



CEU

*Instituto Universitario
de Estudios Europeos*

Universidad San Pablo



Documento de Trabajo

Serie CECOD

Número 15 / 2011

Dos Métodos de Evaluación: Criterios y Teoría del Programa

Juan Andrés Ligeró Lasa

Documento de Trabajo
Serie CECOD
Número 15 / 2011

**Dos métodos de evaluación:
criterios y teoría del programa**

Juan Andrés Ligeró Lasa

El Centro de Estudios de Cooperación al Desarrollo (CECOD) está formado por dos instituciones: la Fundación CODESPA y el Instituto Universitario de Estudios Europeos de la Universidad CEU San Pablo. El CECOD cuenta con el patrocinio de la Agencia Regional para la Inmigración y la Cooperación de la Comunidad de Madrid.

El CECOD es un centro de referencia en materia de cooperación internacional para el desarrollo, que tiene como objetivo fomentar el debate y contribuir al conocimiento e investigación en estos temas. Su actividad se basa en un modelo de colaboración entre diversas esferas de actuación social, como son la comunidad universitaria, la administración pública, las ONGD y el ámbito empresarial.

Serie CECOD de Documentos de Trabajo del Centro de Estudios de Cooperación al Desarrollo

Dos métodos de evaluación: criterios y teoría del programa

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

© 2011, Juan Andrés Ligeró Lasa

© 2011, Fundación Universitaria San Pablo CEU

CEU Ediciones

Julián Romea 18, 28003 Madrid

www.ceuediciones.es

Centro de Estudios de Cooperación al Desarrollo (CECOD)

Julián Romea 22, 28003 Madrid

www.ceciod.org

E-mail autor: jliger@telefonica.net

ISBN: 978-84-15382-08-9

Depósito legal: M-41043-2011

Maquetación: Servicios Gráficos Kenaf s.l.

Índice

1.	INTRODUCCIÓN	5
2.	EVALUACIÓN ORIENTADA POR CRITERIOS	5
	2.1. Elección entre modelos preordenados y modelos ad hoc sensibles a las necesidades de los actores implicados.....	6
	2.2. Identificación de fuentes y actores clave	8
	2.3. Extracción de las necesidades informativas	9
	2.3.1. Fase divergente: recopilación de las necesidades informativas.....	9
	2.3.2. Fase convergente: selección de las necesidades informativas	10
	2.4. Operacionalización.....	10
	2.4.1. Trabajo vertical: definición de los criterios y preguntas.....	11
	2.4.2. Trabajo horizontal: definición de indicadores y fuentes	12
	2.5. Consideraciones sobre la evaluación por criterios	13
	2.6. Críticas a la evaluación orientada por criterios y estándares	15
3.	EVALUACIÓN DE LA TEORÍA DEL PROGRAMA	19
	3.1. Construcción, reconstrucción o identificación de la teoría del programa o modelo lógico.	21
	3.1.1. Resultados	21
	3.1.2. Dimensión de procesos o mecanismos de implementación	24
	3.1.3. Dimensión estructural. Los recursos, la organización y otros elementos estructurales.	28
	3.1.4. La perspectiva sistémica	30
	3.2. Múltiples perspectivas sobre las teorías que subyacen en una intervención	33
	3.2.1. Fuentes para obtener pistas y sugerencias para construir la teoría	33
	3.3. Modelo lógico con perspectiva sistémica	34
	3.4. