



UNIVERSIDAD LOYOLA  
Instituto de Desarrollo  
Fundación ETEA

# **LA EVALUACIÓN DE POLÍTICAS Y PROGRAMAS DE DESARROLLO**

**María José Vázquez de Francisco**

**2018**

## **LA EVALUACIÓN DE POLÍTICAS Y PROGRAMAS DE DESARROLLO**

1. Introducción: la evaluación en el contexto de la gestión de políticas públicas
2. Paradigmas, enfoques y modelos en evaluación
  - 2.1. Los paradigmas y enfoques en evaluación
    - 2.1.1. Clasificación clásica de los enfoques de evaluación
    - 2.1.2. Clasificación según modelos de gestión pública
    - 2.1.3. Clasificación según el enfoque del evaluador
  - 2.2. Los modelos de evaluación
3. La evaluación de la ayuda al desarrollo
  - 3.1. La evaluación tradicional por objetivos de la ayuda al desarrollo basada en el modelo lógico de planificación
    - 3.1.1. Los criterios de evaluación tradicional
    - 3.1.2. El modelo lógico de planificación y gestión del proyecto
    - 3.1.3. Límites y dificultades para evaluar impacto asociados al marco lógico
  - 3.2. La evaluación de impacto en intervenciones de desarrollo
4. Recapitulación

## **1. INTRODUCCIÓN: LA EVALUACIÓN EN EL CONTEXTO DE LA GESTIÓN DE POLÍTICAS PÚBLICAS**

Desde la segunda mitad del siglo XX, se ha desarrollado una abundante literatura académica y no académica sobre evaluación. El motivo es la incorporación de esta actividad (o al menos la intención de hacerlo) de manera cada vez más recurrente y sistemática en la gestión de las políticas públicas, en la que principios como la rendición de cuentas y la transparencia han ido adquiriendo mayor protagonismo.

La actividad de evaluación de políticas públicas se inicia a finales del siglo XIX en Estados Unidos y Gran Bretaña y continúa extendiéndose en la primera mitad del siglo XX, vinculada a políticas educativas y de valoración del rendimiento académico (Larrú, 2000). Durante el primer cuarto del siglo XX, se van perfeccionando los instrumentos de medición, se sistematizan y estandarizan las pruebas (tests) y se mejora la eficacia de las mediciones.

La etapa previa a la Segunda Guerra Mundial la etiqueta Larrú (2000) como la “era tyleriana”, en alusión a Ralph W. Tyler, quien creó un modelo de evaluación centrado en objetivos. La evaluación educativa sigue siendo la principal preocupación gubernamental y se incorpora como método la evaluación por objetivos y la comparación de estos con los resultados alcanzados tras la aplicación de los programas de intervención.

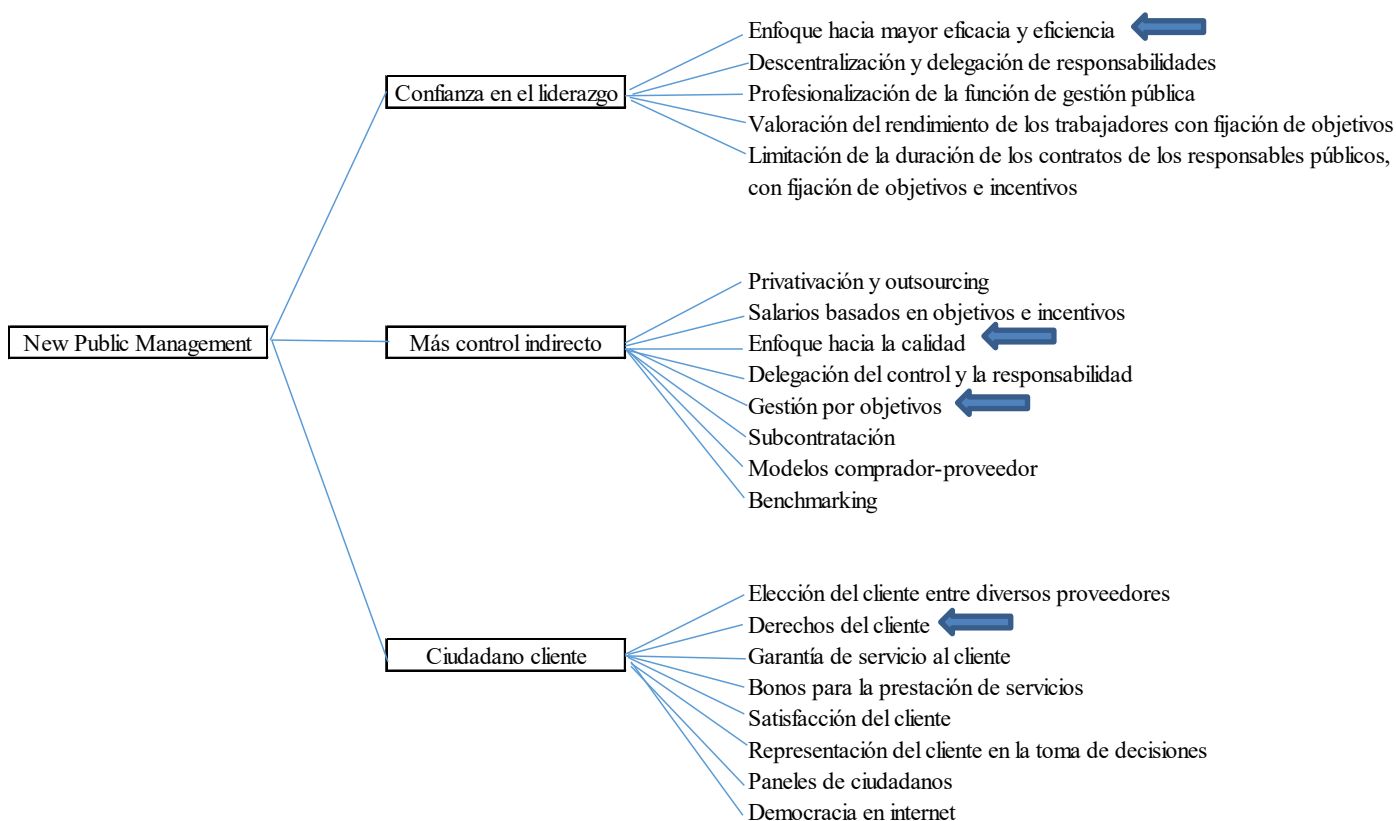
Durante los años de postguerra, la “era de la inocencia” según Larrú (2000), las políticas y los programas sociales se multiplican y se extiende la práctica ya mencionada de la evaluación por objetivos. En los años 60, la evaluación se convierte ya en una práctica generalizada en el mundo anglosajón y en el norte de Europa (Gran Bretaña, Dinamarca, República Federal Alemana y Suecia), donde surge un importante movimiento conocido como “Evaluation Research” (Nioche, 1982). El objetivo principal de la evaluación en esta etapa es valorar el efecto de las políticas sociales que de forma masiva se habían puesto en marcha tras la Segunda Guerra Mundial (AEVAL, 2010; Larrú, 2000). Pero también comienzan a surgir las críticas a esta actividad, ante la falta de capacidad para medir el efecto de las políticas y el poco y mal uso que se hace de las evaluaciones.

En otros países europeos y en América Latina, la práctica de la evaluación se inicia más tarde, principalmente a raíz de la introducción del enfoque de la gestión pública profesional, conocido como *New Public Management* (NPM), que se extiende a partir de las décadas de los años 70 y

80. Según este enfoque, muchos aspectos de la gestión pública pueden asemejarse a la gestión privada, de manera que aquella pueda adquirir las ventajas que se derivan de la gestión eficaz y eficiente, concepción más extendida en el mundo de la empresa.

Según Subirats (2005), una administración pública que rinde cuentas de forma transparente ante la ciudadanía garantiza su legitimidad para seguir ejerciendo su actividad y ostentando la cuota de poder que el estado democrático le atribuye. En la figura 1, se presenta un esquema de los elementos básicos del modelo de gestión pública NPM, donde toman protagonismo, entre otros aspectos, la gestión eficaz y eficiente, de calidad, la gestión por objetivos y la atención al ciudadano, que se concibe como cliente en la aplicación de las políticas.

**Figura 1. Elementos básicos del *New Public Management***



Fuente: Vedung (2010) (traducción propia)

La evaluación en este modelo de gestión viene a reforzar los aspectos y ventajas mencionados, pues con ella se trata de garantizar el cumplimiento de objetivos al menor coste posible y la consideración del ciudadano como parte interesada en la gestión pública.

El Libro Blanco de la Gobernanza europea establece los siguientes principios del “buen gobierno”: apertura, participación, responsabilidad, eficacia y coherencia (AEVAL, 2010). La evaluación de políticas públicas juega en este marco un papel importante como instrumento para garantizar su cumplimiento y para otorgar legitimidad a la acción de la administración pública.

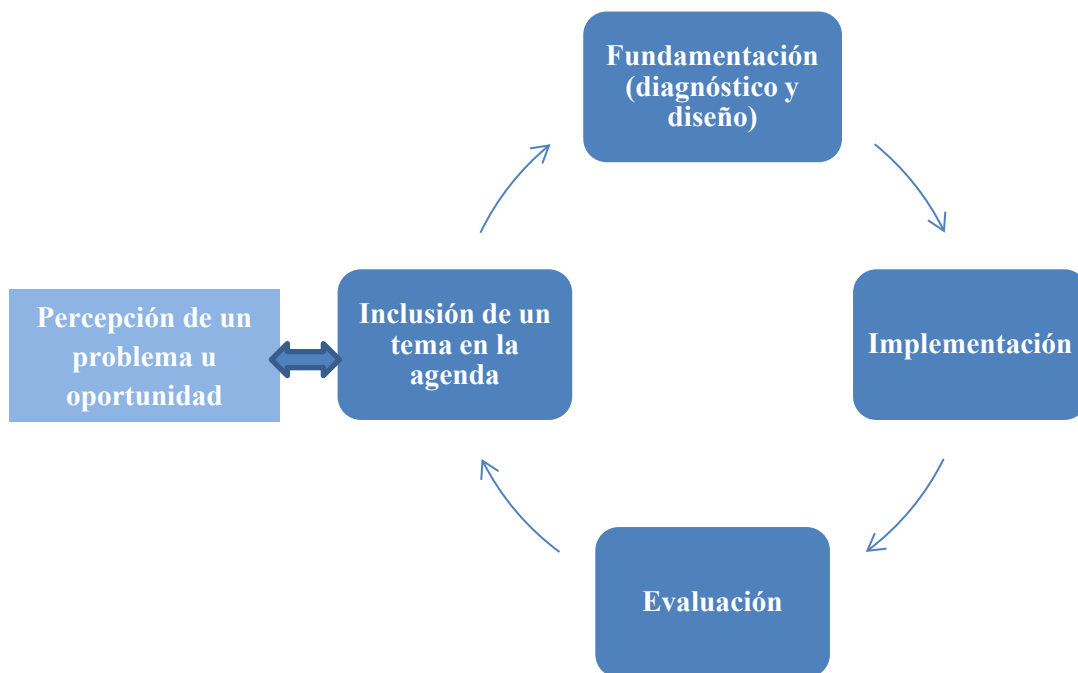
La necesidad de evaluar surge también cuando la administración pública interviene en ámbitos nuevos o con modalidades de intervención nuevas. En este caso, se trata de verificar la utilidad y eficacia de las intervenciones. Nioche indica que “allí donde no existen precedentes o conocimiento acumulado, la evaluación surge como una manera de reducir rápidamente la incertidumbre sobre los efectos de la acción en curso” (Nioche, 1982, pág. 10).

En España, la Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios define la evaluación como:

“un proceso integral de observación, medida, análisis e interpretación, encaminado al conocimiento de una intervención pública –norma, programa, plan o política-, que permita alcanzar un juicio valorativo, basado en evidencias, respecto a su diseño, puesta en práctica, resultados e impactos” (AEVAL, 2010, pág. 13).

De la anterior definición de evaluación se derivan aspectos ontológicos, epistemológicos, axiológicos y metodológicos de la concepción de evaluación. Y como se verá más adelante, existen diversas y variadas definiciones de evaluación, según la institución que la formule y el enfoque que se adopte. No obstante, para el propósito de tener una primera noción de qué es evaluación en materia de políticas públicas, esta definición de la Agencia Estatal de Evaluación española (AEVAL) es suficientemente informativa. Lo que sí es común a todas las definiciones es que con la evaluación se persigue identificar el mérito o valor de una acción, en este caso política, con el objetivo de tomar decisiones sobre la acción en sí. En la figura 2 se presenta el esquema simplificado de funcionamiento del ciclo de políticas públicas.

**Figura 2. Ciclo de políticas**



Fuente: Molina y Torres (2013), pág. 17

El efecto de la evaluación sobre el ciclo de la política puede ser doble. Por un lado, un uso de la evaluación con fines gerenciales permitirá a los gestores de políticas públicas aplicar criterios de gestión encaminados a la consecución de los resultados y objetivos marcados en la planificación de la política o programa, con una utilización racional de los recursos (vertiente gerencial de la evaluación). Este modelo de gestión se conoce como Gestión por Resultados o *Results-Based Management* (Bester, 2012; UNDP, 2007; UNEG, 2007). Por otro, el uso de los resultados de la evaluación también puede tener una vertiente política. Esta concepción es más moderna que la anterior, pues concibe la evaluación como un ejercicio de responsabilidad, una herramienta de comunicación y al servicio de los ciudadanos y un paso previo a la toma de decisiones políticas (AEVAL, 2010).

En esta última vertiente de la evaluación se enmarca el movimiento conocido como “Evidence Based Policy” o de Políticas Basadas en Evidencias (Stern et al., 2012). Este movimiento aboga por una toma de decisiones en materia de políticas públicas, que se fundamente en el conocimiento de qué es lo que funciona (*what works?*) y lo que no y, sobre todo, por qué (*why?*) (Carter, 2012), apelando a esa responsabilidad política a la que se ha apuntado anteriormente. En el capítulo siguiente se identifican los enfoques y metodologías de

evaluación que se relacionan con cada una de estas vertientes en la concepción del uso de la evaluación.

La evaluación se vuelve especialmente compleja cuando de políticas sociales se trata, no solo por la responsabilidad que implica tener que responder con ellas a necesidades humanas básicas o situaciones de exclusión o desigualdad, sino también por la diversidad y heterogeneidad de casos para los que deben funcionar y a los que deben dar respuesta. La tradicional estandarización de las políticas, propia del original modelo de Estado de Bienestar, ha dejado de ser válida (Subirats, 2005).

Lo mismo puede decirse de la política de cooperación al desarrollo y de su evaluación. En este caso, si bien esta política se enmarca en el arco más amplio de la política exterior de un estado, no se debe olvidar que su objetivo último es similar al de las políticas sociales domésticas, solo que en este caso esas políticas sociales son las de otro país que recibe la ayuda<sup>1</sup>. Cuando se habla de pobreza, sea dentro o fuera de las fronteras, se habla de un fenómeno multidimensional que debe ser tratado desde muy diversos aspectos sociales, políticos, económicos, psicológicos o culturales. Saber lo que funciona y lo que no en política social o en políticas de desarrollo, es un ejercicio de responsabilidad para con la ciudadanía, que debe involucrar a todos los actores a lo largo de todo el ciclo de programación. La evaluación debe ser una fuente de información permanente y necesaria para dotar de evidencias a las decisiones políticas.

La teoría de la evaluación ha tenido una evolución prolija, paralela al desarrollo de la ciencia social a lo largo del siglo XX. En el siguiente apartado se abordan los diferentes paradigmas, enfoques y modelos de evaluación que han surgido a partir de la segunda mitad del pasado siglo.

---

<sup>1</sup> Consideramos que las Organizaciones No Gubernamentales de Acción Social (AS) y de Desarrollo (ONGD) que actúan en este u otros países para luchar por la erradicación de la exclusión y la pobreza, entre otros sectores, se ubican en este mismo ámbito de acción de políticas sociales a nivel doméstico y en el exterior. Allá donde la acción de los estados no llega, ya sea por cuestiones de especialización, por eficacia o eficiencia, por proximidad o por cualquier otro motivo, las ONG ocupan ese espacio, financiadas con recursos propios o directamente por los propios estados, por lo que también en este caso se convierten en actores dentro de las políticas y programas.

## 2. PARADIGMAS, ENFOQUES Y MODELOS EN EVALUACIÓN

La revisión de la bibliografía sobre evaluación resulta especialmente compleja, por la abundancia de literatura existente sobre esta actividad, por la variada evolución teórica que ha tenido, según el paradigma imperante en cada momento, y por el gran número de metodologías que investigadores y evaluadores han ido diseñando, según las necesidades de cada caso.

Algunos autores (AEVAL, 2010; Alkin, 2012; Calidoni-Lundberg, 2006; Carden y Alkin, 2012; Carter, 2012; El Dessouky, 2016; Larrú, 2000; Mertens et al., 2012; Palumbo, 2001; Stern et al., 2012; Vedung, 2003, 2010) han realizado revisiones sistemáticas muy completas de autores y modelos con diferentes y útiles clasificaciones según épocas históricas, paradigmas, enfoques y metodologías de investigación, que contribuyen de forma muy útil a entender el contenido y significado de cada uno de ellos para cada autor, su sentido, aplicabilidad y utilidad.

A pesar de la abundancia de producción científica en torno a la evaluación, existe un profundo debate en torno a si esta disciplina puede categorizarse o no como científica, “puesto que su razón de ser es la actividad institucional y su fin último la responsabilidad de los poderes públicos frente a los ciudadanos” (AEVAL, 2010). Sin embargo, no queda tan claro para otros autores, que sí la consideran una disciplina dentro de la ciencia social.

Shadish (1998), quien accedió a la presidencia de la *American Evaluation Association* en 1997, apunta que este ámbito de investigación es “primarily a practice-driven field” (un campo orientado principalmente por la práctica), en el que la teoría es conducida por la práctica y aquella se formula a partir de la experiencia en esta.

Shadish (1998) pone de manifiesto que la teoría de la evaluación es “lo que somos [los evaluadores]” y trata de sistematizar los factores que perfilan la teoría de la evaluación en torno a los siguientes aspectos: 1º) el lenguaje, vocabulario y conceptos propios, que no se encuentran en ningún otro ámbito de la ciencia; 2º) los debates acerca de los distintos paradigmas y metodologías que se utilizan y sus implicaciones; 3º) las cuestiones teóricas que han guiado las conferencias de la *American Evaluation Association*: vínculo entre evaluación y justicia social, empoderamiento, políticas públicas, uso y aplicación a diferentes sectores de intervención, etc.; 4º) la identidad propia del evaluador, diferente a cualquiera otra profesión; 5º) la forma de entender e interpretar la realidad, valorarla, construir conocimiento en torno a

lo evaluado y transmitirlo de forma práctica para que resulte útil; 6º) finalmente, el conocimiento básico del evaluador, propio de su profesión y no de cualquier otra.

Scriven (Donaldson y Scriven, 2003) considera la evaluación como una transdisciplina, puesto que se utiliza y se pone al servicio de muchas otras disciplinas (Mertens et al., 2012); y la asemeja a la estadística y los métodos cuantitativos, que se consideran una disciplina en sí misma y al servicio de otras.

Rossi, Lipsey y Freeman (2004) definen la evaluación (o investigación evaluativa) como una actividad de las ciencias sociales, dirigida a recoger, analizar, interpretar y comunicar información sobre el funcionamiento y efectividad de programas sociales.

Leeuw y Donaldson (2015) consideran que la teoría de la evaluación se divide en dos bloques o tipologías: un primer bloque de teorías, que emergen desde la práctica profesional de *policy-makers, stakeholders* y evaluadores, sobre cómo hacer políticas que funcionan (sobre un modelo lógico de intervención o sobre una teoría de programa) y sobre cómo evaluar intervenciones (según el contexto, la utilidad buscada, etc.); y un segundo bloque de teorías científicas de la evaluación, capaces de explicar las consecuencias de las políticas, programas y las propias acciones evaluadoras con base en teorías sociales.

Aceptando entonces que la evaluación es una subdisciplina aplicada dentro de las ciencias sociales, cuyo fin último y sentido es conectarse y ponerse a disposición de la política, en el siguiente apartado se presentan algunas clasificaciones de autores que resultan de especial interés para conocer la evolución de la evaluación como ciencia práctico-teórica.

## **2.1. Los paradigmas y enfoques en evaluación**

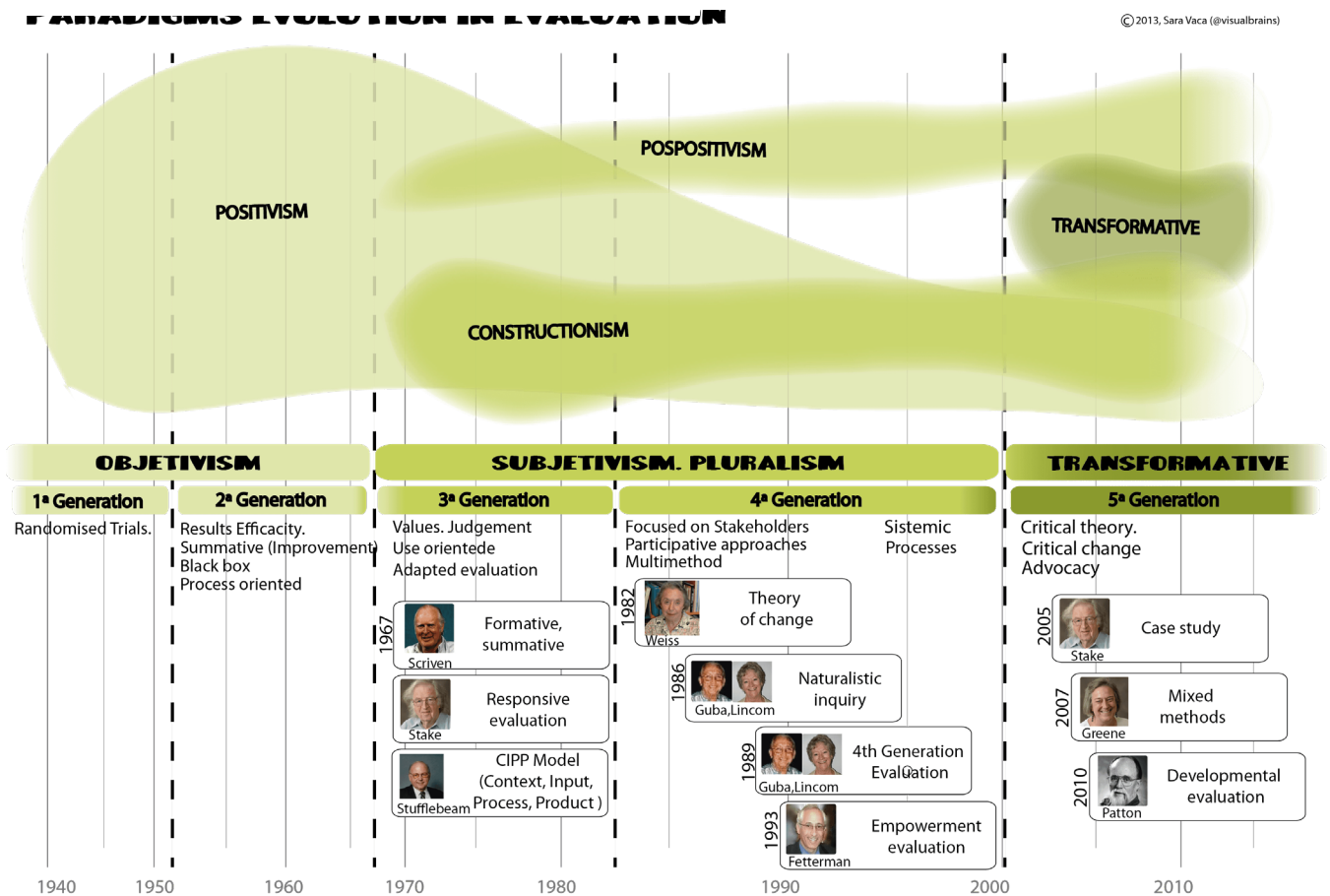
Los paradigmas en evaluación, como en las ciencias sociales, han evolucionado de forma vertiginosa a lo largo de los últimos setenta años, según las escuelas de pensamiento predominantes en cada momento. Como se ha comentado anteriormente, sistematizar el estado del arte en evaluación resulta una tarea ardua. Esta complejidad la han ilustrado varios investigadores en la materia con diferentes imágenes y metáforas, que representan las distintas ramas y ramificaciones que la teoría de la evaluación ha tenido; y esta evolución no se detiene.

### 2.1.1. Clasificación clásica de los enfoques de evaluación

Como en otras ciencias sociales, se considera que la evaluación ha pasado por diferentes paradigmas o corrientes de pensamiento que, si bien nacieron en momentos y contextos concretos de la historia reciente, fueron solapándose conforme avanzó el siglo XX.

La figura 3 (Vaca, 2013) es una representación gráfica muy ilustrativa de esa rica evolución que ha tenido la teoría de la evaluación, donde se observa cómo esas corrientes de pensamiento van sucediéndose (en los siguientes párrafos se comentarán los diferentes enfoques y modelos que aparecen en la imagen).

Figura 3. Evolución de los paradigmas de pensamiento en evaluación



Fuente: Vaca (2013)

El significado de posicionar el enfoque o método de evaluación en cada uno de estos paradigmas es poner de manifiesto la forma en la que se aborda el conocimiento de la realidad, en qué parte de esta se pone el foco de lo que se quiere conocer y cómo se construye el conocimiento a partir de la actividad evaluadora.

Los grandes paradigmas de pensamiento en la investigación social en los que se puede situar a los teóricos (y practicantes) de la evaluación en las últimas décadas son el positivismo, el constructivismo y la transformación o teoría crítica, cuyas características se definen a continuación. Además, estos paradigmas han sido también clasificados por Guba y Lincoln (1989) en “generaciones”, cuando los explican para la teoría de la evaluación:

1) El positivismo o, en rigor, el *post-positivismo* (AEVAL, 2010; El Dessouki, 2016; Guba et al., 1990; Mertens et al., 2012), emerge en EE.UU. en los años 50, vinculado a una interpretación objetiva de la realidad a través del uso de metodologías cuantitativas estadísticas. El evaluador investigador es un observador externo al fenómeno, lo analiza desde fuera. Sin embargo, se acepta que la realidad social es compleja y por tanto, no es posible encontrar una sola relación causa-efecto para explicarla (esta interpretación más purista de la investigación social fue la predominante durante el siglo XIX entre los positivistas). Los métodos experimentales con selección aleatoria, propios de las ciencias naturales, donde se puede medir y aislar la relación causa-efecto para deducir leyes de la Naturaleza, son difícilmente aplicables a los fenómenos sociales, por lo que emergen los métodos cuasi-experimentales, donde grupo de tratamiento y grupo de control no son aleatorios y, por tanto, no son estadísticamente iguales.

El enfoque de los teóricos post-positivistas de la evaluación es hacia el método (*methods-led*, según Mertens et al., 2012): el cuidado del rigor metodológico del proceso de investigación, fundamentado en técnicas cuantitativas, garantizará la validez interna del estudio y, de esta forma, una correcta interpretación y evaluación del fenómeno estudiado. La evaluación en este movimiento se realiza comparando objetivos con resultados alcanzados (evaluación por objetivos). La pregunta de evaluación en este enfoque sería “¿los resultados corresponden a los objetivos?” (Palumbo, 2001).

Según la clasificación de los teóricos de la evaluación por generaciones, este movimiento constituye la *segunda generación* en evaluación. La *primera generación* estaría integrada por los investigadores sociales del siglo XIX, que analizaban la realidad social como si de la naturaleza se tratara, positivistas puros, tratando de encontrar (y medir) una causa única explicativa de los fenómenos humanos, psicológicos o sociales, como Binet, Darwin o Stuart Mill (Larrú, 2000).

Dentro del paradigma post-positivista<sup>2</sup>, se encuentra otra corriente de investigadores evaluadores que sitúan su enfoque en el uso práctico de la evaluación (*use-led*, según Mertens et al., 2012). Preocupados por la falta de utilidad de los informes de evaluación a la que se aludió anteriormente, dirigen su preocupación más hacia el pragmatismo que hacia el método, con el objetivo primordial de que la evaluación sirva realmente para mejorar los programas de intervención. El estudio de la realidad, como ya se ha explicado, se hace desde fuera, de manera objetiva, donde el investigador toma distancia del fenómeno estudiado. La pregunta de evaluación en este enfoque sería “¿los resultados corresponden a criterios de calidad?” (Palumbo, 2001).

Las metodologías de investigación preferidas por estos investigadores evaluadores son, como en el caso anterior, los métodos cuantitativos, preferentemente cuasi-experimentales. Pero dado el carácter eminentemente práctico buscado para la evaluación, también se utilizan otros métodos que el evaluador considere adecuados, incluidos métodos mixtos cuantitativos-cualitativos.

2) El segundo paradigma de la investigación en ciencias sociales más extendido entre teóricos y practicantes de la evaluación es el *constructivismo*. Según estos investigadores sociales, la realidad no es única ni objetiva, sino subjetiva y compleja. Para conocerla, hay que acercarse al fenómeno social, para entenderlo, interpretarlo, valorarlo y juzgarlo en su contexto. Por tanto, en esta forma de acercarse al fenómeno investigado juega un papel primordial el juicio del investigador y la forma que tiene de valorar la realidad. Es un enfoque basado en valores (*values-led*, según Mertens et al., 2012). En este paradigma de investigación, Guba et al. (1989) localizan dos generaciones en evaluación.

La *tercera generación* de evaluadores considera que la realidad (objeto o sujeto evaluado) se debe valorar teniendo en cuenta su mérito propio y el valor que se le otorga (desde fuera), sin fijar unos objetivos a alcanzar previamente. Lo interesante para los defensores de este paradigma es valorar el hecho ocurrido en sí, libre de metas que deban ser conseguidas,

---

<sup>2</sup> Mertens et al. (2012) nombran a este paradigma “pragmático”, aunque se ha optado por mantener este movimiento dentro del mismo post-positivismo, tal y como plantea El Dessouki (2016), dado que este no existe como tal en la investigación en ciencias sociales.

observando la realidad desde fuera, valorándola y juzgándola. Se abandona así la evaluación por objetivos o resultados, propia del paradigma post-positivista.

En la *cuarta generación* en evaluación (Cucuzzella, 2011; Guba et al., 1989; Lay y Papadopoulos, 2007), los evaluadores no solo se acercan a la realidad para valorarla y juzgarla, sino que se adentran en ella, la viven, la sienten, la comparten con el resto de los actores y agentes que intervienen en ella y construyen así su propia interpretación de la realidad, a partir del diálogo, el debate y la reflexión. La pregunta de evaluación en este enfoque sería “¿qué ha sucedido? ¿qué es y en qué consiste el éxito de una política?” (Palumbo, 2001).

Los datos de la investigación emergen a partir de metodologías eminentemente cualitativas y participativas, aunque también pueden ser utilizadas las cuantitativas.

3) El tercer paradigma hacia el que ha evolucionado la teoría de la evaluación se vincula a la teoría crítica y a la transformación social. Se trata del paradigma *transformador*, en busca de la justicia social (*social-justice-led*, según Mertens et al., 2012). Los investigadores sociales, conscientes de las situaciones de distribución inequitativa de riqueza y poder y de las desigualdades sociales y económicas, no solo investigan e interpretan la realidad, sino que pretenden transformarla, para hacerla más justa. Estas situaciones de injusticia se producen no solo en el plano económico y político, sino también en otras realidades relacionadas con la cultura, el sexo o género, raza o etnia, los derechos humanos, las discapacidades físicas o mentales, etc. Surge así una larga lista de investigadores y practicantes de la evaluación, muchos de ellos vinculados a la actividad sobre desarrollo, pobreza y desigualdad, que hunden sus raíces teóricas y enfoques en la teoría crítica, el feminismo, el neo-marxismo o el ambientalismo.

La forma de conocer la realidad que tiene el evaluador en este enfoque es, como en el paradigma constructivista, adentrándose en la realidad y estableciendo una estrecha relación con los participantes (*stakeholders*) en la intervención mediante el diálogo (modelo dialógico) y la discusión. Las metodologías que usan estos evaluadores se basan fundamentalmente en los métodos mixtos cualitativos y cuantitativos, tendiendo a primar aquellos en el acercamiento a la realidad, aunque luego puedan tener cabida los segundos, para completar los análisis de las relaciones causa-efecto.

En este paradigma comienza a vislumbrarse el nacimiento de una nueva generación de evaluadores, la *quinta generación* en evaluación, cuyo enfoque recibe el nombre de *construccionismo* (El Dessouki, 2016; Lund, 2011; Muñoz y Mata, 2017). Según este enfoque, evaluador y evaluado establecen una conexión y aprenden mutuamente en un proceso de evaluación colaborativa. Los evaluados asumen una responsabilidad transformadora en el proceso de evaluación, que se orienta hacia el cambio y a la creación de oportunidades de mejora de las organizaciones o instituciones en las que tiene lugar la intervención (Ivaldi, Scaratti y Nuti, 2015).

En esta nueva generación podría situarse un grupo de evaluadores cuya orientación es el *advocacy* o la voluntad de influir en las esferas de poder para conseguir cambios sociales (Arensmann y Van Wessel, 2018).

En nuestra opinión, el enfoque de capacidades, dentro de la Teoría del Desarrollo Humano, y la evaluación desde este enfoque se situaría en este último paradigma transformador. Así parece también considerarlo El Dessouky (2016), quien incluye a Walker (2005) en este movimiento. El afán de hacer que toda persona pueda vivir una vida que merezca la pena vivir, así como integrar en ese proceso las voces de los que no tienen voz (George, 2015), son indicativos de un enfoque transformador que busca la justicia y el bienestar para todos los seres humanos.

### **2.1.2. Clasificación según modelos de gestión pública**

Si bien la clasificación clásica de los paradigmas en ciencias sociales son los mencionados en los párrafos anteriores (además de las generaciones en evaluación a las que ya se ha aludido), algunos académicos han hecho sus propias clasificaciones de las teorías de la evaluación según los distintos paradigmas y enfoques, lo cual facilita también su entendimiento.

Vedung (2010) considera la evolución de la evaluación en torno a cuatro “olas” que van y vienen entre los paradigmas racionalista (radical) y constructivista; y en ese ir y venir de olas (y paradigmas) se van depositando sedimentos de unos sobre otros, que los alimentan y enriquecen.

**Figura 4. Olas en la teoría de la evaluación**



Fuente: Elaboración propia, a partir de Vedung (2010)<sup>3</sup>

La primera ola en teoría de la evaluación tiene lugar, según Vedung (2010), entre mediados de la década de los sesenta y los setenta. Se basa en la ciencia (*The Science-Driven Wave*) y en los métodos experimentales, con uso de aparato estadístico potente. Vedung lo tilda de racionalismo radical, con una lógica de evaluación mediante la identificación de medios y fines de la intervención pública. Es claramente asimilable al paradigma post-positivista y encuadrable en el modelo de gestión pública por objetivos. Dentro de esta ola, Vedung sitúa el Modelo de “Ingeniería de la Intervención Pública (*The Engineering Model of Public Intervention*), que eleva el modelo de gestión pública a su versión más tecnocrática.

