



ALGUNAS CUESTIONES CONCEPTUALES Y METODOLÓGICAS EN TORNO A LA EVALUACIÓN DE IMPACTO

José María Larrú
Universidad CEU San Pablo

INTRODUCCIÓN

La evaluación de impacto está adquiriendo cada vez más relevancia en el ámbito académico y en la toma de decisiones de políticas públicas. Algunos autores (Leonhardt 2008) han llegado a escribir que la economía del desarrollo ha salido de su ostracismo e irrelevancia gracias al comienzo de la elaboración de evaluaciones de impacto aleatorizadas que, bajo el método experimental, permiten evidencias fuertes sobre lo que funciona y lo que no en las intervenciones de desarrollo (económico y humano). El artículo pretende alcanzar tres objetivos. En primer lugar, aclarar algunos términos que a menudo se usan indistintamente (específicamente resultados, efectos e impactos). En segundo lugar, contribuir a valorar de forma equilibrada las ventajas y limitaciones que tienen las evaluaciones de impacto experimentales o bajo el método de “experimentos aleatorios controlados” (también conocidos por su acrónimo inglés RCTs, de random control trials). La identificación directa de que sólo este método produce conocimiento científico y

por tanto válido, se cuestiona y se pone en relación con la utilidad y enjuiciamiento que toda evaluación debe generar. Por último, se hace una breve presentación del Fondo Español para Evaluaciones de Impacto (SIEF), como apuesta de mayor calado –político y presupuestario- que se ha realizado desde el gobierno español, haciendo notar que su mayor potencial –recibir lecciones aprendidas y articular la información disponible de forma amable y sintética para la toma de decisiones- es una dimensión del SIEF que aún está por implementar. Estos tres objetivos, coinciden con la estructura del resto del artículo, que termina recogiendo las principales conclusiones.

ACLARANDO ALGUNOS TÉRMINOS

En esta sección me propongo aclarar el sentido de algunos términos referidos a la evaluación de impacto que generan confusión. Las fuentes de confusión son dos: por una parte están las debidas a las traduccio-

nes de los términos ingleses; por otra parte, diversos organismos internacionales han utilizado términos iguales para hacer referencia a conceptos distintos.

EL CONCEPTO DE IMPACTO

Creo que el punto de partida debe ser el Glosario del CAD (CAD 2002). En él se define “impacto” como “los efectos de largo plazo positivos y negativos, primarios y secundarios, producidos directa o indirectamente por una intervención para el desarrollo, intencionalmente o no”. Es decir, se adopta un enfoque extenso y holístico. El contexto será altamente condicionante para el juicio evaluativo. Este enfoque es el adoptado por redes internacionales como la “Red de Redes en Evaluaciones de Impacto” (Network of Networks on Impact Evaluation, NONIE (2009:ix).

Algunos autores y organizaciones de desarrollo, como la International Initiative for Impact Evaluation (3ie), el Abdul Latif Jameel Poverty Action Lab (J-PAL) asociado al MIT o Innovations for Poverty, asociada a la Universidad de Yale, “practican” evaluaciones de impacto entendiendo por ellas las que son realizadas bajo diseños experimentales (los que utilizan un grupo de población sobre la que se interviene y otra sobre la que no, a modo de control). Es decir, existe una definición más estrecha de evaluación de impacto. Aquella que en la que evaluar el impacto es responder a la pregunta del contrafactual: cuál sería la situación de la población intervenida de no haber intervenido. Es decir, buscan acercarse al conocimiento causal de la forma más precisa posible. Para algunos esta precisión, será la “científica”, al poder atribuir los efectos netos de una intervención con un máximo grado de validez interna. Es decir, si logramos eliminar los sesgos que perturban la medición del efecto producido por la intervención (especialmente el sesgo de selección, es decir, el producido por la autoselección de beneficiarios o, de forma más general, cuando el efecto de una intervención está correlacionado con el

hecho mismo de intervenir) tendremos una medición insesgada, válida dentro del sistema de intervención. Si eso es “ciencia”, conocimiento insesgado de la realidad medible, las evaluaciones bajo diseños experimentales serán “científicas” y las otras no.

Recordemos que la evaluación no es una disciplina que nazca con el deseo de “hacer ciencia”, sino con la de generar información útil para aprender, controlar procesos y poder tomar decisiones con mejor información. La evaluación pretende levantar la información más ética, útil y rigurosa posible. Por tanto pretende realizar juicios válidos (lógicamente) y verdaderos (epistemológicamente), pero no se va a “obsesionar” por si eso es el fin de la ciencia o no. Las políticas se orientan por muchos valores (desde el “interés general” a “ganar unas elecciones”) entre los que no creo que se encuentre la “cientificidad”.

