecuencias a más largo plazo que las evaluaciones hayan podido tener sobre las propias ONGD (mejora en la formulación de nuevas intervenciones, generación de cultura evaluativa al interior de las organizaciones, entre otras).

BIBLIOGRAFÍA

AECID, (2005). Orden AEC/1303/2005, de 27 de abril, por la que se regulan las bases para la concesión de subvenciones a organizaciones no gubernamentales de desarrollo, para la realización de intervenciones en materia de cooperación internacional para el desarrollo. www.aecid.es

AECID (2009). Guía para evaluaciones de proyectos y convenios de ONGD. www.aecid.es

AECID (2009). Normas de seguimiento y justificación de proyectos y convenios de cooperación al desarrollo subvencionados a las organizaciones no gubernamentales de desarrollo. www.aecid.es

Ligero, J.A (2011) *Dos métodos de Evaluación: Criterios y Teoría del Programa*. Documento de trabajo SERIE CECOD. Número 15/2011

OCDE (2010). *Estándares de calidad para la evaluación del Desarrollo del CAD*. Serie: Directrices y Referencias del CAD. www.oecd.org/editions

Rodríguez, Liliana (2008) *Evaluaciones colaborativas. Un modelo paso a paso para el evaluador*. Llumina Press.

Stake, Robert E. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Crítica y Fundamentos. GRAO.

ANEXOS

Anexo 1. MATRIZ DE EVALUACIÓN

Número pregunta	Dimensión	Criterios	Preguntas de evaluación	Técnicas
1	Resultados	Impacto	¿Las recomendaciones obtenidas en el proceso de evaluación son consideradas útiles para las ONGD?	Encuesta/Entrevista
2	Resultados	Impacto	¿Se aplican recomendaciones al ciclo de vida del Convenio evaluado, mejorándolo de esta manera?	Encuesta/Entrevista
3	Resultados	Impacto	¿Se obtienen recomendaciones útiles para otros proyectos/intervenciones de la ONGD?	Encuesta/Entrevista
4	Resultados	Impacto	El proceso de evaluación, ¿ha resultado útil para la ONGD para otros fines, más allá de la mejora del propio convenio evaluado?	Encuesta/Entrevista
5	Resultados	Impacto	Según el punto de vista de las ONGD, ¿para quién resulta más útil el proceso de evaluación (AECID, beneficiarios, ONGD, otros actores)?	Encuesta/Entrevista
6	Resultados	Impacto	¿La política de gestión de las evaluaciones influye en la mejora de la cultura de evaluación (más capacidades, más interés...) en las ONGD?	discusión
7	Estructura/Resultados	Pertinencia	La obligatoriedad de realizar la evaluación viene determinada por la normativa de AECID. En opinión de las ONGD, ¿la norma favorece que se realice la evaluación?	Encuesta/Entrevista
8	Estructura/Resultados	Pertinencia	La normativa establece que la evaluación debe ser externa. En opinión de las ONGD, ¿la externalización reporta mayor utilidad al proceso de evaluación?	discusión
9	Estructura/Resultados	Pertinencia	La normativa AECID establece que la evaluación debe realizarse en el momento intermedio. ¿Este momento es considerado por la ONGD como el más adecuado para conseguir la mayor utilidad del proceso?	Encuesta/Entrevista
10	Estructura/Resultados	Pertinencia	¿La normativa AECID es lo suficientemente flexible para adaptarse a cada caso de evaluación?	Encuesta/Entrevista
11	Estructura/Resultados	Eficiencia	¿El tiempo utilizado en todo el proceso de evaluación (desde el diseño de la misma hasta la aplicación de recomendaciones) determina la utilidad del propio proceso de evaluación?	Encuesta/Entrevista
12	Estructura/Resultados	Eficiencia	¿Los recursos invertidos: tiempo de las ONGD y sus SL, dinero (financiación AECID) y recursos humanos, son proporcionales en cuanto a la utilidad que proporciona la realización de las evaluaciones?	Encuesta/Entrevista
13	Estructura/Resultados	Eficiencia	¿La metodología (evaluación según los criterios del CAD + procesos) determina el grado de utilidad de las evaluaciones?	Encuesta/Entrevista
14	Estructura/Resultados	Transparencia	¿La ONGD establece el proceso de selección de la empresa evaluadora de acuerdo a criterios transparentes?	Encuesta/Entrevista
15	Estructura/Resultados	Eficacia	¿La experiencia y capacidades de los evaluadores determina el grado de utilidad del proceso de evaluación?	Encuesta/Entrevista
16	Estructura/Resultados	Eficacia	¿La incorporación en el equipo evaluador de expertos sectoriales, de acuerdo a la temática del convenio, agrega utilidad a los resultados de la evaluación?	Grupo de Discusión
17	Estructura/Resultados	Eficacia	¿La incorporación en el equipo evaluador de personal local, de acuerdo al país o países donde se ejecuta el convenio, agrega utilidad a los resultados de la evaluación?	Grupo de Discusión
18	Estructura/Resultados	Eficacia/Eficiencia	¿Existen cuellos de botella en los recursos estructurales (normativa, metodología, recursos, experiencia, etc., que determine el grado de utilidad de las evaluaciones?	Grupo de Discusión
19	Procesos/Resultados	Participación	¿Resulta de utilidad para las ONGD las aportaciones realizadas por la AECID en el momento de elaboración de los TDR?	Encuesta/Entrevista
20	Procesos/Resultados	Participación	¿Resulta de utilidad para las ONGD la participación de la AECID en la selección de la empresa evaluadora?	Encuesta/Entrevista
21	Procesos/Resultados	Participación	¿Resulta de utilidad para las ONGD la participación de la AECID en la discusión de las recomendaciones a aplicar en la marcha del convenio?	Encuesta/Entrevista
22	Procesos/Resultados	Eficacia/Eficiencia	¿Existen cuellos de botella en la relación entre la ONGD y AECID que condicionen el grado de utilidad de las evaluaciones?	Grupo de Discusión
23	Procesos/Resultados	Participación	¿La participación de la ONGD en las distintas fases del proceso de evaluación, determina el grado de utilidad de la evaluación para la misma?	Encuesta/Entrevista
24	Procesos/Resultados	Participación	¿La participación de los socios locales en las distintas fases del proceso de evaluación, determina el grado de utilidad de la evaluación?	Encuesta/Entrevista
25	Procesos/Resultados	Eficacia/Eficiencia	¿Existen cuellos de botella para la participación de la ONGD y/o Socios locales en el proceso de evaluación que afecte a la utilidad de la misma?	Grupo de Discusión
26	Procesos/Resultados	Eficacia/Eficiencia	¿Existen cuellos de botella en la profesionalidad/capacidades/experiencia de los equipos evaluadores, en su hacer en las distintas fases del proceso de evaluación, que determinen el grado de utilidad de la propia evaluación?	Grupo de Discusión
27	Procesos/Resultados	Transparencia	¿La ONGD considera útil la publicación/difusión de los resultados de la evaluación?	Encuesta/Entrevista