La segunda ola surge a mediados de los años setenta y la identifica Vedung como orientada hacia el diálogo (*Dialogue-Oriented Wave*). Según el autor, este movimiento evaluador da preponderancia a la participación en la evaluación de los *stakeholders* de la intervención. Es una evaluación democrática y deliberativa, que basa sus mecanismos de construcción del conocimiento en la discusión y el diálogo, en la racionalidad comunicativa. El paradigma imperante en esta ola es el naturalista, hermenéutico o interpretativo. Sitúa Vedung en este movimiento la *Empowerment Evaluation* (evaluación “empoderadora” o fortalecedora) ideada por Fetterman (Fetterman y Wandersman, 2007).

<sup>3</sup> Fotografía: <https://www.yorokobu.es/%C2%BFpor-que-parece-que-las-olas-viajan-en-grupos-de-siete/>

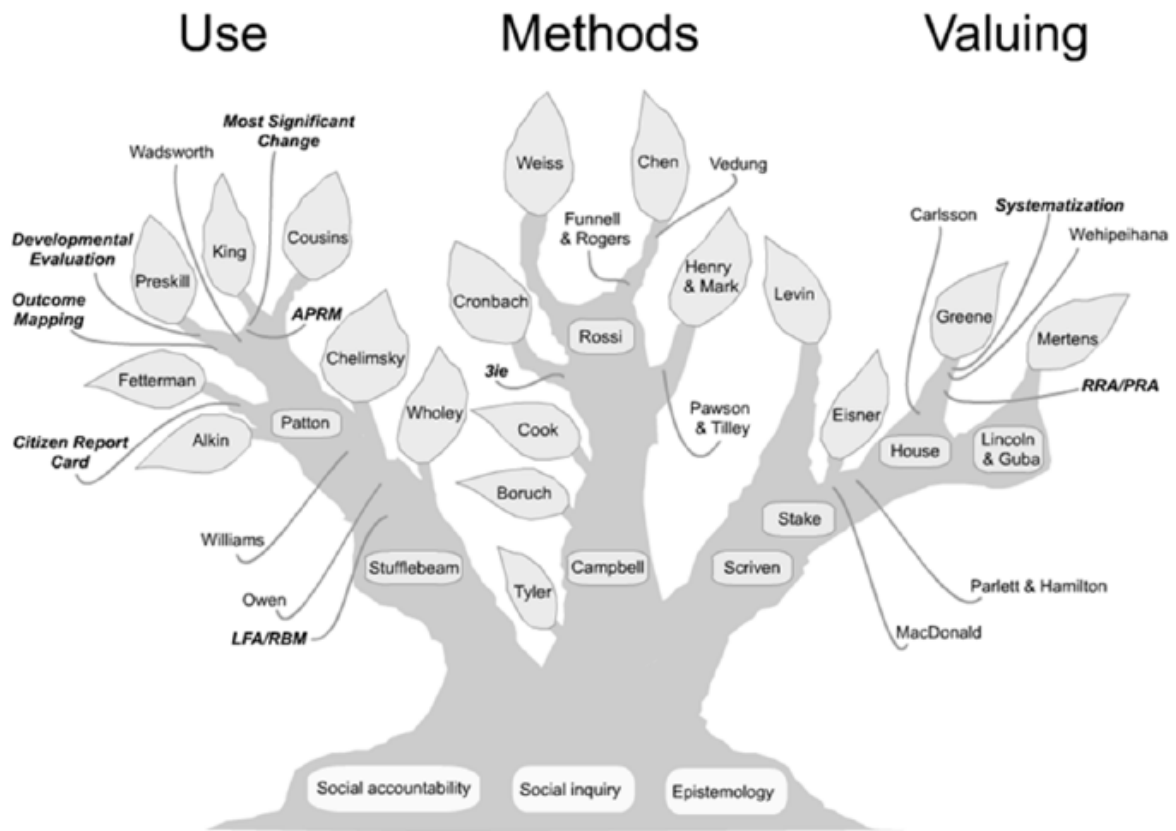
La tercera ola (*The Neo-Liberal Wave*) de evaluadores se vincula al movimiento político neoliberal que surge a finales de los años setenta y se consolida a lo largo de la década de los ochenta. El modelo de gestión pública, como ya se mencionó, se orienta hacia el mercado (*New Public Management*) y a la gestión por objetivos (*Results-Based Management*). Los participantes en las evaluaciones no son ya *stakeholders* de los procesos sino clientes. Los modelos de evaluación se orientan hacia la medición de los resultados, la eficacia y la eficiencia.

La última ola en evaluación según Vedung (2010) es la ola de la evidencia (*The Evidence Wave*). Con esta última regresa el gusto por el experimentalismo, con el objetivo de probar el impacto de las intervenciones mediante la comparación antes y después de los efectos de los programas (racionalidad medios-fines de la evaluación) en grupos que participan con otros que no participan. Las metodologías de evaluación preferidas son las experimentales, con prueba controlada aleatorizada (*randomized control trial –RCT-*) y las cuasi experimentales, donde la selección de individuos no es aleatoria, pero donde la comparación se realiza con técnicas de emparejamiento (*matching*).

### **2.1.3. Clasificación según el enfoque del evaluador**

Si bien la clasificación de los enfoques de evaluación de Vedung corre en paralelo a los modelos de gestión pública y a las ideologías políticas preponderantes durante la segunda mitad del pasado siglo, fundamentalmente en los países desarrollados del Norte, Alkin (2012) y Carden y Alkin (2012) realizan su clasificación según el foco principal de los evaluadores. Así, utilizan la metáfora de un árbol con tres ramas y tronco y raíces comunes, para ordenar teóricos y modelos de evaluación. Las raíces de la acción evaluadora, o su razón de ser, para estos autores son la rendición de cuentas a la sociedad, la investigación social y la manera de conformar el conocimiento (la epistemología). Cada rama representa sobre qué aspecto de la evaluación hacen mayor énfasis los evaluadores: el método de la evaluación, el uso de los resultados de la evaluación y el valor de lo evaluado.

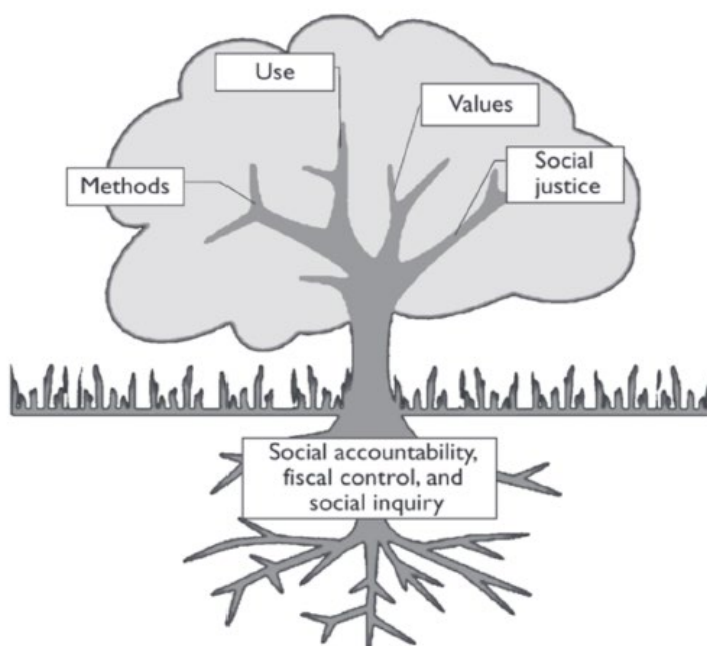
Figura 5. **Árbol de la Teoría de la Evaluación**



Fuente: Carden y Alkin (2012)

Mertens et al. (2012) añaden una rama más al árbol de la evaluación de Carden y Alkin (2012) y sitúan en ella a los teóricos y evaluadores que ponen su enfoque en la justicia social. Considera Mertens que algunos de los autores que Carden y Alkin ubican en la rama del valor, deberían situarse en esa nueva rama de la justicia social, como es el caso de House, Green, Guba y Lincoln o ella misma. Como se explicó anteriormente, la principal motivación de los evaluadores de esta nueva rama de justicia social es lograr que, con los resultados de las evaluaciones, mejore la situación de postergación y desigualdad que sufren algunos grupos humanos (figura 6).

**Figura 6. Árbol de cuatro ramas de la Teoría de la Evaluación**



Fuente: Mertens et al. (2012)

Otras metáforas se han utilizado para reflejar la riqueza y complejidad de los enfoques y modelos en evaluación. Patton (2004) habla de ríos y afluentes que se entrelazan y mezclan sus aguas y sedimentos de ideas. Y Mertens et al. (2012) identifican también esta evolución con el sistema de corrientes marinas.

En definitiva, la teoría de la evaluación ha tenido una evolución rica y compleja, en la que han emergido, se han relegado y han vuelto a aparecer, con diferentes grados de fuerza, los paradigmas y enfoques de las ciencias sociales a lo largo del siglo XX. Esta complejidad ha sido también origen de interés, pues desde cada uno de ellos se han enriquecido la teoría y la práctica de la evaluación, con el fin último común, como reflejan las raíces del árbol, de mejorar la transparencia de la gestión pública para con el ciudadano y crear conocimiento sobre qué es lo que funciona y lo que no en la intervención social.

## **2.2. Los modelos de evaluación**

En el apartado anterior se han recogido los grandes movimientos o paradigmas por los que la teoría y la práctica de la evaluación ha transcurrido a lo largo de las pasadas décadas. En este

apartado se explicitan algunos de los modelos que teóricos y practicantes de la evaluación han diseñado, según su propio enfoque de cómo debe ser ese proceso evaluativo y cómo se construye mejor el conocimiento. En cuanto a las metodologías, ya se han mencionado las principales en el anterior apartado, aunque en éste se aporta alguna información adicional. Se utilizará la clasificación de paradigmas introducido en el apartado anterior (post-positivismo, constructivismo y teoría crítica) para realizar el recorrido por los diferentes modelos.

El paradigma post-positivista en evaluación, con sus ramificaciones en el enfoque sobre el método y el uso, trajo consigo los modelos siguientes:

- 1) El modelo de evaluación basado en el análisis coste-beneficio y coste-efectividad (Comisión Europea, 2015; Ecofin, 2006), con un enfoque eminentemente economicista de las inversiones en proyectos.
- 2) El modelo de evaluación por objetivos (ZOPP, por sus siglas en alemán, *Zielorientierte Projektplanung*), o como la denominara Ralph Tyler, evaluación educativa, en la que se fijan unos objetivos a alcanzar tras una intervención y posteriormente se comprueba si los objetivos han sido alcanzados (El Dessouky, 2016; Larrú, 2000; Mertens et al., 2012;). Contribuye también en este modelo Cronbach (Carden et al., 2012), conocido por la determinación del  $\alpha$  (*alpha*) de Cronbach, como indicador de confianza en las pruebas estadísticas.
- 3) Los modelos experimentales, con grupo de control aleatorio, y cuasi-experimentales, apoyados por Campbell (1991) y más recientemente por Banerjee y Duflo (2009), Blattman, Jamison y Sheridan (2017), Gertler et al., (2011) o Ravallion (2001), o entre muchos otros, en una aproximación rigurosamente científica a la interpretación de los fenómenos sociales, en los que se comprueba la diferente evolución de dos grupos de individuos estadísticamente iguales, siendo uno de ellos intervenido (individuos a los que se aplica alguna intervención) y otro no intervenido, que sirve de control o término de comparación.
- 4) La investigación evaluadora, introducida por Rossi (Rossi et al., 2004) complementa la visión científica de la evaluación, con la utilización de encuestas y metodologías experimentales de investigación para evaluar.

- 5) La evaluación realista generativa, introducida por Henry y Mark (Mertens et al., 2012), en la que se trata de establecer el modelo causal de la intervención, con especial énfasis en los aspectos éticos de la intervención (Westhorp, 2014).
- 6) La evaluación basada en la teoría (Carter, 2012; Schmitt y Beach, 2015), donde la acción evaluadora se basa en la comprobación sobre lo evaluado del modelo causal teórico que produce los cambios sociales. De la evaluación emergen las variables que explican los cambios que se producen después de una intervención social. Algunos teóricos de este modelo son Chen y Lipsey, que introduce el concepto de “teoría del programa”, para referirse al modelo causal de una intervención (citados por Mertens et al., 2012).
- 7) La evaluación de la formación, desde este enfoque sobre el método, la desarrolla Kirkpatrick con su propio método (el modelo de Kirkpatrick) en cuatro niveles (Smidt, Baladin, Sigafos y Reed, 2009): reacción inmediata tras el programa, aprendizaje, cambio en comportamiento y actitudes, y resultados transferidos a la organización, una vez llevado a cabo el programa de formación.
- 8) El modelo CIPP (Context, Input, Process, Product) de Stufflebeam (Stufflebeam y Zhang, 2017), como evolución del modelo de evaluación por objetivos, en el que se introduce la descripción y evolución del contexto de la intervención, lo que da la oportunidad de hacer modificaciones y correcciones al programa, según se vayan produciendo variaciones en el contexto.
- 9) El modelo de evaluación orientado hacia su uso, introducido en los años sesenta por Carol Weiss (El Dessouky, 2016), quien daba prioridad a una tipología de evaluación útil para mejorar los resultados de los programas sociales públicos. El máximo exponente de la evaluación dirigida hacia la utilización (*utilization-focused evaluation*) es Michael Q. Patton (2008), quien considera que la participación de los *stakeholders* desde el inicio en el proceso de evaluación junto al evaluador mejora el conocimiento de estos sobre el proceso mismo, así como del objetivo perseguido con la acción evaluadora y despierta su interés por conocer y utilizar los resultados de la evaluación. El evaluador se convierte en un facilitador del proceso, que guía a los participantes durante el proceso de evaluación y los guía hacia el uso de los resultados. En este mismo sentido se orienta el modelo de evaluación de resultados, introducido por Wholey y Chelimsky; las tipologías

de evaluación mencionadas por Owen: evaluación proactiva, evaluación clarificadora, evaluación de seguimiento y evaluación de impacto; y el modelo de evaluación participativa práctica, de King y Cousins (Mertens et al., 2012).

- 10) El modelo de evaluación “empoderadora” o fortalecedora (*empowerment evaluation*), al que ya se aludió en el anterior apartado, fue introducido por Fetterman (Fetterman et al., 2007). En este modelo, los propios participantes son quienes llevan a cabo la evaluación, acompañados y animados por el evaluador, quien les forma para poder desarrollar su propia evaluación.
- 11) El modelo del mapeo de resultados (*outcome mapping*), diseñado desde el International Development Research Centre por Earl, Carden y Smutylo (2001), lo sitúan Carden et al. en el paradigma post-positivista, focalizado en el uso de las evaluaciones. Se ha vinculado habitualmente a intervenciones de desarrollo. Su objetivo es mapear los efectos que una intervención tiene sobre los cambios en el comportamiento, las relaciones, actividades o acciones de personas, grupos u organizaciones sobre las que la intervención ha actuado directamente. Este modelo es especialmente útil en contextos donde interactúan varias intervenciones al mismo tiempo y donde es difícil atribuir los cambios a una intervención en concreto. Se trata, por tanto, de identificar a qué cambios contribuye una intervención de desarrollo.
- 12) El modelo del cambio más significativo (*most significant change*) de Davies y Dart (2005) se contextualiza igualmente en ambientes complejos, donde varias intervenciones de desarrollo tienen lugar al mismo tiempo y donde los indicadores de avance hacia los resultados previstos no están claramente identificados. En este caso, el cambio más significativo se revela por los propios participantes en la acción mediante narración de historias (Serrat, 2009) de forma iterativa en los distintos niveles de decisión, desde el más inferior hasta los niveles superiores, de manera que se llegue a identificar el más significativo para todos.
- 13) El modelo de evaluación de desarrollo (*developmental evaluation*) (Patton, 2011) es una evolución del modelo de evaluación orientado hacia el uso de Patton, que se presentó en el punto 9. En esta ocasión, la preocupación del académico se sitúa en cómo identificar el efecto de las intervenciones en ambientes complejos y cambiantes de forma

que se puedan modificar las condiciones de las intervenciones conforme se producen variaciones en la realidad en la que se actúa, no para producir resultados de corto plazo, sino para intervenir y lograr resultados e impactos positivos en el largo plazo.

Pasando al marco del paradigma constructivista, encuentran también en él su lugar varios modelos de evaluación en los que el evaluador valora y juzga la realidad, interactuando con ella y valorando el proceso de aprendizaje mismo:

- 1) El modelo libre de objetivos de Scriven (Larrú, 2000) menciona específicamente los conceptos de mérito y valor, en el entendido de que el primero es alcanzado por el sujeto o intervenido y el segundo depende de la convención social sobre qué es lo valioso que se alcanza a través de la intervención según el contexto. En este modelo de evaluación, que Scriven utiliza fundamentalmente para evaluar procesos educativos, lo valorable es el proceso en sí, sin unos objetivos o metas prefijados a los que el sujeto deba llegar.
- 2) Scriven introduce también los modelos de evaluación formativa, que se realiza a lo largo de la intervención con el objetivo de aprender e ir incorporando lecciones aprendidas a lo largo del proceso, y evaluación sumativa, que se lleva a cabo una vez termina la intervención, para sacar conclusiones y aprendizajes.
- 3) El modelo de evaluación experta (*connoisseurship evaluation*) es introducido por Eisner (Eisner, 2017) en los años setenta, como método de evaluación educativa. Contrario a una evaluación basada en calificaciones, aprecia la descripción exhaustiva del aprendizaje; asemeja el proceso evaluativo en educación a la labor de un crítico (también utiliza los conceptos de “ojo iluminado” –*enlightened eye*- y “crítica evaluadora” –*evaluation criticism*-) que describe, de forma experta y minuciosa, una obra de arte.
- 4) El modelo de evaluación sensible (*responsive evaluation*) lo introduce Stake (Stake, 2004; Mertens et al., 2012) para referirse a una forma de evaluar que considera en primer lugar, y durante todo el proceso evaluativo, las valoraciones de los participantes y sus problemas, buscando la calidad del programa. Evita de forma expresa una forma de evaluación preestablecida, dando prioridad a las cuestiones emergentes que preocupan a los participantes en la intervención.