En resumen, la concepción “canónica” (por más antigua) del impacto establecida por el CAD como uno más de los criterios de evaluación, tiene un alcance extenso y holístico. El contexto es un factor determinante del análisis.

En la concepción “estrecha” más moderna el impacto, toma algunas connotaciones diferenciales: se ciñe preferentemente a los diseños experimentales y, en segundo rango, a los cuasi-experimentales; se centra en la medición de los efectos netos de la intervención; toma en serio los sesgos de selección y aborda de forma prioritaria la atribución, por ello utiliza los grupos de tratamiento y control como opción metodológica más perfecta; el concepto de impacto será asociado al de respuesta al contrafactual.

Debemos aclarar entonces que, para la concepción “estrecha” de impacto, responder al contrafactual es el fin primario de una evaluación. Pero utilizar grupos de control no es la única manera de obtener esta respuesta. Se pueden utilizar otros métodos como el análisis de experimentos naturales o la simulación

computacional si disponemos de series históricas de datos. Dicho esto, es cierto que si pretendemos evaluar el impacto de una intervención de desarrollo que se va a desempeñar en el presente y durante unos años en adelante, una forma altamente eficiente de emprender el diseño de la evaluación es adoptando un enfoque experimental que permita recoger información cuantitativa y cualitativa tanto de un grupo sobre el que se interviene como sobre otro “idéntico” en promedio excepto en el “tratamiento” de la intervención. Si el tamaño muestral es lo suficientemente grande (y eso lo determina el tipo de intervención y el número de variables independientes), podré obtener una medición insesgada del efecto promedio debido exclusivamente a la intervención y no a otros posibles factores. En otras palabras, ahora el contexto es que el evaluador pretende “aislar” lo más posible. Siendo poco riguroso, pero con intención de ser gráfico, la evaluación de impacto “estrecha” quiere “un número sin contexto”, mientras que la definición “clásica” del CAD hace del contexto precisamente su objeto de análisis en profundidad. No le importa obtener “un juicio de los hechos en su contexto, sin número”.

Un detalle importante es que para el sentido “amplio” de la evaluación de impacto, es posible diseñar la intervención únicamente ex –post. Pero el sentido estrecho de evaluación de impacto, requiere –en la mayoría de los casos- un sistema de información que empiece a funcionar desde el comienzo mismo de la intervención.

Creo que no tiene sentido preguntarse qué definición es mejor. Simplemente son dos abordajes diferentes. Para algunas evaluaciones, la información requerida de la evaluación de impacto deberá ser cuantitativa, “exacta” y con atribución (por ejemplo, cuando la pregunta por la eficacia de una intervención –¿funciona este subsidio o no?- y su eficiencia respecto a otras alternativas). El contexto deberá “aislarse” lo más posible. Por el contrario, en otras evaluaciones lo que

interese sea precisamente el contexto: Por ejemplo, si el promotor de una evaluación pregunta ¿en qué circunstancias y con qué características poblaciones esta intervención ha mostrado ser eficaz para este grupo concreto? En este caso, el acento queda puesto en la heterogeneidad de los efectos. El diseño experimental devuelve un valor promedio de la intervención insesgado. Pero no puede responder a la heterogeneidad de los efectos, a no ser que se complete con un análisis cualitativo o con un enfoque de evaluación basada en la teoría.

Por último, no hay que olvidar que en muchas de las intervenciones (y quizá en la mayoría de las políticas públicas, no así en sus programas o proyectos) el diseño experimental simplemente no es posible, con lo que el sentido del “impacto” en la evaluación ha de ser necesariamente el clásico del CAD.

IMPACTO Y RESULTADOS

Otra confusión frecuente es la referida a evaluación de resultados. Primero hay error de traducción. A menudo no se cae en la cuenta que los documentos oficiales sobre la eficacia de la ayuda (por ejemplo la Declaración de París y la Agenda para la Acción de Accra) se expresan en términos de orientación hacia resultados y no “de” resultados.

Pero también es cierto que “resultados” es un término que ha sido traducido tanto para outputs (quizá porque productos suene muy “economicista”) como para outcomes. Creo que la principal fuente de confusión (al menos en España) ha sido que el lenguaje del Marco Lógico ha acuñado “resultados” como el fruto de las actividades programadas, de forma que los “resultados” formulados y sobre los que se da cuenta en el informe de seguimiento y final, han sido siempre el grado de cumplimiento de las actividades previstas.