Anexo 2. HERRAMIENTAS DE RECOGIDA DE INFORMACIÓN

Describa de forma breve cómo está organizada la evaluación en su ONGD: (por ejemplo, quién se encarga de hacer los TdR, si existe una comisión de evaluación, qué cuellos de botella o limitaciones encuentran a la hora de plantear las evaluaciones):

(Preguntas relativas a la Participación).

A continuación le voy a hacer algunas preguntas a las que debe responder SÍ/NO

Toda las preguntas se refieren a la ONGD y al SL (se pregunta de forma diferenciada para los dos actores)	ONGD	ONGD	SL	SL
2.1.- ¿La ONGD /SL participó en la elaboración de los TdR?	SI	NO	SI	NO
2.2.- ¿La ONGD/SL participó en la elaboración de las preguntas de la evaluación, es decir, en lo que le interesaba conocer?	SI	NO	SI	NO
2.3.- Respecto a los 5 criterios marcados por AECID-CAD, ¿Hubo cambios por parte de la ONGD/SL en la incorporación, exclusión, de alguno de los criterios?	SI	NO	SI	NO
2.4.- ¿La ONGD/SL participaron en el trabajo de campo para la recogida de la información?	SI	NO	SI	NO
2.5.-¿ La ONGD/SL participó en la elaboración de las conclusiones?	SI	NO	SI	NO
2.6.- ¿ La ONGD/SL participó en la elaboración de las recomendaciones?	SI	NO	SI	NO
2.7.- ¿Participó la ONGD/SL en algún taller realizado en el terreno para la discusión de conclusiones y recomendaciones?				
2.8.- ¿El evaluador fue pidiendo la participación de la ONGD/SL en diferentes etapas del proceso evaluativo?	SI	NO	SI	NO
2.9.- ¿La ONGD/SL participó en las etapas anteriores a la recepción del informe final de la evaluación?	SI	NO	SI	NO

Según su experiencia, ¿Cuáles serían los aspectos clave y los cuellos de botella de la fase de participación?