5) La evaluación naturalista y el modelo de cuarta generación, como fue denominado más tarde por sus representantes más destacados, son introducidos por Guba y Lincoln en los años 80 (Guba et al., 1989). Inicialmente el modelo fue llamado naturalista por los autores, en referencia a que el evaluador se inserta en el ambiente y contexto mismo de los participantes en la intervención, en oposición clara al paradigma post-positivista, en que la realidad se observa desde fuera. El evaluador en este caso se imbuye de la realidad evaluada y los *stakeholders* participan en todo el proceso, desde el diseño de la evaluación hasta el aprendizaje y la construcción de conocimiento.

Para finalizar, también desde el paradigma transformativo (Rodríguez y Monterde, 2014) y la teoría crítica, se han producido interesantes y muy variadas aportaciones de teóricos y practicantes de la evaluación. En este caso, muchos de estos modelos han venido de la mano de políticas, programas y proyectos de desarrollo, financiados con fondos provenientes de donantes de ayuda oficial al desarrollo. El objetivo de estos modelos de evaluación es poner de manifiesto, e intentar corregir, desigualdades e injusticias que sufren los distintos colectivos humanos marginados por razones de etnia, riqueza, sexo, discapacidades o irrespeto por los derechos humanos. Así se han desarrollado entre otros, los siguientes modelos:

- 1) Los modelos de evaluación rápida participativa (*rapid rural appraisal*) y evaluación rural participativa (*participatory rural appraisal*) emergen en los años setenta y ochenta en los países menos desarrollados (Carden et al. 2012) de la mano de políticas de desarrollo con el objetivo de dar voz y hacer emerger las injusticias y desigualdad que sufren los más pobres de las zonas rurales de los países pobres. Chambers es uno de sus mayores impulsores (Chambers, Karlan, Ravallion y Rogers, 2009).
- 2) Los variados modelos de evaluación que hacen referencia a la participación democrática de los *stakeholders*: evaluación democrática (de McDonald y Kushner) y evaluación democrática deliberativa (de House y Howe). Y otros modelos que se enfocan en grupos discriminados o marginados: la evaluación responsable con la cultura (de Frierson), la evaluación con enfoque de género y de grupos LGBTI o la evaluación con enfoque indigenista (de Cram y Chilisa) (todos ellos citados por Mertens et al., 2012).
- 3) Un último modelo de evaluación es cada vez más frecuente en las intervenciones de desarrollo: la sistematización (Carden et al., 2012; Jara, 1994, 2001, 2012; Martinic,

1984; Selener, 1996; Tapella y Rodríguez-Bilella, 2014a y 2014b; Vázquez, Santos, Sianes y Simón, 2015a). Con raíces en la pedagogía crítica y liberadora de Paulo Freire (1967), este método de evaluación se está utilizando cada vez más para evaluar los procesos, obtener lecciones aprendidas y empoderar a los marginados y oprimidos en intervenciones de desarrollo en los países empobrecidos. Chambers reclama la necesidad de dar la importancia que merecen a las voces de los pobres (Chambers, 1995); que ellos definan lo que es bueno o malo para sus vidas; que se les dé voz y que hablen e identifiquen sus percepciones y prioridades (George, 2015). La sistematización de experiencias, como metodología de evaluación del impacto desde un enfoque de procesos, contribuye a identificar relaciones causa-efecto y a formular teorías de cambio de las intervenciones de desarrollo, y permite incorporar la visión y las narrativas de los protagonistas de estos procesos: los receptores de la ayuda.

En este apartado se han presentado los diferentes paradigmas y modelos de evaluación que han surgido en el ámbito de la evaluación más recientemente, aplicables a diferentes tipos de intervención social. En la tabla 1 se presentan, de manera resumida, todos ellos.

En los apartados siguientes se hará especial mención a los modelos más habituales de evaluación que se han venido utilizando en el campo de la ayuda al desarrollo.

**Tabla 1. Modelos de evaluación según paradigmas de pensamiento**

	<i>Paradigma post-positivista</i>	<i>Paradigma constructivista</i>	<i>Paradigma crítico o transformativo</i>
<i>Modelos de evaluación según paradigmas</i>	-Análisis coste-beneficio y coste-efectividad (Ecofin)	-Modelo libre de objetivos (Scriven)	-Evaluación democrática deliberativa (House y Howe)
	-Evaluación por objetivos (ZOPP) o evaluación educativa (Tyler)	-Evaluación formativa y evaluación sumativa (Scriven)	-Evaluación democrática (McDonald y Kushner)
	-Evaluación experimental y cuasi-experimental (Campbell, Ravallion, Banerjee, Duflo)	-Evaluación experta ( <i>connoisseurship evaluation</i> ) (Eisner)	-Modelos de evaluación rápida participativa ( <i>rapid rural appraisal</i> ) y evaluación rural participativa ( <i>participatory rural appraisal</i> ) (Chambers)
	-Evaluación realista generativa (Henry y Mark)	-Evaluación sensible ( <i>responsive evaluation</i> ) (Stake)	-Evaluación responsable con la cultura (Frierson)
	-Evaluación basada en la teoría (Chen y Lipsey)	-Evaluación naturalista (Guba y Lincoln)	-Evaluación con enfoque de género y de grupos LGBTI
	-Evaluación de la formación (Kirkpatrick)		-Evaluación con enfoque indigenista (Cram y Chilisa)
	-Evaluación orientada hacia su uso (Weiss, Patton)		-Sistematización (Jara, Martinic, Selener, Tapella y Rodríguez-Bilella)
	-Investigación evaluadora (Rossi)		
	-Modelo CIPP (Stufflebeam)		
	-Evaluación de resultados (Wholey y Chelimsky)		
	-Evaluación proactiva, evaluación clarificadora, evaluación de seguimiento y evaluación de impacto (Owen)		
	-Evaluación participativa práctica (King y Cousins)		
	-Evaluación “empoderadora” ( <i>empowerment evaluation</i> ) (Fetterman)		
	-Modelo del mapeo de resultados ( <i>outcome mapping</i> ) (Earl, Carden y Smutylo)		
	-Modelo del cambio más significativo ( <i>most significant change</i> ) Davies y Dart		
-Evaluación del desarrollo ( <i>developmental evaluation</i> ) Patton			

Fuente: Elaboración propia

### 3. LA EVALUACIÓN DE LA AYUDA AL DESARROLLO

La práctica de la evaluación tradicional en el ámbito de la cooperación al desarrollo se extendió durante la última década del siglo XX, acompañada del incremento del número de intervenciones en esta materia, de los fondos destinados a ella y de un intenso debate sobre la necesidad de rendir cuentas, para justificar los fondos invertidos y para aprender de lo que funcionó y lo que no (Gudiño, 2016; Kogen, 2018; Porrini y Murciano, 2015). Surgen así un buen número de manuales y guías metodológicas sobre planificación y evaluación de los proyectos de cooperación, elaboradas por agencias e institutos de desarrollo (Bamberger, Rugh y Mabry,

2012; DAC-OCDE, 1995; ETEA, 2001, 2003; EuropeAid, 1999; Larrú, 2000; MAE-SECIPI, 1998, 2001; NORAD, 1990). Su objetivo fundamental es la rendición de cuentas a los propios financiadores sobre la utilización de los fondos invertidos en las intervenciones.

Según la metodología tradicionalmente utilizada para evaluar proyectos y programas ejecutados por agencias de cooperación y por organizaciones no gubernamentales de desarrollo, las evaluaciones deben realizarse conforme a cinco criterios, a saber: pertinencia, eficacia, eficiencia, impacto y viabilidad, cuyo significado y contenido se definen más adelante. Además, se han ido incorporando algunos criterios complementarios a los anteriores, a partir de los discursos sobre *eficacia de la ayuda* (OCDE, 2008) nacidos al amparo de la Declaración de París. Surgen así los criterios de coherencia, coordinación, complementariedad, alineamiento y armonización.

Como se puede observar, ya en los criterios de evaluación mencionados se incluye el criterio de impacto. Sin embargo, en los informes de evaluación tradicionalmente utilizados para rendir cuentas a los financiadores, los análisis que habitualmente se han realizado para extraer conclusiones sobre este aspecto son, en general, livianos y carentes del contenido que se le ha atribuido en los desarrollos teóricos posteriores más específicos sobre esta metodología de evaluación. También el criterio de eficacia en la evaluación tradicional es un elemento interesante para la evaluación de los resultados de una intervención, puesto que con este se pretende confirmar si los objetivos y resultados esperados se han cumplido conforme a lo previsto en la planificación del proyecto. Sin embargo, el alcance temporal de este criterio suele ser de corto plazo, mientras que el de impacto comporta una visión a más largo plazo.

En la primera década del siglo XXI, tienen lugar dos hitos que inician un cierto cambio en la visión sobre la evaluación de la ayuda al desarrollo. El primero de ellos tiene lugar en 2005, cuando se formula el discurso sobre la Eficacia de la Ayuda, que emana de la Declaración de París (OCDE, 2008), donde se hace especial énfasis en la gestión por resultados, la rendición de cuentas y la mutua responsabilidad. La agenda de la eficacia nace del cuestionamiento sobre los beneficios de la ayuda al desarrollo, después de varias décadas de cooperación internacional, y asume dos principios de fondo que explicarían las dificultades que la cooperación tiene para asegurar impactos positivos: por una parte, la falta de protagonismo de los países receptores de la ayuda al definir los objetivos y contenidos de la cooperación; por

otra, la falta de coordinación entre las propias intervenciones de la cooperación internacional y de estas con las políticas nacionales de desarrollo de los receptores. Ambos factores restan eficacia a la ayuda al desarrollo.

El segundo hito tiene lugar en el año 2006, cuando el Grupo de Trabajo “Evaluation Gap”, integrado por reconocidos investigadores y evaluadores<sup>4</sup>, del Center for Global Development (CGD), *think-tank* financiado por la Fundación Bill y Melinda Gates y la Fundación William and Flora Hewlett, publicó el informe “When will we ever learn” (Evaluation Gap Working Group, 2006), que pone en cuestión los resultados de las intervenciones de la cooperación al desarrollo y la necesidad de llevar a cabo evaluaciones de impacto, mediante el uso de instrumentos de evaluación y medición rigurosos basados en metodologías experimentales y cuasi experimentales, para demostrar qué es lo que realmente funciona en intervenciones de desarrollo social.

Por último, el tercer hito se sitúa en los siguientes acuerdos sobre Eficacia de la Ayuda de Accra (OCDE, 2008), que coinciden con el inicio de la crisis económica en el año 2008. A partir de este año comienza el declive de los flujos de ayuda al desarrollo. En sentido contrario, aumenta el número de intervenciones evaluadas, en un afán por encontrar aquello que supusiera un mayor y mejor impacto en el desarrollo de las poblaciones a quienes se destinaba la ayuda.

En los siguientes apartados, se describen los métodos más habitualmente utilizados en evaluación de intervenciones de ayuda al desarrollo y se analizan algunas características de ambos.

---

<sup>4</sup> Los autores del informe son: Co-Organizadores: William D. Savedoff, Ruth Levine y Nancy Birdsall. Miembros: François Bourguignon, Esther Duflo, Paul Gertler, Judith Gueron, Indrani Gupta, Jean Habicht, Dean Jamison, Daniel Kress, Patience Kuruneri, David I. Levine, Richard Manning, Stephen Quick, Blair Sachs, Raj Shah, Smita Singh, Miguel Szekely y Cesar Victora. Coordinadora del proyecto: Jessica Gottlieb.

### **3.1. La evaluación tradicional por objetivos de la ayuda al desarrollo basada en el modelo lógico de planificación**

#### **3.1.1. Los criterios de evaluación en la evaluación tradicional**

La evaluación clásica o tradicional de la cooperación internacional al desarrollo acoge una serie de criterios clave a tener en cuenta a la hora de valorar los alcances y la calidad de una intervención de desarrollo. Esos criterios clásicos, ya mencionados anteriormente, son pertinencia, eficacia, eficiencia, impacto y viabilidad y derivan del modelo de planificación por objetivos conocido como marco lógico.

Además, estos han sido complementados con otros, más recientes, a partir de la *Declaración de París sobre Eficacia de la Ayuda* (OCDE, 2008), en la que se recogen los acuerdos realizados por los países del Comité de Ayuda al Desarrollo (CAD) para mejorar la calidad de las intervenciones, una vez que se han ido conociendo las debilidades y amenazas, así como los vicios, que la ayuda al desarrollo generó en las décadas anteriores, tanto en los países donantes como en los países socios. Se trata de los criterios de alineamiento, armonización, apropiación, gestión por resultados y mutua responsabilidad.

Existe una abundante literatura (DAC-OCDE, 1995; ETEA, 2001, 2003; EuropeAid, 1999; Larrú, 2000; MAE-SECIPI, 1998, 2001; NORAD, 1990) en la que se detalla el significado de los criterios tradicionalmente utilizados, por lo que aquí solo se definen brevemente para conocer su contenido básico.

- 1) **Pertinencia:** mide la adecuación de la intervención a los problemas y necesidades detectados en la población destinataria de la acción, a las políticas propias de la zona en la que se desarrolla la acción, a las políticas de cooperación de otros agentes que trabajan en la zona y a las capacidades humanas y técnicas de los *stakeholders* intervinientes en el programa o proyecto.
- 2) **Eficacia:** mide el grado de consecución de los objetivos y resultados planteados en una intervención de desarrollo. Para determinar el grado de alcance de objetivos y resultados es necesario planificar desde un inicio los indicadores para cada uno de estos, su nivel de partida y la meta esperada.

- 3) Eficiencia: mide el grado de consecución de objetivos y resultados planteados en una intervención de desarrollo en relación con el consumo de recursos financieros, humanos y materiales necesarios para ello. La eficiencia será mayor en tanto en cuanto el consumo de recursos sea el mínimo necesario para alcanzar los objetivos y resultados esperados.
- 4) Impacto: mide los efectos de la acción a medio plazo, tanto los esperados como los no esperados, planificados o no planificados, positivos y negativos, sobre la población destinataria o sobre otra población afectada, aunque no haya sido intervenida. Con la medición del impacto se trata de conocer el efecto neto (atribución de efectos) que la intervención ha producido, una vez finalizada esta.
- 5) Viabilidad: con la evaluación de la viabilidad se trata de identificar la probabilidad de que los efectos y resultados provocados por la intervención se mantengan en el tiempo. En esta permanencia en el tiempo pueden influir factores socio-culturales, políticos, institucionales, económico-financieros, tecnológicos, de género o medioambientales.
- 6) Alineamiento: mide la correlación entre la intervención de desarrollo del donante y las políticas públicas y estrategias propias del país receptor de la ayuda al desarrollo.
- 7) Armonización: mide el grado en que las acciones financiadas por los donantes de ayuda al desarrollo están coordinadas y se complementan entre sí, son más transparentes y colectivamente eficaces.
- 8) Apropiación: mide el grado en que el país receptor de la ayuda al desarrollo lidera y coordina los procesos de diseño e implantación de las intervenciones financiadas por los donantes de ayuda.
- 9) Gestión por resultados: mide el grado en el que los recursos se administran de la forma más adecuada para alcanzar los resultados planificados y mejorar así los procesos de toma de decisiones orientadas.
- 10) Mutua responsabilidad: mide el grado en que donantes y receptores se responsabilizan conjuntamente de la consecución de los resultados de desarrollo planificados.

Como se comentó al inicio de este epígrafe, los criterios de evaluación tradicionales en cooperación al desarrollo se derivan del modelo lógico de planificación por objetivos, que es el

más extendido en intervenciones de este tipo. Sobre este modelo de planificación, siendo el más usado, existe una abundante bibliografía. En algunas referencias se tratan los problemas que supone este modelo de planificación, para poder realizar evaluaciones de impacto útiles, que contribuyan no solo a la rendición de cuentas, sino también al aprendizaje sobre lo que funciona y lo que no en cooperación al desarrollo en el largo plazo (Muñiz y Gasper, 2009). Esta problemática se aborda en el siguiente epígrafe.

### **3.1.2. El modelo lógico de planificación y gestión del ciclo del proyecto<sup>5</sup>**

Como ya se ha apuntado anteriormente, muchos autores sugieren que una de las fuentes de problemas y obstáculos para realizar evaluaciones de impacto de la cooperación al desarrollo se encuentra en el origen de las intervenciones, es decir, en la propia planificación de los proyectos. Esto se debe, en buena parte, a la utilización de una herramienta que es recurrentemente usada, pero también criticada (Bakewell et al., 2005), puesto que se considera poco adecuada para este tipo de intervenciones: el marco lógico. Estos teóricos apuntan a que esta metodología de planificación dificulta la posterior realización de evaluaciones de impacto, no sólo por el mal uso que se hace de ella, sino también por su propia naturaleza.