Mi opción es traducir los outcomes como los efectos o cambios en la vida de las personas (con independencia de la atribución a la intervención o no). El Glosario del CAD que dedica una sección a la gestión “basada en” resultados, define el término como “Producto, efecto o impacto (intencional o no, positivo y/o negativo) de una intervención para el desarrollo. Términos conexos: efecto directo, repercusión, impacto” y “Cadena de resultados” como “la secuencia causal de una intervención para el desarrollo que estipula la secuencia necesaria para lograr los objetivos deseados, comenzando con los insumos, pasando por las actividades y los productos, y culminando en el efecto directo, el impacto y la retroalimentación. En

algunos organismos, los destinatarios (o el « alcance ») forman parte de la cadena de resultados. Términos conexos: supuestos, marco de resultados” (CAD 2002: 33).

Queda aquí claramente mostrada como la confusión está más que justificada.

La oficina comunitaria de cooperación para el desarrollo, EuropeAid, aborda esta cuestión en sus metodologías de evaluación. Para esta agencia, existe un matiz temporal entre resultados (efectos iniciales o de corto plazo) y los impactos que se producen a más largo plazo.

Figura 1. Significados polisémicos de impacto y resultados entre la OCDE y la Comisión Europea.

Riesgo de malos entendidos		
	OCDE	CE
Impacto	... efectos a largo plazo producidos por una intervención para el desarrollo...	Un término general utilizado para descubrir los efectos de una intervención en la sociedad...
Outcome (traducido como efecto directo)	Los efectos, a corto o a medio plazo probables o logrados por los productos de una intervención.	Los impactos a más largo plazo, por lo general expresados como consecuencias socio económicas en el sentido amplio de la palabra.
Resultado	... productos, efecto o impacto (intencional o no, positivo y/o negativo)	El impacto inicial de una intervención...

Fuente: Glosario de evaluación de EuropeAid: http://ec.europa.eu/europeaid/evaluation/methodology/glossary/glo_es.htm

La polisemia se agudiza cuando se comparan el significado de producto, resultado e impacto entre la guía

de evaluación de EuropeAid, el glosario del CAD y las directrices del Presupuesto comunitario. Como refleja

Figura 2. Diversidad de acepciones de producto, resultado e impacto.

Lógica de intervención: comparación de los conceptos clave							
Nivel	Esta guía		OCDE CAD		CE DG Presupuesto		
1	Producto	Efecto	Producto	Resultado	Producto		
2	Resultado		Outcome		Resultado	Resultado	Impacto
3	Impacto		Impacto		Outcome		
1	Bienes y/o servicios producidos/entregados por la intervención						
2	Cambio inicial atribuible a la intervención pública						
3	Cambios a más largo plazo atribuibles a la intervención pública						

Fuente: Glosario de evaluación de EuropeAid: http://ec.europa.eu/europeaid/evaluation/methodology/glossary/glo_es.htm

la Figura 2, en el ámbito comunitario se mantiene la asociación de resultados al corto plazo (nivel 2 de la comparación) aunque incluye la atribución del efecto, cosa que no ocurre con la acepción del CAD. Nuestra opción ha sido la de reservar la atribución para el impacto ni con nuestra definición “estrecha” de impacto.

En conclusión y de aquí en adelante, el sentido de los términos que emplearemos es el siguiente:

productos (outputs) serán los indicadores que informen del grado de cumplimiento de las actividades previstas en la formulación del Marco Lógico de un proyecto o programa;

resultados o efectos (outcomes) serán los cambios que las intervenciones han producido en la vida de las personas, con independencia de que puedan atribuirse de forma exclusiva o causal a la intervención; así una evaluación de resultados indagará en los cambios habidos en la vida de las personas utilizando testimonios recogidos en entrevistas o indicadores cuantitativos observables (mejoras en la infraestructura

del hogar, grado de morbilidad de los miembros de una familia; desnutrición observada a través de medidas antropométricas, etc.); esta evaluación informará de la contribución de la intervención a dicho cambio pero suspenderá el juicio sobre la atribución causal;

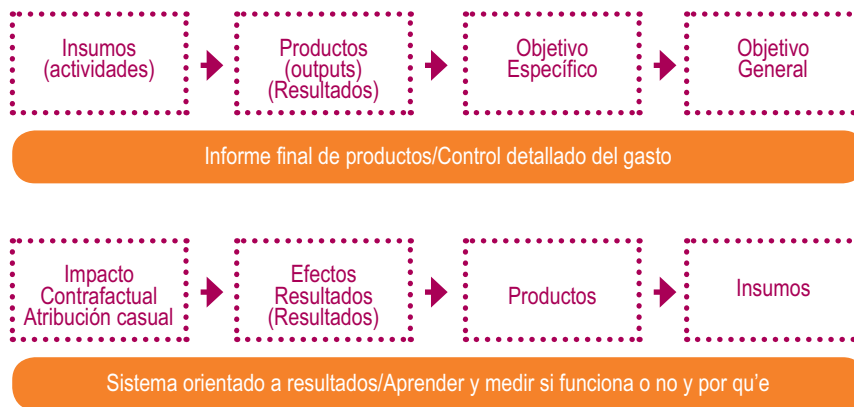
impacto será la respuesta a un contrafactual con capacidad para atribuir el cambio medido a la intervención realizada.