Preguntas relativas a la opinión/evaluación de las ONGD sobre el trabajo realizado por el evaluador externo)

A continuación le voy a hacer 10 preguntas a las que debe contestar en una escala de: Malo, regular, bueno y muy bueno. Valore el trabajo del equipo evaluador con respecto a:

		Malo	Regular	Bueno	Muy Bueno
3.1	Diseño de las preguntas de la evaluación por cada criterio analizado (se valoran aspectos como haber tenido en cuenta los intereses de la ONGD y del SL, las preguntas eran suficientes, era una información interesante de conocer)				
3.2	Recogida de información (Se valoran aspectos como: tiempo empleado, información recogida suficiente, información fiable, adecuados métodos de recogida, etc.)				
3.3	Objetividad y credibilidad de la interpretación de la información recogida				
3.4	Calidad del sistema de comunicación utilizado con la ONGD (Se valoran aspectos como que la información haya sido fluida, se hayan dado espacios a la participación, haya sido clara, etc.)				
3.5	Utilidad y pertinencia de las recomendaciones				
3.6	¿En qué grado considera que la evaluación ha sido completa (se han identificado fortalezas y debilidades)?				
3.7	¿En qué grado considera que se han tenido en cuenta los intereses de la ONGD y del SL?				
3.8	¿Cómo calificaría al equipo evaluador?				
3.9	¿Cómo calificaría el ajuste entre el trabajo de evaluación finalmente hecho y la propuesta que el equipo evaluador realizó en la oferta inicial?				

3.10.- Valore la evaluación del 1 al 10.

1	2	3	4	5	6	7	8	9	10

Según su experiencia, ¿Cuáles serían los aspectos clave y los cuellos de botella del trabajo de los evaluadores?

(Preguntas relativas a la utilidad y credibilidad de la evaluación)

		Nada	Poco	Algo	Mucho
4.1	¿Le ha resultado útil todo el proceso de evaluación?				
4.2	¿Le han resultado creíbles las conclusiones y juicios?				
4.3	¿Le ha resultado útil las recomendaciones?				

Si ha contestado nada/poco en alguna de las tres preguntas, preguntar las causas

4.1.-

4.2.-

4.3.-

Conteste: SÍ o NO (Con las respuestas tratamos de contrastar el grado de utilidad expresado por la ONGD con la puesta en práctica de la misma)

4.4.- Se ha elaborado un plan de mejora por escrito	SI	NO
4.5.- Tras la evaluación se solicitaron algunos cambios a AECID	SI	NO
4.6.- ¿Recuerda alguna(s) recomendación(es) que se haya aplicado a lo que restaba del convenio a una fase posterior del mismo?	SI	NO
4.7.- ¿Se han extraído conclusiones útiles para ser aplicados en programas similares de la organización?	SI	NO

Si ha contestado NO en alguna de las preguntas 4.6 ó 4.7, preguntar las causas

4.6.-

4.7.-

4.8.- Conteste lo primero que le viene a la cabeza: ¿Cuál fue el principal aprendizaje de la realización de ésta evaluación?

.....
.....

4.9. ¿Para quién cree que ha sido más útil?

- AECID
- ONGD
- OL
- Beneficiarios
- Otros:.....

4.10.- ¿Por qué razón/es?

.....

4.10.- ¿Considera que ha habido alguna limitación/dificultad importante que ha condicionado el alcance de los resultados de la evaluación?.

- Sí. ¿Cuál?.....

- No

A continuación valore en una escala de cuatro grados: Insuficiente, suficiente, proporcional, y excesivo, las siguientes cuestiones:

Considera que los recursos invertidos por la ONGD en cuanto a: (se lee una por una): tiempo, recursos humanos y dinero han sido:

		Insuficiente	Suficiente	Proporcional	Excesivo
4.11	Tiempo				
4.12	Recursos humanos				
4.13	Dinero				

Según su experiencia, ¿Cuáles serían los aspectos clave y los cuellos de botella para que las evaluaciones sean útiles?