A pesar de las críticas, el marco lógico es una herramienta de planificación por objetivos que se ha convertido en el principal instrumento de planificación en cooperación al desarrollo. Los antecedentes remotos de este método los encontramos en el uso militar de la planificación por objetivos (Hummelbrunner, 2010, en Fujita, 2010). Los antecedentes y evolución más cercana los clasifica Sartorius (1996) (citado por Ferrero, 2003) en tres generaciones:

- 1) En la Primera Generación (1970-1980), a partir del trabajo de diseño realizado por la agencia estadounidense de cooperación (USAID), se utilizaba tan solo la matriz de planificación del proyecto (MPP).
- 2) Durante la Segunda Generación (1980-1990), el marco lógico se amplía al enfoque de marco lógico (EML), en el que se hace un mayor énfasis en el proceso de planificación, estructurando este en diferentes fases y con la participación de diferentes agentes

---

<sup>5</sup> Parte del contenido de este apartado fue publicado en Vázquez et al. (2015b).

interesados (la agencia alemana de cooperación GTZ crea el método ZOPP –Zielorientierte Projektplanung-).

- 3) Finalmente, en la Tercera Generación, en la que se puede decir que nos encontramos, se incorpora la gestión del ciclo del proyecto, lo que implica una forma específica no solo de planificar, sino también de gestionar, hacer seguimiento y evaluar la cooperación al desarrollo.

### **La matriz de marco lógico**

La matriz de marco lógico (MML) (Gasper 1997, 1999, 2000a, 2000b) o matriz de planificación del proyecto (MPP) es una tabla de cuatro filas por cuatro columnas en la que se recogen diversos componentes ordenados según dos lógicas complementarias:

- la lógica vertical (las columnas), basada en relaciones de causa-efecto y medios-fines;
- y la lógica horizontal (las filas), encaminada a la comprobación o medición de los resultados alcanzados (*accountability*), a través de la definición de indicadores (objetivamente verificables) y las fuentes donde verificar su cumplimiento (fuentes de verificación).

En la lógica vertical se encuentra además una casi siempre denostada última columna de hipótesis, en la que se plasman posibles contingencias, imponderables o premisas que, en caso de que ocurrieran, romperían la correcta consecución de la lógica vertical e impedirían, por tanto, que se pudieran alcanzar los resultados u objetivos previstos. Esta última columna, generalmente olvidada, es sin embargo fundamental de cara a la realización de futuras evaluaciones de impacto, precisamente por la razón anteriormente expuesta.

En la primera columna de la matriz, de arriba abajo y en cuatro niveles, se recogen los elementos principales definitorios de un proyecto de cooperación al desarrollo, consistentes en:

- 1) El objetivo general, o fin último al que se pretende contribuir con una intervención de desarrollo, que suele hacer referencia a una mejora en el nivel de bienestar de una población, territorio, etc.; según la terminología inglesa, este sería el impacto (*impact*) de la intervención, una vez cuantificado el cambio; por tanto es esencial tenerlo en cuenta en una futura evaluación de impacto: una mala o errónea formulación del

objetivo general supondría ya de por sí un obstáculo o la imposibilidad de realizar la evaluación de impacto de la intervención (proyecto o programa).

- 2) El objetivo específico (pueden ser varios, aunque no es recomendable según esta metodología) o propósito concreto del proyecto (*outcome*).
- 3) Los resultados (*outputs*) o productos o servicios que el proyecto va a generar.
- 4) Las actividades y acciones que es necesario llevar a cabo para conseguir generar esos productos o servicios<sup>6</sup>.

Al respecto de estos cuatro niveles, uno de los problemas que presenta la simplificación de la realidad es la sobre-agregación de conceptos o *jamming* (Gasper, 1997) en los niveles de resultados y objetivos, resultando en muchas ocasiones confusos y mal formulados. Según Cordingley (1995), Cracknell, (1986) y Eggers (1992, 1994), se confunden resultados con objetivos y viceversa; o como indica Gasper (2000), en ocasiones faltan niveles para representar correctamente las relaciones de causa-efecto (*lack-frames*). Además, esa simplificación de la realidad representa de manera muy pobre los procesos de cambio (Davis, 2002), pues el marco lógico no incorpora la dimensión temporal (Gasper, 1997).

Los elementos anteriormente mencionados pueden ser divididos en dos estratos o pisos de importancia distinta, de cara a una futura evaluación de impacto de una intervención, en la medida en que lo realmente relevante en esta será la comprobación del alcance o contribución a la consecución del objetivo general y el objetivo específico. Sin embargo, la práctica indica que, donde las entidades ejecutoras de proyectos focalizan los esfuerzos durante una intervención de desarrollo, es en la consecución de los resultados y las actividades necesarias para ello, pues es ahí donde los financiadores pondrán mayor atención para revisar que el proyecto se ha ejecutado conforme a lo planificado y para lo que se concedieron los fondos (Solem, 1987). Es más, autores como Cracknell (1989), Gasper (2000) y Solem (1987) afirman que los formuladores muchas veces cumplimentan la matriz de marco lógico después de haber

---

<sup>6</sup> Se ha añadido la traducción inglesa de los diferentes componentes por dos razones: en primer lugar, porque es más abundante la literatura inglesa sobre este campo y para los castellano-hablantes resulta a veces confusa esta terminología; en segundo lugar, porque a nuestro entender, es mucho más evidente que el cambio en el objetivo general es el impacto esperado o la contribución al mismo, y por tanto, una buena definición del mismo es fundamental de cara a la realización de una futura evaluación de impacto.

concebido el proyecto, para cumplir con el requisito exigido por los financiadores, obteniendo así matrices ilógicas que pierden las relaciones de causalidad entre niveles, a lo que Gasper llama *logic-less frames* o marcos lógicos sin sentido.

En cuanto a los indicadores de medición del grado de avance hacia los objetivos específicos y los resultados, se encuentra también aquí una fuente de futuros problemas para cuantificar el impacto de los proyectos, pues en ocasiones se diseñan indicadores poco realistas o difíciles de medir, no solo por su propia definición, sino también por el alto coste que supondría la recolección de datos y su valoración (Chambers, 1995; Chambers y Guijt, 1995).

### ***El enfoque de marco lógico***

El enfoque de marco lógico (EML) es una versión más avanzada del método, pues a la matriz se le añade una metodología de trabajo, estructurada en diferentes fases y en la que se espera que participen no solo el equipo técnico que diseña el proyecto, sino la población destinataria; se basa en lo que se conoce como diagnóstico y planificación participativos. Fruto de ese diagnóstico, se obtiene el amplio y complejo conjunto de problemas que afectan a una población concreta, que se disponen en un simple esquema según relaciones lineales de causa-efecto, que se conoce como árbol de problemas. Estos problemas, cuando se formulan en positivo, como situación ideal a la que aspirar, y combinando acciones que se relacionan mediante la lógica de medios-fines, se convierten en el árbol de objetivos. De las posibles líneas de intervención (o ramas del árbol de objetivos, que reflejan sucesiones de causas y efectos) se elige una. Esta se traslada a la primera columna de la matriz de marco lógico, que supone, por tanto, el corazón de esta herramienta de planificación.

Las críticas principales que se hacen a esta metodología hablan del falso consenso que se alcanza a través de las metodologías participativas para la elección de una alternativa de intervención, así como de la imposición por parte de los diseñadores de los proyectos de una forma de discurrir propia del pensamiento occidental, que no tiene por qué ser válida y universal para todas las culturas (Chambers, 1995; Chambers et al., 1995).

Esta simplificación de la realidad (Chambers, 1995; Chambers et al., 1995; MacArthur, 1994), si bien ayuda a la hora de focalizar las acciones de cooperación, supone la pérdida de grandes cantidades de información, de cara a una futura evaluación de impacto. La realidad no consiste en relaciones lineales unívocas de causas y efectos y no responde a una lógica piramidal

(Gasper, 1997). Los problemas reales aparecen como nudos o marañas en los que una multitud de factores sociales, económicos, culturales, políticos, institucionales... [según Crittenden y Lea (1991), citado por Wiggings y Shields (1995), esta metodología está más orientada hacia los aspectos físicos –infraestructuras- y cuantitativos de los proyectos], tanto internos como externos, controlables e incontrolables, afecta al resto.

Esto tiene un triple efecto: por una parte, ocurrirán impactos no previstos, tanto positivos como negativos; por otra parte, habrá impactos previstos que, a pesar de alcanzar los resultados esperados formulados en el proyecto, no lleguen a producirse (de ahí la importancia de la columna de hipótesis). Por último, a la hora de medir el impacto del proyecto, se planteará el problema de si se podrá medir el efecto neto del proyecto sobre la realidad (atribución) o si el evaluador tendrá que asumir su incapacidad para ello y conformarse con considerar el efecto bruto o contribución al cambio (Ferrero, 2003; Stern et al., 2012).

En todos los casos anteriormente descritos, estaría jugando un papel importante tanto la cadena de resultados como la columna de hipótesis, si se hubiera tenido suficientemente en cuenta; pero como ya se mencionó, esta columna se olvida, una vez diseñado el proyecto (Cracknell y Rednall, 1986; Gasper, 1997; Solem, 1987; Wiggings y Shields, 1995). La simplificación de la realidad en pro de la facilitación del control de gestión hace concebir esta como un túnel normativo, útil para el diseño y la planificación, pero poco ventajoso durante la ejecución o de cara a una evaluación de impacto.

### ***La gestión del ciclo del proyecto***

Finalmente, la gestión del ciclo del proyecto (Biggs y Smith, 2003) es la forma de gestionar la cooperación al desarrollo, en la que el mismo enfoque de marco lógico se utiliza para las diferentes fases de la gestión, desde la planificación hasta la evaluación, pasando por la ejecución y el seguimiento. El trasfondo de esta forma de gestionar es incorporar los cambios del entorno y la corrección de errores y, en menor medida, las lecciones aprendidas, en las distintas etapas del ciclo, de cara a futuras intervenciones, lo cual vendría a solventar el problema anteriormente mencionado sobre la falta de dimensión temporal en el EML (Gasper, 2000). Por eso es habitual encontrar representaciones gráficas del ciclo del proyecto de manera circular, como su propio nombre indica.

**Figura 7. El ciclo del proyecto**



Fuente: EuropeAid (2001)

No obstante, a lo largo de los últimos años se han venido proponiendo otras metodologías de planificación, principalmente en el mundo anglosajón, que atienden mejor el enfoque de proceso y aprendizaje en las intervenciones de cooperación al desarrollo (Ferrero, 2008), como el mapeo de resultados (Earl et al., 2001) o el modelo de cambio más significativo (Davies et al., 2005). Sin embargo, estas nuevas metodologías de planificación han tenido hasta el momento poco predicamento, no solo entre los financiadores, sino entre los propios agentes ejecutores de acciones de ayuda al desarrollo.

La realidad es que muy escasas veces la gestión del ciclo del proyecto es realizada con fines de aprendizaje y, en raras ocasiones, un proyecto planificado según la metodología de marco lógico sufre alguna modificación o corrección profunda. La razón la encontramos en que los financiadores de los proyectos no facilitan esta práctica, pues quieren mantener bajo control aquello para lo que se concedieron los fondos, sin que se pueda dar lugar a desviaciones importantes de lo ejecutado con respecto a lo planificado. El marco lógico se convierte así en un corsé (*lock-frame*) (Ferrero, 2003) para la compleja realidad social y económica que rodea cualquier intervención de desarrollo, cuando lo que se necesita es huir del enfoque de *blueprint* o diseño del proyecto, y dotar de flexibilidad a la intervención, para adaptarla a la realidad.

Asociado a esa rigidez del marco lógico, presionadas las organizaciones ejecutoras de proyectos por las instituciones financiadoras, aparece el problema de encontrar indicadores adecuados de logro de los resultados y objetivos mal o ambiguamente enunciados, poco ambiciosos y poco cerrados cuantitativamente, lo que supone un obstáculo más a la hora de realizar una evaluación de impacto.

Por último, se debe mencionar que, en el ciclo del proyecto, el eslabón de la evaluación (y, por ende, la evaluación de impacto) aparece en el último lugar. Esto significa que las organizaciones no incorporan desde el inicio de la planificación de una intervención la lógica de la evaluación de impacto, por lo cual difícilmente se dan las condiciones óptimas para que se puedan acometer, transcurridos unos años desde el cierre de un proyecto o programa. Como ejemplo de ello, se puede mencionar el hecho de que en el marco lógico no se exige la formulación de un indicador para el objetivo general, por lo que difícilmente se podrá dar seguimiento a una variable que no ha sido pensada desde el principio.

En la bibliografía consultada, se han podido identificar dos grupos de autores claramente diferenciados: aquellos a favor o levemente críticos del uso de esta metodología de planificación y gestión, y otros claramente en contra y a favor de una planificación de los proyectos basada en el proceso y el aprendizaje (figura 8).

Resulta interesante observar la evolución temporal de los posicionamientos teóricos desde los años 70 hasta la actualidad, que coinciden con la evolución de paradigmas de pensamiento que se presentaron en la figura 3 (Vaca, 2013). Mientras a lo largo de la segunda mitad del siglo XX había una clara mayoría de autores que se posicionaba a favor del uso de la herramienta del marco lógico para la formulación de las intervenciones de desarrollo, las críticas han ido siendo cada vez más frecuentes desde los años noventa y en los primeros años del XXI. Se podría concluir que la experiencia acumulada del uso de este instrumento en el ámbito del desarrollo ha permitido identificar con mayor claridad sus debilidades y limitaciones. Así mismo, la mayor madurez en los teóricos sobre cooperación al desarrollo y gestión del desarrollo, y del sector de la cooperación en general, han venido a apoyar también, de manera evidente, la idea de que es necesario incorporar nuevas visiones en la gestión de los procesos de desarrollo.

**Figura 8. Autores defensores y detractores del uso del EML**

Autores y publicaciones a favor o levemente críticos del EML	Año	Autores y publicaciones críticos del EML, a favor del proceso
	1967	Hirschman
Baum	1970	
Biggs	1981	Slade
Casley and Lury	1982	
Guttinger	1982	
Rondinelli	1983	
Cracknell	1986	Clay and Schaffer
Casley and Kumar	1987	Chambers
Solem	1987	
Coleman	1987	
Daniels	1987	
NORAD	1990	Coleman
Eggers	1992	Coleman
Crittenden y Lea	1992	
Eggers	1993	
Cameron	1993	
ISNAR	1993	
Eggers	1994	Maddock
Analoui	1994	Bell
MacArthur	1994	
Chambers	1995	Eyben and Ladbury
Wiggings and Shields	1995	Davis
Sartorius	1996	
Cracknell	1996	
Toffolon-Weis et al.	1996	
Farrington et al.	1997	Gasper
Cummings	1997	Platt
Coles et al.	1998	Mosse, Farrington and Rew
Horton	1998	Horton
IDS	1998	Guijt and Shah
	1998	Thin
	1998	Wood
Eggers	2000	
Cracknell	2000	
Gasper	2000	Gasper
Bell	2000	Cooke y Kothari
HARP	2000	Horton et al.
Horton, Mackay, Anderson & Dupleix	2000	
Alex and Byerlee	2001	Amis
HARP	2001	Hummelbrunner
	2002	Den Hayer
	2003	Ferrero
	2005	Bakewell y Garbutt
	2005	Davis
	2007	Chambers
	2008	Mayne
	2009	Woolcock
	2010	Hummelbrunner
	2010	Fujita
	2011	Collier
	2011	Fors et al.
	2011	Gauck
	2011	Holma et al.
	2011	Piccioto
	2012	Stern
	2012	Delahais et al.
	2012	Wimbush et al
White	2013	

Fuente: Elaboración propia, a partir de Biggs (2003), Ferrero<sup>7</sup> (2003) y Stern et al. (2012)

<sup>7</sup> Ferrero resulta de especial relevancia en este estudio, pues su tesis doctoral fue dirigida por el Dr. José Juan Romero, en ETEA. Por tanto, forma parte de una de las líneas de investigación principales de esta institución universitaria desde hace 25 años: el desarrollo y la cooperación.

Si bien el marco lógico se adoptó como un instrumento para facilitar y simplificar la formulación de los proyectos, el paso del tiempo ha revelado que, por las causas que se explican más adelante, no es la herramienta más adecuada para planificar procesos de desarrollo humano de largo plazo. En muchos casos, la herramienta se ha convertido en el fin y ha dejado de ser el medio.

Para finalizar, y sintetizando el contenido del marco teórico expuesto, según Ferrero (2003), las críticas al enfoque de marco lógico que diversos autores han hecho a lo largo de las últimas décadas se pueden clasificar según las siguientes categorías:

- a. Asociadas a la naturaleza del modelo lógico.
- b. Asociadas a la definición de los elementos del marco lógico (contenidos en la matriz del proyecto).
- c. Asociadas al método (EML-ZOPP).
- d. Asociadas a los efectos inducidos por el uso del EML sobre el carácter de las intervenciones. En algunos casos, el problema está en el erróneo uso que se hace de la herramienta; pero, en otros casos, el problema es la propia herramienta, es decir, su naturaleza, que no está diseñada para planificar intervenciones complejas de desarrollo.