La figura 3 expresa dos vías para encadenar las fases de una intervención. La primera recoge el sistema de gestión predominante hasta ahora en la Cooperación Española en el que el protagonismo recae sobre el informe final. Éste debe dar cuenta de los productos obtenidos y justificar de forma precisa el gasto incurrido para obtener dichos productos. Pero apenas informa sobre la sostenibilidad futura de estos productos ni de su utilidad real en la vida de las personas. Pongamos un ejemplo ilustrativo. El informe final de un programa de formación de microempresarios revelará el número de capacitaciones realizadas sobre un número determinado de “beneficiarios”. Pero desconocemos

si esta formación ha servido para las microempresas estrenadas fruto de la formación sobrevivan dos o tres años y si los receptores de las capacitaciones mejoraron su situación personal en cuanto ingresos o calidad de vida. Es decir, podemos estar apoyando procesos que no dan fruto en términos de desarrollo.

Siguen siendo pobres. Capacitados, pero pobres. Por no ahondar en que han podido existir efectos perjudiciales como un aumento de la carga familiar o personal de trabajo, menos tiempo libre, menor calidad de vida por mayor estrés y preocupación por sacar adelante un negocio que no es rentable.

Figura 3. Encadenamiento de productos y de resultados.



Fuente: elaboración propia.

LA PREGUNTA PRECEDE AL MÉTODO

Lo primero que debemos preguntarnos cuando diseñamos una evaluación de impacto, no es qué herramientas debemos usar o qué diseño metodológico emplear, si es el experimental o cuasi-experimental. La supuesta "jerarquía de métodos", si es que existe, no es más una clasificación funcional bajo criterios de "cientificidad". Pero el valor de la ciencia y sus criterios de verdad no tienen por qué coincidir con los de una evaluación. La evaluación busca un juicio práctico que ayude a responder a una (o varias) pregunta concreta. La utilidad es básica y ésta queda marcada por el contexto socio-político de cada momento. La ciencia, busca responder de forma lo más universal posible a la verdad, entendida como adecuación a la reali-

dad, con independencia (hasta donde sea posible) del contexto.

En la discusión que tuvo lugar en El Cairo en un seminario organizado por NONIE (29-marzo-2 abril, Perspectives on Impact Evaluation) varios expertos fueron inquiridos a diseñar tres intervenciones muy diversas. La mayoría de ellos comenzó su intervención apuntando que el diseño que se iba a comentar a continuación era secundario. Lo primero era preguntarse por qué se promueve esta evaluación, quién se beneficiará de ella, y cómo será utilizado y entendido el "impacto" (Chambers 2009). Ravallion (2009) sintetizó las preguntas clave previas al diseño concreto de una evaluación de impacto en las siguientes:

- ▶ ¿Por qué se ha realizado esta intervención?
- ▶ ¿Cuál es el conocimiento actual sobre la cuestión y cuáles son las incógnitas más importantes (knowledge gaps)?
- ▶ ¿Qué contrafactual es el que resulta ahora más idóneo? ¿Cuál es realmente el que le interesa al hacedor de políticas?
- ▶ ¿Cuáles son los efectos deseados, sobre qué periodo y cuáles los potencialmente indeseados?
- ▶ ¿Cuáles son los parámetros más relevantes a estimar (los intencionalmente tratados o los realmente tratados)?
- ▶ ¿Cuáles son los mecanismos de transmisión de efectos más probables?
- ▶ ¿Cuáles son las opciones metodológicas más adecuadas, dadas las restricciones existentes de recursos, tiempo e información?