(Preguntas relativas a la difusión)

Conteste: SÍ o NO

5.1.- ¿La ONGD ha difundido externamente esta evaluación realizada en algún medio cómo en su web u otro?	SI	NO
--	----	----

Solo si contesta a la pregunta 5.1 con NO, ¿Por qué?.....

Muchísimas gracias por su colaboración

PUNTOS TRATADOS EN EL GRUPO DE DISCUSIÓN

1.- Obligatoriedad del momento intermedio:

¿Existe flexibilidad de la normativa cuando se explica la casuística a la AECID?. Si no, ¿creen que esto ha condicionado la menor utilidad de las evaluaciones?.

2.- Obligatoriedad de evaluación externa:

¿Perciben flexibilidad en la normativa (hubo casos en donde la ONGD acompañó incluso en el trabajo de campo)?, ¿creen que la externalización resta utilidad a la evaluación?.

3.- Participación de AECID en la elaboración de los TdR

¿Cómo se percibe ésta? (¿intromisión?, ¿tedioso?)

4.- ¿Difícil encontrar buenos expertos en evaluación en España, y sobre las empresas que subcontratan?.

5.- ¿Qué es preferible: experto en evaluación o en el tema/sector?. (Compartir experiencias)

6.- Opinión sobre la “obligatoriedad (no recogida en normativa) de incorporar expertos locales. Pros y contras del trabajo realizado por locales.

7.- ¿Objetivo de la evaluación demasiado amplio?. ¿Procesos vs 5 criterios del CAD?. ¿El tener objetivos tan amplios resta utilidad a la evaluación?.

8.- Cultura y conocimiento de evaluación. ¿Se sienten con capacidad para discutir, negociar el trabajo realizado por los evaluadores externos?.

9. ¿Por qué no se publica?

10. los principales cuellos de botella que encuentran para hacer de las evaluaciones un instrumento útil y de cambio en sus Organizaciones

GUION ENTREVISTA CON AECID

Respecto al momento:

Se debería dar más margen a la organización para realizarlas en el momento que puedan ser de más utilidad para la propia intervención y para la ONGD, en función de la duración prevista inicialmente de la intervención (hay convenios que están pensados como una primera fase de un proceso más largo y en estos casos hubiera sido de mayor utilidad una evaluación final (del primer convenio) en vez de la intermedia.

1.- ¿Alguna ONGD lo planteó, habría flexibilidad en este aspecto por parte de la normativa?

Motivo principal:

Claramente la mayoría de las ONGD ha visto en las evaluaciones, una posibilidad de aprendizaje y de mejora, más allá de la obligatoriedad de la norma. La mayoría de las ONGD han vivido los intereses de la AECID como “imposición” lo que posiblemente haya restado utilidad para las ONGD. La AECID ha tratado de que sus intereses se vieran reflejados en las evaluaciones, especialmente a través de su participación en la definición de los TdR. En este sentido, la apreciación de las ONGD es que la participación de la AECID en esta fase ha generado procesos demasiado largos y que han generado poco valor añadido, produciendo un efecto “desgaste” en la mayoría de las ONGD.

2.- ¿Cuál es la opinión de la AECID en este sentido?.

Externalidad y elección de empresa evaluadora:

En relación con la anterior cuestión, para las ONGD la función principal de la evaluación era el aprendizaje y no la rendición de cuentas. En este sentido, se considera que una evaluación mixta hubiera permitido un mayor aprendizaje y nivel de apropiación. Además, los equipos mixtos hubieran permitido generar sinergias ya que la ONGD y SL pueden aportar el conocimiento del país, del contexto, y del sector de la intervención, que muchas veces los evaluadores no tienen. Mientras que los evaluadores aportan el conocimiento en evaluación (metodología y herramientas) que la ONGD y SL no suelen tener.

3.- ¿Cabría flexibilidad de la normativa para este aspecto, no resultaría mucho más interesante para el aprendizaje y la generación de cultura evaluativa?.

Exigencia de evaluadores locales

Claramente en convenios multi-país, este requisito ha sido contraproducente ya que se ha perdido la mirada regional del convenio y no se garantiza que una persona de origen de uno de los países del convenio aporte un conocimiento en profundidad del resto de los países del convenio.