### ***3.1.3. Límites y dificultades asociados al marco lógico para evaluar impacto***

Teniendo como referencia la bibliografía académica sobre el marco lógico y la evaluación de impacto, Vázquez, Torres y Caldentey (2015a) identifican los principales problemas y dificultades que la naturaleza y el uso del marco lógico plantean a la hora de llevar a cabo evaluaciones de las intervenciones de la cooperación al desarrollo, cuando lo que se busca no es solo la rendición de cuentas, sino la identificación de los mecanismos de éxito o fracaso de las acciones para favorecer el aprendizaje. Sus conclusiones derivan de un meta-análisis (Waddington et al., 2012) de cincuenta evaluaciones realizadas por la Fundación ETEA para el Desarrollo y la Cooperación<sup>8</sup> a proyectos e intervenciones de desarrollo llevadas a cabo entre

---

<sup>8</sup> La Fundación ETEA para el Desarrollo y la Cooperación es el Instituto de Desarrollo de la Universidad Loyola Andalucía. [www.fundacionetea.org](http://www.fundacionetea.org). La doctoranda desarrolla su labor profesional en esta institución desde el año

2000 y 2009. El objetivo de este meta-análisis fue comprobar la evaluabilidad de los efectos a medio o largo plazo (evaluación de impacto) de las intervenciones, con el objetivo de aprender más que de rendir cuentas al financiador. En los siguientes párrafos se enumeran los problemas y dificultades encontrados.

- Límites y dificultades asociados a la propia naturaleza del marco lógico:

*A) Problemas en la identificación de los componentes principales de la lógica de intervención: los objetivos*

En numerosas ocasiones, en las intervenciones de cooperación al desarrollo se pretende un objetivo general o meta muy genérica o elevada, existiendo una distancia insalvable entre este y el objetivo específico. El objetivo general se convierte así en un elemento retórico, un “adorno” del proyecto, al que se vacía completamente de importancia. Esto supone ya en sí un problema básico a la hora de plantearse realizar una evaluación de impacto, pues el objetivo general se formula habitualmente en términos de “contribución” y no de “atribución”. Un problema adicional en los diseños de proyectos suele ser la formulación del objetivo general como objetivo específico.

Existe todo un debate en torno a la medición de los impactos macro de un proyecto o programa de desarrollo, cuando las intervenciones están formuladas en su mayoría desde un punto de vista micro, tanto de ONGD como de agencias o bancos de desarrollo. Un objetivo específico bien formulado resulta un mejor elemento para una evaluación de impacto que un objetivo general vago e inconcreto.

*B) Enfoque reduccionista y simplificado de la realidad, sobre-agregación de conceptos en los niveles de resultados esperados y objetivo específico y falta de niveles en la matriz*

El marco lógico tiene su origen en la planificación por objetivos para realidades concretas, controlables y poco cambiantes. Como se comentó anteriormente, es muy adecuado para planificar proyectos de infraestructuras. Sin embargo, por su aparente simplicidad, se ha adoptado la herramienta de manera general para la planificación de intervenciones de carácter socioeconómico, donde las relaciones causales son complejas, enredadas, recursivas,

---

2002 y ha llevado a cabo múltiples evaluaciones de intervenciones de desarrollo, lo que le ha permitido profundizar sus conocimientos en esta materia, desde la teoría y desde la práctica.

reticulares, etc. La propia naturaleza de la herramienta resulta forzada. Se necesitarían no sólo más niveles en la matriz, sino nuevos diseños, no matriciales, tanto de la matriz como de los árboles de problemas y objetivos: de redes, circulares, bi o multidireccionales, etc. En definitiva, un nuevo diseño de herramienta.

Los proyectos de desarrollo a menudo presentan objetivos y resultados que se confunden (y por tanto sus indicadores) o faltan niveles en la matriz de planificación para reflejar correctamente las relaciones causa-efecto, puesto que es difícil reflejar la realidad con toda su complejidad en un escueto esquema lineal y unidireccional. Tratar de realizar una evaluación de impacto sobre un proyecto en el que no está claro cuáles son los objetivos o los resultados, o cuál es la relación entre estos y aquellos, supondrá un problema, por ejemplo, a la hora de intentar establecer relaciones causales aplicando métodos cuantitativos poco flexibles. Por otra parte, podría resultar que el impacto alcanzado por un proyecto no estuviera ni tan siquiera en el diseño del proyecto o que fuera muy distinto al previsto. Probar el impacto de una intervención sobre la variable elegida en la formulación resultaría imposible en este caso, con metodologías poco flexibles como el marco lógico.

### *C) Ausencia de la dimensión temporal en la matriz de marco lógico*

Uno de los principales problemas que presenta el marco lógico es la ausencia de dimensión temporal en la matriz de planificación; es decir, es una herramienta estática de planificación. Este enfoque ha tratado de ser solventado con la incorporación del ciclo del proyecto, pero no se ha logrado, por los motivos que a continuación se exponen.

La matriz de marco lógico no refleja los procesos de cambio que se van sucediendo en el transcurso de una intervención (está por tanto relacionado también directamente con el problema del reduccionismo de la realidad); tampoco aquello que sigue ocurriendo una vez que el proyecto o programa se ha cerrado y que es importante detectar cuando se realiza una evaluación de impacto.

Cuando las intervenciones planificadas en los proyectos son concretas y específicas (por ejemplo, la construcción de casas), el marco lógico resulta más adecuado (como dijimos antes, es apto para construcción de infraestructuras) y robusto para la evaluación que cuando se acometen procesos con características cualitativas, intangibles y “sutiles”, como el empoderamiento de mujeres, el fortalecimiento de organizaciones o algún otro aspecto

relativo a las capacidades humanas (Nussbaum, 2000, citada por Gough, 2007). En estos casos, los informes de evaluación enfocados sólo desde la perspectiva del marco lógico resultan pobres y limitados, y no logran captar cambios en actitudes, preferencias, valoraciones, etc.

En cuanto a la posibilidad que ofrece el enfoque del ciclo del proyecto como forma de incorporar procesos y aprendizajes a las intervenciones de desarrollo, el principal problema se encuentra en que, aun a pesar de los cambios detectados durante la ejecución, en la mayor parte de los casos no se reformulan las matrices de planificación para incorporar las modificaciones y adaptar el diseño a las nuevas realidades. En este caso, el problema está asociado no solo a la naturaleza estática de la herramienta, sino también al uso de la misma. Las entidades, aun a pesar de contar con la posibilidad de cambiar las matrices a través de las llamadas “reformulaciones” que se presentan al financiador, se ven constreñidas por los procedimientos de estos y, por tanto, dificultadas para incorporar a su gestión modificaciones derivadas de los aprendizajes, por temor a que, en una futura revisión de la intervención, el financiador considere no adecuado el cambio y solicite la devolución de fondos.

#### *D) Falta de atención a la columna de hipótesis*

Esta es una de las grandes limitaciones del marco lógico. En un proyecto o programa de cooperación, donde los aspectos sociales y humanos cobran el protagonismo, las hipótesis serían incontables y la mayoría de ellas imprevisibles e incontrolables. Sería prácticamente imposible recoger toda la casuística de posibles derivaciones que podría tener una intervención de desarrollo en la columna de hipótesis. Y precisamente en esta columna estriba buena parte del interés de los futuros impactos de las intervenciones, puesto que son determinantes en el cumplimiento o no de una cadena de resultados y es necesario controlar estas variables, si se quiere aplicar una metodología de evaluación de impacto cuantitativa. Sin embargo, más habitualmente de lo deseable, la columna de hipótesis no refleja una reflexión profunda sobre la complejidad que rodea las intervenciones y los condicionantes para el éxito de estas.

- Límites y dificultades asociados al uso del marco lógico para una futura evaluación de impacto:

*A) Dificultad de comprensión de los conceptos de los componentes, objetivos y resultados, y de la formulación de sus indicadores*

La confusión de conceptos ha venido siendo habitual en los proyectos de ayuda al desarrollo, en especial el objetivo general y el objetivo específico y objetivos específicos con resultados, pues en la realidad estos niveles de la matriz actúan de forma recursiva, influyendo unos en otros.

Este aspecto, íntimamente relacionado con el que ya se ha expuesto sobre el reduccionismo de la realidad y los fallos en la lógica vertical, tiene una relación directa con la errónea formulación de los indicadores para todos los niveles. Si los indicadores no están bien definidos para objetivos y resultados, no será posible medir al final de la intervención los alcances realizados, por lo que no será posible determinar el impacto.

Por otra parte, si no existe una variable indicadora del logro del objetivo general formulada desde el inicio, no será posible medir su avance, una vez ejecutado el proyecto. En los últimos años se ha tratado de resolver este problema con los estudios de la línea de base; pero también en este caso se encontraron problemas derivados del hecho de que estos estudios se basan en información secundaria poco aterrizada en las zonas de intervención, por lo que el impacto tampoco podrá ser evaluado de forma exhaustiva, sino aproximada.

El problema de los indicadores está también muy relacionado con la falta de pericia de los formuladores a la hora de diseñarlos y con la falta de interés real de los financiadores por controlarlos. En muchas ocasiones, por tanto, no es posible medir el avance de los indicadores de los objetivos específicos y la eficacia se mide en términos de resultados obtenidos.

#### *B) La eficiencia por encima de la eficacia y el impacto*

Cuando la evaluación se centra en el logro de los resultados, en la ejecución de las actividades y en el control de los recursos utilizados, se pierde el interés por la medición de la consecución de los objetivos, que además casi siempre resulta difícil, por la errónea formulación de estos y sus indicadores. Se entra así en un bucle no virtuoso en el que los formuladores empiezan el diseño de las intervenciones por las actividades, en lugar de por los objetivos y es en ellas donde depositan sus mayores esfuerzos; mientras que los financiadores se centran en controlar los recursos empleados. Se deja de atender el proyecto desde la lógica del impacto. Y, si en la planificación no se cuida la formulación de los niveles superiores de la matriz, difícilmente se podrá hacer una correcta evaluación de impacto.

En definitiva, las limitaciones del marco lógico, unas por el mal uso y otras por la propia naturaleza de la metodología, analizadas a través del estudio de la bibliografía académica sobre el marco lógico y la evaluación de impacto, y refrendadas por la evidencia empírica (Vázquez et al., 2015b), plantean dificultades para la realización de evaluaciones de impacto basadas solo en métodos cuantitativos, donde las variables deben estar claramente controladas, definidas, ser mensurables y susceptibles de ser relacionadas matemáticamente entre sí.

### **3.2. La evaluación de impacto en intervenciones de desarrollo**

Existe un importante debate en torno a la evaluación de impacto de la ayuda al desarrollo, pues suscita desacuerdos no solo en su propia definición, sino en los métodos y estándares para llevarla a cabo y en la utilidad que se le puede atribuir. Incluso algunos autores plantean dudas sobre su sentido ético, dependiendo de la metodología que se utilice (Alonso, 2012a). Aun así, existe un consenso generalizado alrededor de la necesidad de evaluar el impacto de la ayuda ya prestada.

Existen varias definiciones de impacto y de evaluación de impacto, desarrolladas por financiadores de la ayuda al desarrollo, *think tanks*, institutos de investigación y académicos (Alonso, 2012a; ETEA, 2001, 2003; Evaluation Gap Working Group, 2006; Gertler et al., 2011; Guijt, 2013; Kusters, Van Vugt, Wigboldus, Williams y Woodhill, 2011; Larrú, 2014; NORAD, 1990; Stern et al, 2012). Y cada una de ellas implica una visión y un enfoque metodológico, muchas veces no coincidentes, aun a pesar de estar definiendo un mismo concepto.

La más extendida, por la institución que la formula, es la realizada por el Comité de Ayuda al Desarrollo (CAD) de la Organización para la Cooperación y el Desarrollo Económico (OCDE), aunque existen otras definiciones que van ocupando un lugar importante en la literatura relacionada con el tema, como la aportada por *3ie*<sup>9</sup>.

Entre otras definiciones de evaluación de impacto se encuentran las siguientes:

---

<sup>9</sup> *3ie*, *International Initiative for Impact Evaluation*, está financiado por tres patrocinadores principales: Bill & Melinda Gates Foundation, UKAid, a través del Department for International Development, y William and Flora Hewlett Foundation. Su misión es la financiación de evaluaciones de impacto y revisiones sistemáticas y la generación y difusión de evidencias sobre qué funciona y qué no en programas de desarrollo y por qué. Su especial significación es su nacimiento a partir del informe del Evaluation Gap Working Group que elaboró el informe "When will we ever learn", del que se habló al inicio del capítulo.

- Según el CAD (1995): determinación de los efectos buscados y no buscados, positivos y negativos, a corto y a largo plazo, no sólo sobre la población objetivo, sino sobre la sociedad o comunidad en su conjunto.
- Según Gertler et al. (2011): aquella que pretende establecer de forma nítida las relaciones de causalidad asociadas a la intervención, de modo que pueda determinarse hasta qué punto los cambios en las condiciones de vida de la población pueden ser atribuidos a los efectos del proyecto, programa o política en cuestión.
- Según Alonso (2012a): comprobación de la relación de causalidad entre la intervención y el impacto, una vez que se dejan de lado las consecuencias provocadas por otras acciones, ya sean intrínsecas a la población analizada o provocadas por una política ajena a la intervención que se evalúa.
- Según 3ie: impacto es el resultado de la siguiente relación:  $I = Y_1 - Y_0$ , siendo  $Y_0$  la situación alcanzada por los indicadores de desarrollo cuando no se ha producido ninguna intervención e  $Y_1$  la situación alcanzada por los indicadores de desarrollo cuando sí se ha realizado una intervención.

Por tanto, el principal objetivo de una evaluación de impacto es determinar los efectos de una intervención y la relación de causalidad entre una intervención y sus impactos, es decir, la correcta atribución de efectos. Compartiendo la visión de Stern et al. (2012) y el DFID (Department for International Development, cooperación oficial británica), la evaluación de impacto no debe quedar ahí. También es importante inferir y entender las causas particulares de los cambios producidos (Guijt, Brouwers, Kusters, Prins y Zeynalova, 2011).

Como se puede observar, cada definición implica una visión (paradigma) de la evaluación de impacto, unas más holísticas e incluyentes de todos los efectos posibles, frecuentemente encaminadas a la comprensión de los procesos, a las relaciones causa-efecto y al aprendizaje (como las de Alonso o Gertler), y otras, que focalizan su atención en el método cuantitativo y en la cuantificación del impacto, orientadas normalmente a la medición de lo que funciona y lo que no (como la de 3ie).

Los motivos por los que en los últimos años ha aumentado el interés por este tipo de estudios son, entre otros: la crisis de la cooperación, el cuestionamiento de la eficacia de la ayuda, la

crisis financiera y los cambios en el escenario de la cooperación al desarrollo (actores y factores económicos). Pero tal vez los motivos más importantes son el afán por el aprendizaje y la rendición de cuentas, no siempre compatibles (Regeer, de Wildt-Liesveld, Van Mierlo y Bunders, 2016), no sólo de y hacia los financiadores, sino también a la parte interesada más importante en cualquier acción de desarrollo: los destinatarios del bien o servicio. Según Guijt (2013), se puede considerar que la evaluación de impacto es útil en los siguientes casos:

- Cuando se esté probando una innovación y su posible “multiplicación”.
- Cuando se esté invirtiendo en cooperación en contextos complejos y arriesgados.
- Para rendir cuentas a financiadores, de cara a una posible continuidad o aumento de la ayuda.
- Para rendir cuentas “hacia abajo”, a beneficiarios finales, contrapartes, etc.
- Como parte del proceso de aprendizaje en las intervenciones de desarrollo, tanto para actores del Norte como del Sur.
- Cuando se trate de intervenciones amplias o escalables.

En nuestra opinión, el motivo principal para llevar a cabo una evaluación de impacto es contribuir a la definición, construcción, confirmación o refutación de la *teoría de cambio* de una intervención, con el fin de mejorar los niveles de desarrollo en una población a través de unos mecanismos concretos. La evaluación de impacto debe conducir necesariamente hacia alguna toma de decisión, bien hacia continuar con un tipo de intervención para el desarrollo, bien para descartar una repetición de esta, si se comprueba que el impacto obtenido no es el deseado.

Uno de los debates principales entre teóricos y practicantes de la evaluación desde diferentes paradigmas gira en torno al rigor de los resultados obtenidos, exactitud frente a utilidad, objetividad frente a relevancia, reivindicados con fuerza por quienes abogan por los métodos estadísticos y experimentales. Sin embargo, quienes se posicionan en el paradigma constructivista prefieren un profundo conocimiento de la realidad y el contexto y un buen análisis causal, realizado con rigor en los procedimientos, aun no basado en métodos cuantitativos exhaustivos, ya que estos pueden conducir igualmente a la identificación de las relaciones causales y principalmente a la comprensión de los procesos y al aprendizaje de los mecanismos que las provocan.

AEVAL (2010), Alonso (2012a), García (2008) y Stern et al. (2012) muestran las diferentes metodologías de evaluación de impacto tradicionalmente utilizadas en ciencias sociales, que se pueden clasificar en:

1. Evaluaciones basadas en métodos estadísticos (*multiple causation*, según Stern et al., 2012): a partir de una toma de datos suficientemente representativa de la población, se determinan las relaciones causales entre inputs (intervención) e impactos, a través del cálculo de correlaciones, análisis de panel, análisis de tendencias, análisis de series temporales, tablas de contingencia, modelos lineales o log-lineales, etc.
2. Comparación entre grupos (*counterfactual frameworks*, según Stern et al., 2012): con presencia de grupos de control, aleatorios (método experimental) o no aleatorios (métodos cuasi experimentales o naturales).
  - La evaluación con grupo de control aleatorio (*randomized control trial - RCT*) implica seleccionar individuos estadísticamente iguales (con iguales características) y asignarlos aleatoriamente a un grupo de tratamiento, sobre el que se interviene con la acción de desarrollo, y a un grupo de control, sobre el que no se interviene. Según la teoría (García, 2008; White, 2013), al finalizar la intervención, la diferencia observada en la variable de impacto elegida (por simple diferencia de medias), comparando grupo de tratamiento y grupo de control, se deberá al efecto de la intervención de desarrollo sobre los intervenidos. Esta metodología recibe numerosas críticas, principalmente por cuestionamientos de tipo ético, ya que, por su propio diseño, suponen dejar fuera de la intervención a un grupo de personas, para que actúen como control (cuando se trata de intervenciones para aliviar la pobreza, dejar fuera a un grupo de personas puede suponer problemas éticos).