Como vemos son preguntas que deben preceder a toda investigación. Pero refuerzan la idea anterior y el título de esta sección. Las metodologías (y sus correspondientes herramientas cuantitativas y cualitativas) son posteriores a las respuestas estratégicas (más políticas) de la evaluación. Es decir, el diseño debe responder a la utilidad que se le va a dar a la evaluación. Si ante una intervención lo principal que se desea conocer es dónde poder gastar recursos limitados (eficiencia), el método será diferente a si lo que predomina es cómo aumentar la escala de una programa, o detallar los efectos indeseados (directos o secundarios) que provoca para minimizarlos. Lo más probable (siempre el contexto puede matizar esta valoración) es que para responder al primera pregunta lo primero que debiéramos considerar es un diseño experimental (o experimento aleatorio con control, Randomized Control Trial). Pero el RCT no es

la mejor opción para predecir si los efectos se mantendrán al ampliar la escala de una intervención concreta. En este caso y en el de los efectos secundarios y heterogéneos en la población meta, los diseños con abundancia de herramientas cualitativas y participativas pueden ser más útiles y necesarios.

Una falacia que conviene advertir es adjudicar “linealmente” las categorías de: experimental-cuantitativo-riguroso-científico-válido-verdadero frente a lo no experimental-cualitativo-flexible-opinático-inválidomenos creíble.

Karlan (2009) y White (2009), entre otros, han señalado que el diseño experimental bajo los RCT no son la única forma de evaluar (nadie lo sostiene a mi modo de ver) y no se nutre sólo de cálculos sin ninguna teoría o ninguna evidencia cualitativa. Por otra parte, los diseños participativos son también muy capaces de generar datos y cálculos estadísticos que midan el impacto (Mayoux y Chambers 2005; Chambers 2007).

Rogers (2008, 2009) diferencia entre evaluaciones de intervenciones sencillas, complicadas y complejas en función de la implementación de la intervención. Cada una de ellas elabora una forma de causalidad diferente, y complementaria si es el caso. La sencilla se caracteriza porque la lleva a cabo una sola organización. El impacto se centra en “la causa” más probable que ha producido los efectos. Responde predominantemente a la pregunta de qué funciona.

La intervención complicada está formada por múltiples componentes (objetivos específicos en terminología del EML). Varias organizaciones trabajan de forma contractual bajo responsabilidades bien definidas. El análisis causal es múltiple, secuencial y/o simultáneo y en diferentes niveles. Podemos conocer qué funciona, para quién y bajo qué circunstancias.

La intervención compleja es aquella en la que inter-

vienen varias organizaciones relacionándose en un marco informal de asociación y cooperación (o competencia). Pensemos en la multitud de organizaciones que trabajaron en la reconstrucción de las zonas afectadas por el tsunami de 2005. En este caso, las rela-

ciones causales pueden ser recurrentes y “emergen” efectos dinámicos (es decir, el todo es mayor la suma de sus partes). La respuesta buscada es qué es lo que está funcionando en este contexto y circunstancias determinadas.

Figura 4. Evaluación de intervenciones simples, complicadas y complejas.

Intervención	Fines de la evaluación	Cadena causal	Ejemplo
SIMPLE (un OE y un agente involucrado lidera la ejecución de la intervención).	Saber si funciona y en qué contexto. Mejorar un programa o política ya en marcha.	A -> X No A ->no X	Transferencias condicionadas
COMPLICADA (varios OEs y varios agentes involucrados repartidas las responsabilidades de forma clara y contractual).	Valorar los impactos de un programa (múltiples componentes y agentes involucrados).	A, B, C -> X A->B B->C pero no necesariamente A->C; A^B -> X B^noC->X etc.	Construcción de infraestructuras
COMPLEJA (múltiples objetivos y acciones dinámicas; efectos dinámicos e intervenciones emergentes respondiendo a diversas necesidades y oportunidades).	Entender los procesos y cadenas causales; mejorar programas en marcha y adoptar ajustes dinámicos y contextuales. Varios efectos sobre múltiples causas producidos por varios agentes que interactúan y modifican los efectos.	<p>A <-> B <-> C ↓ ↓ ↓ Xi Xj</p>	Comisión anti-corrupción

Fuente: Elaboración propia, basada en Rogers (2009).

En resumen, la pregunta principal en la evaluación de impacto es por qué se promueve esta evaluación, quién la usará y qué tipo de evidencia necesita para dar respuesta a su iniciativa. Aclaradas estas cuestiones, se determinará a quién se debe involucrar (para Chambers es innegociable que los afectados-usuarios-beneficiarios estén directamente involucrados y participando) y por qué método optar. En tercer lugar se selecciona la combinación de herramientas cuantitativas que sea más oportuna para cada caso.

No debemos olvidar que estas respuestas no deberían obtenerse únicamente de forma ex -post, o una vez terminada o ya iniciada la intervención. La recogida de información es clave y esta debe ser levantada antes, durante y después de la intervención. En los RCT, es donde quizá esto queda más evidenciado dada la gran cantidad de datos cuantitativos a recoger y que el tamaño muestral queda determinado por el poder muestral (y margen de error) que pretendemos alcanzar.