En el resto de convenios, la mayor dificultad encontrada ha sido la falta de conocimiento del evaluador local del sistema de la cooperación española, aunque esto no ha sido siempre así.

4.- ¿Cuál es la percepción de AECID?, ¿evaluadores locales han reportado mayor calidad a los proceso de evaluación?.

Participación de los Socios Locales

La participación de los socios locales sigue siendo baja. En opinión de las ONGD, esta menor participación se debe a la falta de interés porque no se concibe como una oportunidad de mejora sino como un trámite más, una imposición, debido a la falta de cultura evaluativa en sus organizaciones.

5.- ¿Cuál es la percepción de la AECID, no podría regularse por normativa una mayor participación?

Sobre la calidad del trabajo hecho por los evaluadores:

El trabajo realizado por los evaluadores, en términos generales, es calificado de bueno por las ONGD (aunque la cuarta parte de las ONGD califican este trabajo como regular). En general, la valoración global sobre el trabajo realizado se calificaría de notable en una escala del 1-10.

6.- ¿Cómo lo calificaría de media la AECI?

La menor satisfacción de las ONGD sobre el trabajo de los evaluadores se muestra en la fase de recogida de información y en la que la evaluación haya sido completa (hay una tendencia a recoger lo negativo y no identificar las fortalezas de las intervenciones), mientras que la mayor satisfacción se muestra en la comunicación mantenida entre los equipos evaluadores y las ONGD.

7.- ¿Cuáles podrían ser los puntos fuertes y débiles, en opinión de la AECID, sobre la calidad del trabajo hecho por los evaluadores?.

Sobre la utilidad

La percepción revelada sobre la utilidad del proceso, las conclusiones y recomendaciones parece ser alta y muy alta en la mayoría de las ONGD. Esta alta valoración conduce a querer conocer *las consecuencias* concretas que ha tenido la evaluación sobre las intervenciones de las ONGD. La primera consecuencia es contar con un plan de mejora que la mayoría de las ONGD cuentan ya con él o están en proceso de elaboración. Muchas de las ONGD afirman que su plan de mejora es más completo y de mayor alcance que la adenda firmada con la AECID.

8.- ¿Las adendas firmadas, fruto de las evaluaciones, responden a las expectativas de la AECID, consideran que introducen cambios importantes para la mejora de las intervenciones?

En cuanto al agente para el que la evaluación ha sido más útil, las ONGD perciben que ellas mismas son el principal beneficiario de la evaluación, seguidos de sus socios locales, quedando en último término la propia AECID y los beneficiarios, aunque la mayoría reconoce que de manera indirecta, en la medida en que se mejoran los procesos, la evaluación también resulta útil para los beneficiarios últimos.

9.- Las ONGD perciben que la AECID no tiene capacidad para asimilar todas las evaluaciones realizadas y por tanto esto resta utilidad de las mismas para este agente. ¿Cómo lo ve la AECID y cuál era la finalidad principal cuando se puso en marcha esta normativa: utilidad para las ONGD, para la AEI, la rendición de cuentas, la mejora de los procesos, etc.?

Sobre los recursos utilizados.

No se deduce del análisis cuantitativo una falta de recursos (económicos, humanos y de tiempo) que correlacione con la utilidad-credibilidad de la evaluación. Sin embargo, desde un punto de vista cualitativo hay reflexiones interesantes sobre el condicionante de los recursos para obtener trabajos de calidad: “El dinero para hacer la evaluación, determina la composición del equipo evaluador, y el tiempo que se emplea para recoger información en terreno, son dos cuestiones básicas para hacer un trabajo de calidad”. En este sentido, afirmar que sí existen algunas quejas generalizadas por parte de las ONGD de la escasez de tiempo dedicado al trabajo de campo por parte de los evaluadores lo que limita el alcance de la recogida de información, así como deficiencias importantes en la composición de los equipos evaluadores.

10.- ¿Cuál es la opinión de la AECID al respecto?

Con respecto al tiempo también existe la percepción generalizada de que es escaso no para realizar la evaluación pero sí para procesar el aprendizaje obtenido y para implementar mejoras al quedar poco tiempo de vida del convenio una vez obtenido el informe de evaluación.

11.- ¿Qué opina la AECID sobre este aspecto y qué posibles soluciones vería para ello?