La metodología de RCT presenta también algunos problemas relacionados con la validez interna de los resultados y validez externa del experimento. Entre los problemas de validez interna se encuentran: problemas con la aleatorización de los grupos; abandono de participantes (*attrition*); la contaminación entre participantes y no participantes; y efectos experimentales, cuando los individuos se comportan de manera diferente a la habitual, al saber que son parte de un

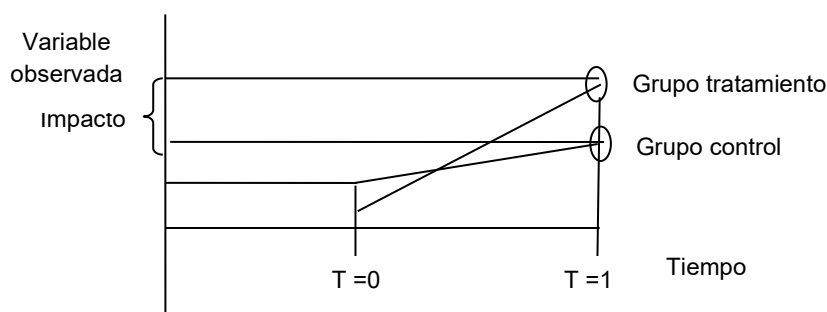
experimento. En todos ellos la comparación no sería válida. Entre los problemas de validez externa se encuentran: utilización de una muestra no representativa del universo poblacional; que la intervención estudiada no represente el programa evaluado; y efectos de equilibrio general relativos al contexto, cuando una intervención funciona en un contexto, pero no en otro.

- Dentro de los diseños cuasi experimentales, se identifican así mismo las siguientes metodologías (Alonso, 2012a; Gertler et al., 2011):
  - a. Diferencias en diferencias (*differences in differences*): se trata de analizar las diferencias que se producen a lo largo del tiempo (normalmente al inicio y al fin de la intervención) entre un grupo de tratamiento y un grupo de control, para eliminar otros efectos que, de forma natural o estructural, se producirían en los grupos, tanto si se interviniera como si no se interviniera con una acción de desarrollo.
  - b. Correspondencia o emparejamiento (*matching*): cuando no ha sido posible la asignación aleatoria al grupo de tratamiento y grupo de control, se trata de crear un grupo de control artificial, identificando aquellos individuos con características (demográficas, socioeconómicas, etc.) lo más parecidas posible a las del grupo con el que se trabajará en la intervención de desarrollo.
  - c. Regresión discontinua o discontinuidad en la regresión (*discontinuity in regression design*): cuando en las intervenciones se fija un índice para acceder a ellas (por ejemplo, un nivel de renta máximo por encima del cual no se puede participar), el grupo de control se puede conformar con aquellos individuos que estén muy cercanos al punto de corte, pues se supone que tienen características que, si bien no son exactamente iguales, casi lo son.
  - d. Métodos de selección aleatoria (*randomized selection methods*): en ocasiones, cuando el acceso a un programa es voluntario, puede haber causas diferenciales de partida entre los individuos que se postularon y los que no, por ejemplo, más espíritu emprendedor, menor aversión al riesgo, mayor autoconfianza, etc. En estos casos, la forma de construir el grupo de

control no sería entre quienes se postularon y los que no lo hicieron, precisamente porque no serían estadísticamente iguales, sino entre quienes se ofrecieron voluntariamente, asignar aleatoriamente (mediante sorteo) a las personas al grupo de tratamiento y al grupo de control.

Utilizando una representación gráfica de la evaluación de impacto mediante comparación entre grupos, el significado de impacto sería el siguiente:

**Figura 9. Representación gráfica de impacto según el método experimental**



Fuente: Guijt (2013)

Se puede decir que no existe un patrón ideal en las metodologías para la realización de evaluaciones de impacto. La elección de una u otra metodología dependerá de cada situación concreta que se pretenda evaluar. Y no en menor medida, del propósito que se persiga con la realización de la evaluación de impacto.

Los métodos experimentales no son necesariamente mejores en cualquier caso que aquellos en los que se puede aplicar la observación de las relaciones causa-efecto, aunque esta carezca de representatividad numérica. El “experimentalismo”, entendido como el esfuerzo por la determinación de una muestra representativa en la que se puedan controlar los efectos de una intervención de desarrollo, no tiene por qué ser siempre la mejor opción.

El origen de estos métodos se asocia principalmente con la investigación médica, donde un fármaco se prueba en una muestra de población (grupo de tratamiento), mientras se determina otra muestra de similares características en la que no se administra (grupo de control), para observar los efectos del mismo en el grupo tratado (Banerjee y Duflo, 2009).

Pero como ya otros autores han criticado (ver los enumerados cuando se presentó el paradigma constructivista), en una acción de desarrollo no resulta sencillo utilizar este tipo de metodología con garantías y solo es aplicable a nivel micro (García, 2008). Entre otros factores, porque encontrar un grupo humano exactamente igual a otro es, si no imposible, ciertamente muy complicado.

Pese a las críticas recibidas por la evaluación experimental, Banerjee y Duflo (2009) argumentan que siempre será mejor, en términos de validez, el método experimental que los datos obtenidos por observación; y que, de hecho, este no es un debate nuevo, en alusión al trabajo de Heckman (1992) "Randomization and social policy evaluation" (citado por Banerjee et al., 2009).

Los problemas que Banerjee y Duflo (2009) identifican en la evaluación experimental son:

- Aspectos relacionados con la dependencia del entorno: los experimentos nunca podrán ser iguales, puesto que el entorno siempre variará, aunque sea levemente.
- Aspectos relacionados con el exceso de confianza en el cumplimiento de los resultados previstos, por sesgo de selección de la muestra: cuando a las personas se les ofrece participar en el proyecto y aceptan entusiasmadas esa participación. O cuando las entidades que implementan los programas, los ponen en marcha allí donde piensan que pueden tener un mayor nivel de participación o de éxito.
- Aspectos relacionados con la aleatoriedad en la selección de los participantes: habitualmente se plantea un problema ético cuando se pone en marcha un proyecto del que únicamente se benefician algunos individuos o familias, y se deja conscientemente a otro grupo, de similares características, fuera del proyecto, solo para fungir como grupo de control.
- Efectos de equilibrio o de escalamiento del experimento: los efectos producidos en un proyecto a nivel micro pueden no ser los mismos (y probablemente no lo sean) cuando la intervención se escale a nivel nacional (Vivalt, 2014).
- Heterogeneidad de efectos en el grupo de tratamiento: puesto que las conclusiones de un experimento hacen referencia a los resultados medios obtenidos en el grupo de

tratamiento, claramente se estarán dejando de lado aquellos casos que tengan una mayor covarianza con respecto a la media.

Sin embargo, también para todos estos aspectos encuentran contra-argumentos por los que, aun así, recomiendan el método experimental. Se trata de hacer asunciones sobre la simplificación de los resultados, pero en su opinión, consideran que para el aprendizaje es más favorable diseñar las evaluaciones de impacto desde esa perspectiva.

La cuestión clave en la determinación de la metodología más adecuada para la evaluación es poder identificar correctamente los efectos o impactos de un proyecto o programa (*program attribution*) que, en algunos casos, tan sólo podrán definirse como “contribuciones” (*contribution*), por la dificultad de aislar los efectos de aquel de otras actuaciones; o incluso de la evolución propia e inherente a la población con la que se trabaja. No se debe olvidar que cualquier colectivo humano evoluciona a lo largo del tiempo y habrá cambios inherentes a su propia naturaleza, no debidos necesariamente a las acciones externas.

La realidad es cada vez más compleja y los estudios sobre complejidad están siendo también una referencia para teóricos y practicantes de la evaluación desde el paradigma constructivista (Gerrits y Verweij, 2015; Jones, 2011): las relaciones de causalidad son múltiples y no lineales, los sistemas de causalidad están abiertos a influencias externas y las causas interactúan unas con otras. El contexto lo cambia todo; por tanto, se necesita una teoría para entender cómo, cuándo y dónde las causas marcan una diferencia en los efectos. Sin embargo, las evaluaciones de impacto se focalizan en una parte muy pequeña de la realidad, aquella que es más fácil de medir y, en consecuencia, donde la posibilidad de obtener información rigurosa y válida (al menos internamente) es más factible. Sin embargo, con esta forma de evaluar, como con toda simplificación, se pierden datos, principalmente sobre los efectos intangibles o cualitativos, que son una buena parte de los efectos de la ayuda al desarrollo de no poca importancia, porque afectan a seres humanos. Estos efectos, que resulta fundamental descubrir y valorar para aprender, se producen tras un proceso de cambio que normalmente responde a una estrategia, lo que implica un plazo considerable de tiempo (de ahí la importancia de *cuándo* evaluar) (Fitz-Gerald y Neal, 2002).

Desde los teóricos y practicantes de la evaluación que defienden métodos de investigación alternativos a los experimentales y cuasi experimentales para la evaluación de impacto, han

surgido otras metodologías alternativas, que buscan el aprendizaje de la experiencia y su posible réplica (Stern et al., 2012):

- Evaluación realista (*realist evaluation framing*): basada en la identificación de los factores y mecanismos causa-efecto a través de estudios de caso, según el contexto, para la formulación de una teoría causal.
- Análisis de contribuciones (*contribution analysis*): basada en la identificación y confirmación de los procesos causa-efecto. Tanto el anterior método como este, están fundamentados en la evaluación basada en la teoría (Carter, 2012).
- Entrevistas y encuestas de actitudes (*people's narratives and surveys for attitudes*).
- Análisis cualitativos comparativos (*qualitative comparative analysis*)
- Evaluación participativa enfocada al aprendizaje (*participatory impact assessment and learning approach*), evaluación rural participativa (*PRA –Participatory Rural Appraisal*, Chambers et al., 2009).

#### **4. RECAPITULACIÓN**

El análisis del marco teórico de la evaluación es complejo y variado, con múltiples modelos y metodologías, según el paradigma desde el que se plantee la investigación. Las limitaciones del marco lógico, unas por el mal uso y otras por la propia naturaleza de la metodología, suponen un problema para la realización de evaluaciones de impacto. Se han detectado varios puntos críticos en torno a la metodología de planificación de la cooperación, si se quieren promover intervenciones que puedan ser adecuadamente evaluadas. Destacan, entre otros: contar con una línea de base centrada específicamente en las variables sobre las que se interviene y que serán evaluadas; revisar la correcta enunciación de objetivo general, objetivos específicos y resultados, para asegurar su coherencia; formular adecuadamente los indicadores; y definir mecanismos de flexibilidad en la formulación y ejecución de las intervenciones, que respondan especialmente a las hipótesis de riesgo, cuya formulación debe ganar relevancia en la práctica, dado que suelen convertirse en factores críticos para explicar los impactos obtenidos.

Se han presentado modelos alternativos de evaluación, basados en metodologías cuantitativas y cualitativas. Sin embargo, en la actualidad muchos practicantes de la evaluación de impacto están incluyendo en las evaluaciones ambos tipos (*mixed methods*) (Stern et al., 2012). De esta forma, no solo se evalúa el efecto neto de las intervenciones, sino que se conocen sus causas y qué cambios de tipo cualitativo o intangible han producido en las personas.

## REFERENCIAS BIBLIOGRÁFICAS

- AEVAL (2010). *Fundamentos de evaluación de políticas públicas*. Madrid: Ministerio de Política Territorial y Administración Pública. Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios.
- Alkin, M. C. (Ed.) (2012). *Evaluation roots: A Wider Perspective of Theorists' Views and Influences* (2nd Ed.). Thousand Oaks, CA: Sage.
- Alonso, J.A. (2012a). "La evaluación en la cooperación internacional para el desarrollo". *Presupuesto y Gasto Público* 68/2012: 239-255. Madrid: Secretaría de Estado de Presupuestos y Gastos, Instituto de Estudios Fiscales.
- Arensmann, B. y Van Wessel, M. (2018). "Negotiating effectiveness in transnational advocacy evaluation". *Evaluation*, Vol. 24(1) 51–68. DOI: 10.1177/1356389017733210
- Bakewell, O. y Garbutt, A. (2005). *The use and abuse of the logical framework approach*. Stockholm: SIDA.
- Bamberger, M., Rugh, J., Mabry, L. (2012). *Real world evaluation. Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage.
- Banerjee, A. y Duflo, E. (2009). "The Experimental Approach to Development Economics", *Annual Review of Economics*, 1:151–78, DOI: 10.1146/annurev.economics.050708.143235.
- Bester, A. (2012). *Results-Based Management in the United Nations Development System: Progress and Challenges*. New York: United Nations Department of Economic and Social Affairs.
- Biggs, S. y Smith, S. (2003). "A paradox of learning in project cycle management and the role of organizational culture". *World Development*, Vol. 31, nº 10, pp. 1743-1757. Elsevier.
- Blattman, C., Jamison, J. y Sheridan, M. (2017). "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia". *American Economic Review*, Nº 107(4): 1165–1206.
- Calidoni-Lundberg, F. (2006). "Evaluation: definitions, methods and models". *Working Paper*, R2006:002. Östersund, Sweden: ITPS, Swedish Institute For Growth Policy Studies.

- Campbell, D. (1991). "Methods for the Experimenting Society". *Evaluation Practice*, Vol. 12, Nº 3, pág. 223-260.
- Carden, F. y Alkin, M. (2012). "Evaluation roots: an international perspective". *Journal of MultiDisciplinary Evaluation*, Volume 8, Nº 17. ISSN 1556-8180.
- Carter, B. (2012). *Theory-based evaluation approach*. Helpdesk Research Report. Londres: Governance and Social Development Resource Centre.
- Chambers, R. (1995). "Poverty and livelihoods: whose reality counts?" *Environment and Urbanization*, Vol. 7, Nº 1, April.
- Chambers, R. y Guijt, I. (1995). "PRA Five years later. Where are we now? Forests, trees and people". *Newsletter* Nº 26/27. En línea en <https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/277/rc174.pdf?se>. Acceso: 21/05/2018.
- Chambers, R., Karlan, D., Ravallion, M. y Rogers, P. (2009). "Diseño de evaluaciones de impacto: Perspectivas diversas". *Working Paper* Nº 4. Nueva Delhi: Howard White, International Initiative for Impact Evaluation (3ie).
- Comisión Europea (2015). *Guide to Cost-Benefit Analysis of Investment Projects*. Bruselas: Dirección General de Política Regional y Urbana de la Unión Europea.
- Cucuzzella, C. (2011). "Why is Fourth Generation Evaluation essential for sustainable design?" *Design Principles and Practices*, Vol. 5 (1): 239-252.
- Davies, R. y Dart, J. (2005). *The "Most Significant Change" technique: A guide to its use*. UK: CARE International y otros.
- Development Assistance Committee (1995). *Principios del CAD para una ayuda eficaz*, Comité de Ayuda al Desarrollo, OCDE. Madrid: Mundiprensa.
- Donaldson, S. y Scriven, M. (2003). *Evaluating Social Programs and Problems. Visions for the New Millennium*. New York: Routledge.
- Earl, S., Carden, F. y Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs*. Ottawa: International Development Research Centre.

- Ecofin (2006). *Financial and economic project analyses: A decision-making tool*. European Commission. En línea: [https://ec.europa.eu/europeaid/sites/devco/files/methodology-financial-and-economic-project-analyses-decision-making-tool\\_en\\_2.pdf](https://ec.europa.eu/europeaid/sites/devco/files/methodology-financial-and-economic-project-analyses-decision-making-tool_en_2.pdf)
- Eisner, E. (2017). *The enlightened eye: qualitative inquiry and the enhancement of educational practice*. New York: Teachers College Press.
- El Dessouky, N. (2016). "Public policy evaluation theory: from first to fifth generation". *International Journal of Economic and Business Review*, Vol. 4, Issue 4 (April). e-ISSN: 2347.
- ETEA (2001). *Desarrollando la cooperación (I). Evaluación, seguimiento y difusión de proyectos de cooperación internacional al desarrollo*. Ayuntamiento de Córdoba. Córdoba.
- ETEA (2003). *Desarrollando la cooperación (II). Evaluación, seguimiento y difusión de proyectos de cooperación internacional al desarrollo*. Ayuntamiento de Córdoba. Córdoba.
- EuropeAid Co-operation Office (1999). *Guidelines for evaluation methodology, criteria and suggested layout for evaluation reports, 02/99, SCR/F/5 D98*.
- EuropeAid Co-operation Office (2001). *Manual de Gestión del Ciclo de Proyecto*. Bruselas: Comisión Europea.
- EuropeAid Co-operation Office, Directorate General External Relations, Directorate General Development (2006). *Evaluation Methods for the European Union's External Assistance. Evaluation Tools*, Vol. 4, European Communities, Luxemburg. En línea: [file:///D:/Users/mjvazquez/Downloads/evaluation-methods-guidance-vol4\\_en.pdf](file:///D:/Users/mjvazquez/Downloads/evaluation-methods-guidance-vol4_en.pdf). Acceso: 1/9/2017.
- Evaluation Gap Working Group (2006). *When will we ever learn? Improving Lives through Impact Evaluation*. Ed. Center for Global Development, Washington.
- Ferrero, G. (2003). *De los proyectos de cooperación a los procesos de desarrollo. Hacia una gestión orientada a resultados*, PhD diss, Valencia: Universidad de Valencia.
- Ferrero, G. (2008). *Apoyando los procesos de desarrollo. Enfoques y métodos para una ayuda inclusiva*. Documento de trabajo número 17. Madrid: Fundación Carolina.