LA APORTACIÓN ESPAÑOLA: EL SIEF

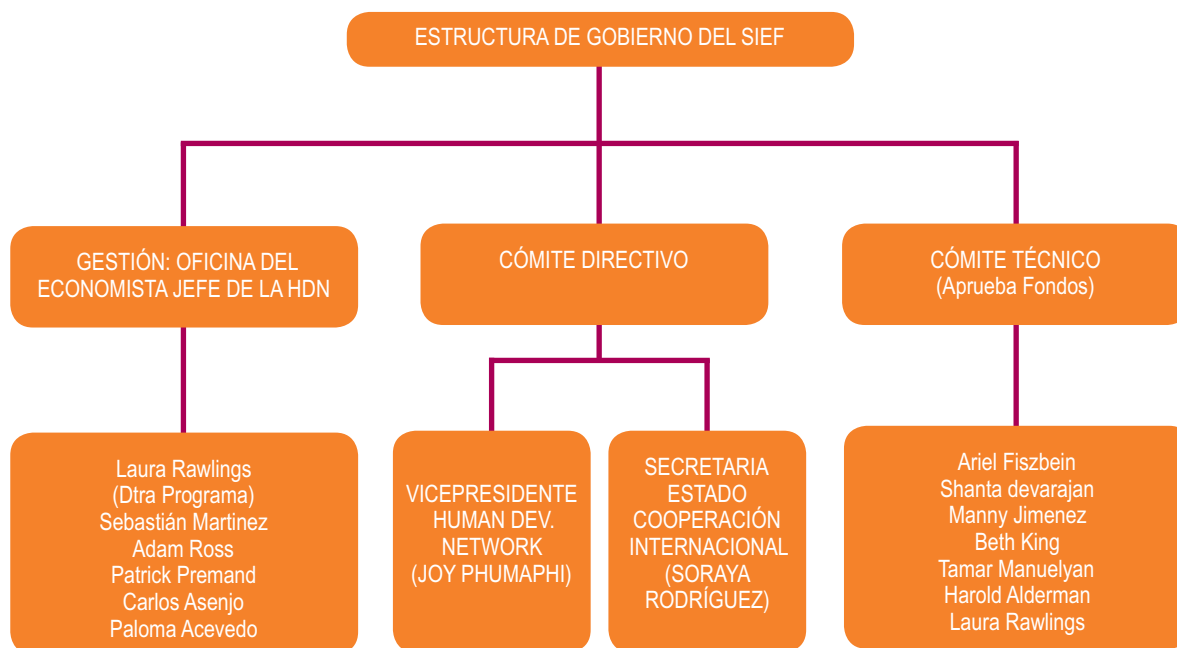
Hasta la fecha, la Cooperación Española no ha promovido ni realizado de forma directa ninguna evaluación de impacto en el sentido estricto que aquí venimos utilizando.

Sin embargo, en 2007 decidió crear el Fondo Español de Evaluaciones de Impacto (SIEF en acrónimo inglés Spanish-Wold Bank Trust Fund for Impact Evaluation)¹.

Es una iniciativa que se enmarca dentro del Grupo de Desarrollo Humano del Banco Mundial y que está dotada con 10,4 millones de dólares para el periodo 2007-2010. España ha establecido, en diálogo con el Grupo de Desarrollo del Banco Mundial, un listado de 72 países prioritarios donde realizar evaluaciones de impacto, la mayoría bajo metodología aleatorizada o experimental.

La organización y gestión del fondo queda sintetizada en la Figura 5.

Figura 5. Organigrama del SIEF.



Fuente: elaboración propia a partir de la web del SIEF.

¹ Véase <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:21407362~menuPK:384336~pagePK:148956~piPK:216618~theSitePK:384329,00.html>

Como refleja el gráfico, el fondo está organizado en tres niveles. En el nivel directivo, las decisiones de máximo nivel se toman de forma consensuada entre el Vicepresidente de la Red de Desarrollo Humano del Banco Mundial y –por parte española- la Secretaría de Estado de Cooperación Internacional. La gestión diaria queda a cargo de la Oficina del Economista Jefe de la Red de Desarrollo Humano y su equipo de apoyo. En el nivel técnico, encargado de la selección y aprobación de los fondos destinados a cada evaluación, ayudan a la economista jefe un cuerpo de seis expertos.

El SIEF tiene encomendada tres finalidades: realizar evaluaciones de impacto, realizar capacitaciones sobre las metodologías de evaluación de impacto y difundir los aprendizajes surgidos de las evaluaciones realizadas. A continuación describimos brevemente cada una de las tres.