Limitaciones, cuellos de botella:

Los principales cuellos de botella recogidos por las ONGD para restar utilidad a las evaluaciones realizadas son los siguientes (se recogen solo aquellos que más se repetían en la información cualitativa):

Metodología inadecuada para convenios regionales, se pierde la visión de Región y dificulta tener que evaluar en todos los países con la misma contratación de equipo de evaluación.

12.- ¿Qué opinión tiene la AECID una vez tenida esta primera experiencia?, ¿cabría una mayor flexibilidad en la normativa para este tipo de convenio?

Limitaciones en el propio equipo evaluador: Falta de dominio del idioma local, falta de coordinación entre los miembros del equipo, escasa formación en metodología y herramientas de recogida de información de evaluación, falta de evaluadores expertos en el sector, materia que se evalúa.

13.- ¿Qué opinión tiene la AECID sobre las limitaciones de los equipos de evaluación, en su opinión, cuáles serían las debilidades de los mismos?.

14.- ¿Correspondencia entre el producto final entregado y lo firmado en los TdR-oferta?.

15.- ¿Opinión en el dilema entre experto en evaluación y experto en el sector?

Objeto de la evaluación demasiado amplio: El hecho de que se obligue a evaluar bajo los criterios del CAD no permite dar tanta importancia a los procesos que en una evaluación intermedia debería ser el elemento central. Este punto parece ser que ha sido uno de los que la AECID ha mostrado mayor inflexibilidad.

16.- ¿Cuál es la finalidad de la AECID para obligar a evaluar según los criterios CAD en una evaluación intermedia cuya finalidad parece ser el aprendizaje y la mejora continua y por tanto parecería más adecuado centrarse en los procesos?

17.- ¿La AECID identifica algún otro cuello de botella en los procesos de evaluación ejecutados?

Difusión.

Los datos muestran que sólo 6 ONGD (21%) han respondido que han difundido de alguna forma su evaluación. El principal motivo de la falta de publicación parece ser la falta de cultura evaluativa (a mayor cultura evaluativa, mayor difusión de los resultados de las evaluaciones o de las propias evaluaciones), de hecho muchas ONGD contestaron que éste era un tema que aún no se había debatido, reflexionado al interior de su organización. Aunque se recogen otras opiniones como la falta de calidad de los informes de evaluación y el temor de la publicación de malos resultados, pero estas opiniones se recogen en mucho menor número de casos.

18.-¿Qué opinión tiene la AECID sobre este tema?, ¿se plantea la difusión de las evaluaciones en su propia página web, por qué?

Principales aprendizajes para las ONGD.

Generar cultura evaluativa

Conocimiento profundo de las intervenciones: hacerlas más realistas y con mejores sistemas de información.

Aprender de los propios procesos de evaluación para el futuro (acotar los Cvs en las ofertas, acotar las técnicas de recogida de información, etc.)

19.- ¿Y para la AECID cuál ha sido su aprendizaje del proceso y cuál ha sido en su opinión el aprendizaje para las ONGD?

20.- Utilidad para AECID: ¿Cuál era la principal finalidad de la AECID al poner en marcha esta normativa?, ¿cuál es su balance ahora tras haberse aplicado, ha respondido a sus expectativas?

Anexo 3. ESTADÍSTICOS DESCRIPTIVOS

	N	Mínimo	Máximo	Media	Desv. típ.
MOMENTO	28	1,00	3,00	1,8571	,44840
MOTIVO	23	1,00	4,00	2,6913	,96996
3OFERTAS	22	1,00	3,00	1,9500	,91482
AECID_acepta	28	,00	1,00	,9643	,18898
PO_TdR	28	,00	1,00	,9643	,18898
PSL_TdR	26	,00	1,00	,5385	,50839
PO_preg	28	,00	1,00	,9643	,18898
PSL_preg	26	,00	1,00	,5769	,50383
PO_crit	27	,00	1,00	,4444	,50637
PSL_crit	25	,00	1,00	,0800	,27689
PO_campo	28	,00	1,00	,5714	,50395
PSL_campo	26	,00	1,00	,5769	,50383
PO_concl	28	,00	1,00	,6429	,48795
PSL_concl	26	,00	1,00	,5385	,50839
PO_recom	28	,00	1,00	,5357	,50787
PSL_recom	26	,00	1,00	,3077	,47068
PO_taller	28	,00	1,00	,7857	,41786
PS_taller	26	,00	1,00	,6923	,47068
Ev_pregO	28	,00	1,00	,9286	,26227
EV_pregSL	26	,00	1,00	,7692	,42967
PO_informe	28	,00	1,00	,9643	,18898
PSL_informe	26	,00	1,00	,7308	,45234
Diseño	27	1,00	4,00	2,8519	,60152
Informac	28	1,00	4,00	2,5714	,61935
ObjCredib	28	2,00	4,00	2,8036	,61372
Comunic	28	2,00	4,00	3,0536	,67136
Recomends	28	2,00	4,00	2,8036	,55007
completa	28	2,00	3,00	2,5714	,48523
intereses	28	2,00	4,00	2,8036	,61372
equipo	28	1,00	4,00	2,7857	,61507
ajuste	28	1,00	4,00	2,7857	,64447
Valor	28	3,00	9,00	6,5893	1,36798