- Ferrero y de Loma-Osorio, G., Zepeda, C.S. (2014). "Rethinking Development Management Methodology: Towards a "Process Freedoms Approach". *Journal of Human Development and Capabilities*, Vol. 15:1, 28-46, DOI: 10.1080/19452829.2013.877425.
- Fetterman, D. y Wandersman, A. (2007). "Empowerment Evaluation: Yesterday, Today, and Tomorrow". *American Journal of Evaluation*, Vol. 28 Nº 2, June 2007, 179-198. DOI: 10.1177/1098214007301350.
- Fitz-Gerald, A. and Neal, D. (2002). *A Strategic Management Framework for Improved Aid Delivery*. Royal Military College of Science, Cranfield University. En línea: <http://www.unssc.org>. Acceso: 18/3/2014.
- Freire, P. (1967). *La educación como práctica de la libertad*. Madrid: Ed. Siglo XXI.
- Fujita, N. (2010). *Beyond logframe; using systems concepts in evaluation. Issues and prospects of evaluation for international development*. Series IV. Tokyo: Foundation for Advanced Studies on International Development.
- García, J. (2008). *El análisis experimental de la ayuda al desarrollo. La evaluación de lo que funciona y lo que no funciona*. Bilbao: Fundación BBVA.
- Gaspar, D. (1997). "Logical frameworks: a critical assessment managerial theory, pluralistic practice". *Working Series Paper 264*. The Hague: Institute of Social Studies.
- Gaspar, D. (1999). *The logical framework approach. Beyond the project enclave: questions and warnings for monitoring and evaluation of humanitarian assistance in complex emergencies*. The Hague: Institute of Social Studies.
- Gaspar, D. (2000a). "Evaluating the 'Logical Framework Approach' towards learning-oriented development evaluation", *Public Administration and Development*, Vol. 20, Issue 1: 17-28. DOI: 10.1002/1099-162X(200002)20:1<17::AID-PAD89>3.0.CO;2-5
- Gaspar, D. (2000b). *"Logical frameworks": problems and potentials*. Institute of Social Studies. The Hague. En línea: [http://www.petersigsgaard.dk/PDFfiler/gasper\\_logical\\_framework\\_problems.pdf](http://www.petersigsgaard.dk/PDFfiler/gasper_logical_framework_problems.pdf). Acceso: 2/2014.

- George, A. (2015). "Explicating the capability approach through the Voices of the Poor: A case study of waste-picking women in Kerala", *Journal of Human Development and Capabilities*, Vol. 16, Nº 1, 33-46, <http://dx.doi.org/10.1080/19452829.2014.938728>.
- Gertler, P., Martínez, S., Premand, P., Rawlings, L., Vermeersch, C. (2011). *Impact evaluation in practice*. Washington Ed. World Bank. En línea: acceso el 23/07/2013 en [http://soterespirces-wprñdbaml-prg/EXTHDOFFICE/Resources/5485726-1295455628620/Impact\\_Evaluation\\_in\\_practice.pdf](http://soterespirces-wprñdbaml-prg/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_practice.pdf).
- Gerrits, L. y Verweij, S. (2015). "Taking stock of complexity in evaluation: A discussion of three recent publications". *Evaluation*, Vol. 21(4), pp. 481–491
- Gough, I. (2007). "El enfoque de las capacidades de M. Nussbaum: un análisis comparado con nuestra teoría de las necesidades humanas". *Papeles de Relaciones Ecosociales y Cambio Global*, Nº 100, CIP-Ecosocial/Icaria, invierno 2007/2008.
- Guba, E. y Lincoln, Y. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage Publishing.
- Gudiño, F. (2016). *Evaluar para aprender. Un análisis de utilización de las evaluaciones de la Cooperación Española para la toma de decisiones*. Madrid: Ministerio de Asuntos Exteriores y de Cooperación. Secretaría de Estado de Cooperación Internacional y para Iberoamérica.
- Guijt, I. (2013). "Impact evaluation for International Development: Why we need it, why it is hard". Key-Speech. Wageningen: Conference Impact Evaluation 2013: Taking stock and looking ahead.
- Haug, P. y Shwandt, T. (Eds.) (2003). *Evaluating educational reforms: Scandinavian perspectives*. pp. 41-71. Greenwich, Connecticut: Information Age Publishing.
- Heckman (1992). "Randomization and social policy evaluation". *NBER Technical Working Paper Series*, nº 107. Cambridge, MA: National Bureau of Economic Research.
- Hummelbrunner (2010). "Beyond logframe: critique, variations and alternatives", en Fujita, N. (2010). *Beyond logframe; using systems concepts in evaluation. Issues and prospects of evaluation for international development*. Series IV. Tokyo: Foundation for Advanced Studies on International Development.

- Ivaldi, S., Scaratti, G. y Nuti, G. (2015). "The practice of evaluation as an evaluation of practices". *Evaluation*, Vol. 21(4) 497–512. DOI: 10.1177/1356389015606538.
- Jara, O. (1994). *Para sistematizar experiencias. Una propuesta teórica y práctica*. Lima: Ediciones Tareas.
- Jara, O. (2001). *Dilemas y desafíos de la sistematización de experiencias*. Costa Rica: Centro de Estudios y Publicaciones-Alforja.
- Jara, O. (2012). "Sistematización de experiencias, investigación y evaluación: aproximaciones desde tres ángulos". *Educación Global Research* Nº 1. En línea: <http://educacionglobalresearch.net/wp-content/uploads/02A-Jara-Castellano.pdf>.
- Jones, H. (2011). "Taking responsibility for complexity. How implementation can achieve results in the face of complex problems". *Working Paper*, Nº 330. London: Overseas Development Institute.
- Jones, H. (2013a). "Complexity 101 Part 1: behind de hype, what do we actually know?" *Blog From Poverty to Power*. En línea: <http://oxfamblogs.org/fp2p/complexity-101-behind-the-hype-what-do-we-actually-know/>. Acceso: 5/11/2013
- Jones, H. (2013b). "Complexity 101 Part 2: Getting to the so what". *Blog From Poverty to Power*. En línea: <http://oxfamblogs.org/fp2p/complexity-101-part-2-getting-to-the-so-whats/>. Acceso: 6/11/2013
- Kogen, L. (2018). "What have we learned here? Questioning accountability in aid policy and practice". *Evaluation*, Vol. 24(1), pp. 98–112.
- Kusters, C., Van Vugt, S., Wigboldus, S., Williams, B., Woodhill, J. (2011). *Making evaluation matter: a practical guide for evaluators*, Wageningen: Centre for Development Innovation, Wageningen University & Research Centre.
- Larrú, J.M. (2000). *La evaluación en los proyectos de cooperación al desarrollo de las ONGD españolas*. Tesis doctoral. Madrid: Universidad San Pablo CEU.
- Larrú, J.M. (2014). "Evaluaciones de impacto en la cooperación para el desarrollo". *Revista de Evaluación de Programas y Políticas Públicas*, Nº. 3. Madrid: Universidad Nacional de Educación a Distancia, pp. 117-153.

- Lay, M. y Papadopoulos, I. (2007). "An exploration of fourth generation evaluation in practice". *Evaluation*, Vol. 13 (4): 486-495. DOI 10.1177/1356389007082135.
- Leeuw, F. y Donaldson, S. (2015). "Theory in evaluation: Reducing confusion and encouraging debate". *Evaluation*, Vol. 21(4), pp. 467-480. DOI: 10.1177/1356389015607712.
- Lund, G. (2011). "Fifth-generation evaluation". *Social constructionism in practice*. En línea: <http://www.haslebo-partnere.dk/en/publications/articles>. Acceso: 4/2018.
- Martinic, S. (1984). *Análisis de algunas categorías de sistematización*. Talagante: CIDE-FLACSO.
- MAE-SECIPI (1998). *Metodología de evaluación de la Cooperación Española*. Madrid: Ministerio de Asuntos Exteriores, Secretaría de Estado para la Cooperación Internacional y para Iberoamérica.
- MAE-SECIPI (2001). *Metodología de evaluación de la Cooperación Española II*, Ministerio de Asuntos Exteriores, Madrid: Secretaría de Estado para la Cooperación Internacional y para Iberoamérica.
- Mertens, D. y Wilson, A. (2012). *Program evaluation theory and practice: a comprehensive guide*. New York: The Guilford Press.
- Molina, H. y Torres, A. (2013). *Gestión de políticas y programas de desarrollo infantil temprano*. Washington: Banco Interamericano de Desarrollo.
- Muñiz, M. y Gasper, D. (2009). "Looking for long-run development effectiveness: an autonomy-centred framework for project evaluation". *Working Paper*, MGSOG/2009/WP005. Maastricht: Maastricht Graduate School of Governance.
- Muñoz, G. y Mata, R. (2017). "The fifth generation of evaluation: evaluating for quality". Hawaii University International Conferences. Science, Technology and Engineering, Arts, Mathematics and Education, June 2017. En línea: en <https://huichawaii.org/wp-content/uploads/2017/08/Mata-Toledo-Ramon-A.-2017-STEAM-HUIC.pdf>. Acceso: 4/2018.
- Nioche, J.P. (1982). "De la evaluación al análisis de las políticas públicas". *Revue Francaise de Science Politique*, Vol. 33, Nº 1, pág. 32-61.

- Norwegian Agency for Development Cooperation (NORAD) (1990). *The logical framework approach – Handbook for objectives-oriented project planning*, Oslo: NORAD.
- Nussbaum, M.C. (2000). *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- Organización para la Cooperación y el Desarrollo Económico (2008). *Declaración de París sobre la eficacia de la ayuda al desarrollo y programa de acción de Accra 2005-2008*. En línea: <http://www.oecd.org/development/effectiveness/34580968.pdf>. Acceso: 30/7/2013.
- Palumbo, D. (2001). “Il Porcesso di valutazione. Decidere, programmare, valutare”, en AEVAL (2010). *Fundamentos de evaluación de políticas públicas*. Madrid: Ministerio de Política Territorial y Administración Pública. Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios.
- Patton, M. (2004). “The roots of utilization-focused evaluation”, en Alkin, M. (2004). *Evaluation roots*. Thousand Oaks, CA: Sage.
- Patton, M.Q. (2008). *Utilization focused evaluation*. Thousand Oaks, CA: Sage.
- Patton, M.Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: The Guilford Press.
- Porrini, L. y Murciano, J. (2015). “Evaluación de políticas públicas. Editorial”. *Studia politicæ*. Nº 34, primavera-verano 2014/2015.
- Ravallion, M. (2001). “The mystery of the vanishing benefits: an introduction to impact evaluation”. *The World Bank Economic Review*, Vol. 15, Nº 1, pp. 115-140.
- Regeer, B., De Wildt-Liesveld, R., Van Mierlo, B. y Bunders, J. (2016). “Exploring ways to reconcile accountability and learning in the evaluation of niche experiments”. *Evaluation*, Vol. 22(1), pp. 6–28
- Rodríguez, C. y Monterde, R. (2014). “Estrategias cualitativas de evaluación en la Cooperación para el Desarrollo: viejos debates y nuevos retos”. *Documento de Trabajo, Serie CECOD*, Número 29 / 2014.
- Rossi, P. H., Freeman, H. E., y Lipsey, M. W. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

- Sartorius, R. (1996). "The third generation logical framework approach: dynamic management for agricultural research projects". *The Journal of Agriculture Education and Extension*, 2 (4).
- Schmitt, J. y Beach, D. (2015). "The contribution of process tracing to theory-based evaluations of complex aid instruments". *Evaluation*, Vol. 21(4), pp. 429–447.
- Selener, D. (1996). *La sistematización de proyectos de desarrollo. Una metodología de evaluación participativa para fortalecer la capacidad institucional de ONGs y organizaciones populares*. Preval. En línea: <http://preval.org/documentos/00513.pdf>.
- Serrat, O. (2009). "The Most Significant Change Technique". Knowledge Solutions. Asian Development Bank. En línea: <https://www.adb.org/sites/default/files/publication/27613/most-significant-change.pdf>. Acceso: 19/05/2018.
- Shadish, W. (1998). "Evaluation theory is who we are". *American Journal of Evaluation*, Vol. 19, Nº 1, pp. 1-19. ISSN: 1098-2140.
- Smidt, A., Balandin, S., Sigafoos J. y Reed, V. (2009). "The Kirkpatrick model: A useful tool for evaluating training outcomes". *Journal of Intellectual & Developmental Disability*, Volume 34, Issue 3, pp. 266-274. DOI: 10.1080/13668250903093125.
- Solem, R.R. (1987). "The logical framework approach to project design, review and evaluation in A.I.D.: Genesis, impact, problems and opportunities". *Working paper*, Nº 99. Washington: USAID.
- Stake, R. (2004). *Standard-based and responsive evaluation*. Thousand Oaks, CA: sage.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B. (2012). "Broadening the range of designs and methods for impact evaluations". *Working paper* Nº 38. London: Department for International Development.
- Stufflebeam, D. y Zhang, G. (2017). *The CIPP Evaluation Model: How to Evaluate for Improvement and Accountability*. New York: The Guilford Press.

- Subirats, J. (2005). "Catorce puntos esenciales sobre evaluación de políticas públicas con especial referencia al caso de las políticas sociales". *Ekonomiaz*, Nº 60, Vol. 1, 3er. cuatrimestre.
- Subirats, J. y Gomà, R. (2002). "Evaluación de políticas públicas". *Planificación y evaluación de políticas de información*. Barcelona: Ed. Fundació per a la Universitat Oberta de Catalunya.
- Tapella, E. y Rodríguez-Bilella, P. (2014a). "Shared learning and participatory evaluation: The sistematización approach to assess development interventions". *Evaluation*, 2014, Vol. 20 (1), pp. 115–133.
- Tapella, E. y Rodríguez-Bilella, P. (2014b). "Sistematización de experiencias: una metodología para evaluar intervenciones de desarrollo". *Revista de Evaluación de Programas y Políticas Públicas*, Nº. 3: 80-116. Madrid: Universidad Nacional de Educación a Distancia.
- United Nations Development Program (UNDP) (2007). *Evaluation of results-based management in UNDP*. New York: United Nations Development Program.
- United Nations Evaluation Group (UNEG) (2007). *The role of evaluation in Results-Based Management*. New York: United Nations Evaluation Group.
- Vázquez, M.J., Caldentey, P., Romero, J.J. (2013). "Hacia una concreción de término y metodología de evaluación de impacto de la ayuda universitaria al desarrollo coherentes con el desarrollo humano". *Full Academic Paper*. Managua: HDCA Conference.
- Vázquez, M.J., Santos, F., Sianes, A., Simón, P. (2015a). *Sistematización, evaluación y rendición de cuentas: reflexiones y aportes para las ONGD andaluzas*, ISBN: 978-84-606-7254-8. Córdoba: Fundación ETEA para el Desarrollo y la Cooperación.
- Vázquez, M.J., Torres, M., Caldentey, P. (2015b). "Límites del marco lógico y deficiencias de la evaluación tradicional de la cooperación al desarrollo para medir impacto". *Revista Iberoamericana de Estudios de Desarrollo*, Vol. 4, Nº 2 (2015), pp. 80-105. ISSN: 2254-2035.
- Vázquez, M.J., Torres, M., Caldentey, P., Nekhay, O. (2017a). "Evaluating impacts of University Cooperation for Development from the Voice of the South". *Revista de Economía Mundial*, Nº 47, 2017, pp. 95-116.

- Vedung, E. (2003). "Models of Evaluation", en Haug, P. y Schwandt, T. (Eds.) (2003). *Evaluating educational reforms: Scandinavian perspectives*. pp. 41-71. Greenwich, Connecticut: Information Age Publishing.
- Vedung, E. (2010). "Four Waves of Evaluation Diffusion". *Evaluation*, 16(3), pp. 263–277. DOI: 10.1177/1356389010372452.
- Vivalt, E. (2014). "How Much Can We Generalize from Impact Evaluations?". *Working Paper*. New York: New York University.
- Waddington, H., White, H., Birte Snilstveit, B., Garcia-Hombrados, J., Vojtkova, M., Davies, P., Bhavsar, A., Evers, J., Perez-Koehlmoos, T., Petticrew, M., Valentine, J.C. y Tugwell, P. (2012). "How to do a good systematic review of effects in international development: a tool kit". *Journal of Development Effectiveness*, 4:3, pp. 359-387, DOI: 10.1080/19439342.2012.711765.
- Walker, M. (2005). "Amartya Sen's capability approach and education". *Educational Action Research*, Vol. 13, Nº 1.
- Westhorp, G. (2014). *Realist impact evaluation. An introduction*. Methods Lab. London: Overseas Development Institute.
- White, H. (2013). "An introduction to the use of randomised control trials to evaluate development interventions". *Journal of Development Effectiveness*, Vol. 5, Nº 1, pp. 30–49, <http://dx.doi.org/10.1080/19439342.2013.764652>.
- Wiggings, S., Shields, D. (1995). "Clarifying the 'logical framework' as a tool for planning and managing development projects". *Project appraisal*, 10 (1): pp. 2-10.
- World Bank (2004). *Instrumentos, métodos y enfoques*. Washington: Departamento de Evaluación de Operaciones, Seguimiento y Evaluación del Banco Internacional de Fomento, Banco Mundial.