Para realizar evaluaciones de impacto, se han establecido 72 prioridades geográficas y once sectoriales o temáticas. Las prioridades geográficas se distribuyen entre 29 países sudafricanos, 20 latinoamericanos, 8 del norte de África y Oriente Medio, 6 del Sur de Asia, 6 del Este asiático y 3 de Europa Central y Oriental. Todos los países prioritarios de la Cooperación Española están dentro de esta extensa lista, elaborada bajo criterios ad hoc². Por su parte, las once prioridades temáticas son las siguientes:

1. Soberanía alimentaria y lucha contra el hambre
2. Educación
3. Protección Social y laboral
4. Salud
5. HIV/SIDA, Malaria y otras enfermedades

2 Al menos, en la web no aparece ningún tipo de argumento técnico para esta selección, con lo que puede deducirse que ha sido la discrecionalidad política quien ha fijado éstas prioridades.

6. Igualdad de género
7. Salud Sexual y Reproductiva
8. Niñez y Juventud
9. Discapacitados
10. Agua y Saneamiento
11. Vivienda

Como puede comprobarse, el conglomerado nutrición-salud-educación, que está en la base del concepto de Naciones Unidas de seguridad alimentaria, queda ampliamente recogido pues las prioridades 1,2,4,5,6,7 y 10 tienen cabida en dicho concepto³.

El SIEF ha establecido a su vez tres tipos de fondos o convocatorias.

► Las evaluaciones innovadoras o de “resultados rápidos” (quick wins) que fueron las primeras en aprobarse⁴. Están organizadas en torno a siete temas comunes. Se presentaron 24 propuestas y fueron seleccionadas 13. Cuentan con un presupuesto conjunto de 1,5 millones de dólares.

► Convocatoria de evaluaciones dentro de los “clusters” seleccionados. Mediante concurso, han sido seleccionadas 26 Evaluaciones dentro entre las 126 propuestas, que se han agrupado en torno a 6 clusters temáticos: prevención del SIDA, transferencias

3 La definición global de seguridad alimentaria (o alimentaria y nutricional), oficializada unánimemente por los Jefes de Estado y de Gobierno de los países miembros de la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) durante la Cumbre Mundial de la Alimentación (1996) indica que existe seguridad alimentaria “Cuando todas las personas tienen en todo momento acceso físico, social y económico a los alimentos suficientes, inocuos y nutritivos que satisfagan sus necesidades energéticas diarias y preferencias alimentarias para llevar una vida sana y activa”.

4 Véase el detalle de esta modalidad en http://sitere-sources.worldbank.org/INTISPM/Resourses/383704-1184250322738/SIEF_QuickWins_Program.pdf

condicionadas en efectivo, educación básica, control de la malaria, contrataciones de desempeño en salud y, mercados de trabajo activos y para jóvenes. Está dotada con 5,6 millones de dólares.

► La última en implementarse ha sido el Fondo de Innovación. Dotado con 2,6 millones de dólares, está dividido en dos ventanillas. La primera es de Desarrollo Humano y deben realizarse sobre intervenciones dentro de los 9 sectores seleccionados que inciden de forma estrecha sobre el desarrollo humano. Bajo esta ventanilla se están realizando 5 evaluaciones. La segunda ventanilla, de Desarrollo Social, financia actualmente 8 evaluaciones dentro del ámbito temático del acceso a agua, saneamiento y vivienda⁵.

La segunda misión del SIEF es realizar cursos y sesiones formativas que aumenten la capacidad local para promover y, en su caso, realizar evaluaciones de impacto rigurosas. Desde su creación en 2007 se han llevado a cabo nueve cursos en ciudades de todo el mundo y están programados otros 3 para 2010. En Madrid se celebró el curso entre el 23-28 de junio de 2008 con 184 participantes⁶.

La tercera misión encomendada al SIEF es la de difundir los resultados obtenidos en las distintas evaluaciones de impacto. Para ello están previstas actividades como estudios de política de alta calidad que sinteticen la evidencia generada sobre aquello “que funciona” para alcanzar resultados en el ámbito de desarrollo humano; acceso público a las bases de datos y traducción al español de los resultados de las

investigaciones. Lamentablemente, hasta la fecha y que conozca, no hay disponible ningún producto que pueda incluirse dentro de esta tercera misión, lo cual es una pena ya que es la forma más directa de cómo la Cooperación Española se beneficia del SIEF. Dado que no hay ningún evaluador español dirigiendo las evaluaciones de impacto en marcha, la apropiación de los conocimientos generados por los demás debería ser la prioridad para rentabilizar la inversión del SIEF. En 2009, se destinó a Washington a un miembro de la Unidad de Evaluaciones de la DGPOLDE, para poder tener un seguimiento más cercano y directo del Fondo. Sería deseable que esta dimensión de producción de buenas y malas prácticas y resultados puedan reflejarse en decisiones estratégicas e intervenciones de desarrollo de probado mayor impacto. Evaluaciones que no conducen a cambios operativos y estratégicos no sirven de mucho, y la utilidad –junto con la credibilidad- son las dos características esenciales de las evaluaciones (y que las diferencian en parte de las investigaciones).