Util_proceso	28	2,00	4,00	3,4464	,56665
Creib_concl	28	2,00	4,00	3,4107	,62440
Util_recom	28	2,00	4,00	3,3393	,60940
Pmejora	28	,00	1,00	,8214	,39002
cambiosAECID	28	,00	1,00	,4643	,50787
recuerda	28	,00	1,00	,8929	,31497
Conclutil	28	,00	1,00	,9286	,26227
Másútil	24	1,00	4,00	2,0167	,68504
Limitdific	28	,00	1,00	,7143	,46004
tiempo	28	1,00	4,00	2,4643	,92224
RRHH	28	2,00	4,00	2,6071	,73733
dinero	28	1,00	4,00	2,5536	,97505
difundida	28	,00	1,00	,2143	,41786
N válido (según lista)	15				

Anexo 4. PRESENTACIÓN PPT



LA INTEGRACIÓN DE LA EVALUACIÓN EN EL CICLO DEL PROYECTO EN LAS ONGD ESPAÑOLAS

José María Larrú y María Méndez
Universidad CEU San Pablo y Evaluadora
independiente

ESTRUCTURA DE LA PRESENTACIÓN

- 1.ACOTACIÓN DEL OBJETO DE EVALUACIÓN**
- 2.PREGUNTA PRINCIPAL DE LA EVALUACIÓN**
- 3.METODOLOGÍA UTILIZADA**
- 4.TÉCNICAS DE RECOGIDA DE INFORMACIÓN**
- 5.PRINCIPALES CONCLUSIONES**



1. ACOTACIÓN DEL OBJETO DE EVALUACIÓN

- Marco de los Convenios con AECID (intermedia externa obligatoria)
- Convocatorias de 2006 y 2007 con convenio

CECOD



2. PREGUNTAS DE EVALUACIÓN

- ¿Las evaluaciones han sido útiles y por qué?
 - Fortalezas y cuellos de botella del proceso
- ¿Se integra la evaluación en el ciclo del proyecto?
 - Fortalezas y cuellos de botella en la integración de los aprendizajes

Matriz de evaluación con 27 subpreguntas

CECOD



3. METODOLOGÍA UTILIZADA

- Modelo no experimental: Modelo a estimar

$$UTIL = \alpha + \beta ORG_PROM_i + \gamma PART_i + \delta EQUIPO_i + e DIFU + u_i$$

- Variables tenidas en cuenta:

- las condiciones determinadas por la propia **normatividad** de AECID para su desarrollo (obligatoriamente externa e intermedia)
- la **participación** de las ONGD y/o de sus socios locales;
- la calidad profesional de los **evaluadores** externos;
- y los **recursos** (tiempo, dinero y personal) empleado para la ejecución y difusión

CECOD



4. LA RECOGIDA DE INFORMACIÓN

- Modelo mixto de combinación de técnicas cuantitativas y cualitativas

- **CUESTIONARIO** CUANTI-CUALITATIVO

- UNIVERSO=37
- MUESTRA CONSEGUIDA=28 (84,8%)
 - 2 NO TIENEN AUN INFORME DE EVALUACIÓN.