CONCLUSIONES

En este artículo se ha propuesto diferenciar una concepción más amplia y canónica de impacto (la contenida en el Glosario del CAD) de otra más estrecha o estricta que es aquella que responde al contrafactual. También debe diferenciarse resultados (en el sentido frecuente de productos siguiendo la terminología del Enfoque del Marco Lógica habitual en la formulación de las intervenciones de cooperación para el desarrollo) de resultados o efectos (outcomes) que se utiliza en los documentos doctrinales de la eficacia de la ayuda como la Declaración de París o de Accra.

Por otra parte, la respuesta al contrafactual puede realizarse bajo diversas metodologías y diseños evaluativos, no debiendo ceñirse exclusivamente al experimental o aleatorizado. La pregunta evaluativa

5 El detalle de las 12 evaluaciones del Fondo de Innovación puede consultarse en: http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1256762343506/SIEF_Innovation_Fund_Budget_Allocation_17nov09.pdf

6 Los detalles del programa y materiales ofrecidos pueden consultarse en la web del SIEF: <http://web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/ORGANIZATION/EXTHDNETWORK/EXTHDOFFICE/0,,contentMDK:22383415~pagePK:64168445~piPK:64168309~theSitePK:5485727,0.html>

precede al método para lograr una respuesta útil y creíble. El método experimental, como cualquier otro diseño, tiene ventajas y limitaciones (véase una discusión más amplia en Larrú 2007, 2009). Una opción, preferente para algunos evaluadores e investigadores, es responder al contrafactual bajo un diseño de RCT. Pero existen otras posibilidades, en función del contexto o de la complejidad de la intervención, como propone Rogers.

Por último, España a través del Fondo de Evaluaciones de Impacto (SIEF) ha hecho una apuesta muy

fuerte por este tipo de evidencias. Creado en 2007 y bien dotado económicamente, se echa de menos una mayor presencia de evaluadores españoles realizando evaluaciones y, sobre todo, que la dimensión de difusión de buenas prácticas a instrumentos de desarrollo exitosos o erróneos tenga una mayor entidad, pues hasta la fecha no se ha producido un ejercicio sistemático en esta materia que enriquezca a la comunidad evaluadora española, ni a los tomadores de decisiones implicados en la política de cooperación y/o de desarrollo internacional.

REFERENCIAS BIBLIOGRÁFICAS

CAD (2002) Glosario de los principales términos sobre evaluación y gestión basada en resultados. Evaluation and Aid Effectiveness N°6. OECD-DAC. París.

CHAMBERS, R. (2007) "Who Counts? The Quiet Revolution of Participation and Numbers", IDS Working Paper 296.

CHAMBERS, R. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

KARLAN, D. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

LARRÚ, J.M. (2007) "La Evaluación de Impacto: ¿qué es, cómo se mide y qué está aportando en la Cooperación al Desarrollo", Colección Escuela Diplomática N°12, 109-133.

LARRÚ, J.M. (2009) La ayuda al desarrollo, ¿reduce la pobreza? Eficacia y evaluación de la cooperación para el desarrollo. Biblioteca Nueva. Madrid.

LEONHARDT, D. (2008) "Making Economics Relevant Again", New York Times, feb, 20.

MAYOUX, L. & CHAMBERS, R. (2005) "Reversing the Paradigm: Quantification, Participatory Methods and Pro-Poor Impact Assessment", Journal of International Development 17 (2), 271-298.

NONIE (2009) Impact Evaluations and Development. NONIE Guidance on Impact Evaluation. Frans Leeuw & Jos Vaessen (authors). The Network of Networks on Impact Evaluation, Washington.

RAVALLION; M. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

ROGERS, P. (2008) "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", Evaluation 14(1), 29-48.

ROGERS, P. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

WHITE, H. (2009) "Some Reflections on Current Debates in Impact Evaluation", 3ie Working Paper 1. [también publicado en Evaluation, Vol. 16, No. 2, 153-164].