- **ENTREVISTAS** EN PROFUNDIDAD SEMI-ESTRUCTURADAS

- **GRUPO DE DISCUSIÓN** (7 ONGD)

- Inscripción voluntaria
- Triangulación

- **ENTREVISTA ACTORES INVOLUCRADOS:**

- AECID, DGPOLDE

LIMITE: NO SOCIOS LOCALES

CECOD



5. PRINCIPALES CONCLUSIONES (I)

- **RESPECTO AL MOMENTO EN QUE SE REALIZA LA EVALUACIÓN**
 - la obligatoriedad de la normativa ha servido para generar mayor cultura evaluativa
 - PERO intermedias ha podido restar utilidad
- **MOTIVO PRINCIPAL PARA REALIZAR LA EVALUACIÓN**
 - aprendizaje y de mejora más allá de la obligatoriedad de la norma
 - fase de elaboración de los TdR :
 - ONGD: participación de la AECID ha generado procesos demasiado largos, generando poco valor añadido
 - AECID: que sus intereses se vieran reflejados (financiador)
- **EXTERNALIZACIÓN DE LA EVALUACIÓN**
 - Selección de consultora combinando transparencia con experiencias previas
 - ONGD preferencia por la mixta (mayor auto-aprendizaje y apropiación)

CECOD



PRINCIPALES CONCLUSIONES (II)

- **EXIGENCIA DE EVALUADORES LOCALES**
 - Dificultad en convenios multi-país (pérdida de la mirada regional)
 - que los locales pueden tener valor añadido en cuanto a conocer mejor el contexto regional, resultados controvertidos (conoce el sistema de la Cooperación Española? => recomendaciones poco prácticas o contrarias a la normativa)
- **PARTICIPACIÓN DE LAS ONGD**
 - Clave para su utilidad (sobre todo en primeras fases ciclo evaluativo: preguntas y TdR)
 - tanto AECID como ONGD: mayor concreción de las preguntas e involucrar a los SL
- **BAJA PARTICIPACIÓN DE LOS SOCIOS LOCALES**
 - *“falta de interés porque la evaluación no se concibe como una oportunidad de mejora sino como un trámite más, una imposición, debido a la falta de cultura evaluativa”*

CECOD



PRINCIPALES CONCLUSIONES (III)

● CALIDAD DEL TRABAJO DE LOS EVALUADORES

- La más explicativa de utilidad con los datos estadísticos.
- **ONGD**: notable (promedio de 6,5). AECID: suficiente (5)
- **ONGD**: satisfacción en comunicación; menor en recogida de información (tiempo en terreno; técnicas empleadas (sólo cuali); tendencia a resaltar sólo lo negativo
- **AECID**: pobre justificación de la metodología elegida y las técnicas de recogida de información; casos de débiles matrices de evaluación; satisfacción en respuestas a los criterios de evaluación y las recomendaciones
- **PROBLEMAS**:
 - subcontratación de evaluadores (no el “senior” de la consultora)
 - ¿expertos en evaluación o expertos en el sector? En evaluación con técnicos en el equipo
 - debilidad formativa en metodologías

CECOD



PRINCIPALES CONCLUSIONES (IV)

● SOBRE LA UTILIDAD

- Satisfacción promedio en los tres fines: comprensión de las intervenciones, iluminación para la mejora y rendición de cuentas
- Jerarquía: ONGD y AECID -> socios locales -> beneficiarios
- Principales limitaciones
 - **obligatoriedades**: momento, amplitud (cinco criterios CAD, resultados y procesos ...)
 - las de **algunos equipos evaluadores**: no entrevistar en idioma local, debilidades metodológicas, poco tiempo en terreno
 - Estadísticamente, no aparecen tiempo+RR.HH.+dinero como determinantes, pero: *“el dinero para hacer la evaluación parece determinar la **composición del equipo evaluador**, y el tiempo que se emplea para **recoger información en terreno**. Estas son dos cuestiones básicas para hacer un trabajo de calidad”*.

CECOD



PRINCIPALES CONCLUSIONES (V)

● **DIFUSIÓN DE LOS RESULTADOS**

- Sólo 6 ONGD (21%) la han difundido de alguna forma
 - falta de cultura evaluativa
 - es un tema que aún no se había debatido y/o se está reflexionado ahora de forma interna (ligado a transparencia)

● **FACTORES DETERMINANTES DE LA UTILIDAD**

- CUANTITATIVO: orientar con precisión el trabajo del equipo evaluador
- CUALITATIVO:
 - la amplitud y profundidad de la participación de la propia ONGD en todo el proceso de evaluación.
 - no se refleja, la de los Socios Locales

CECOD



VISION SISTÉMICA

¿IMPOSICION- OPORTUNIDAD?



CECOD

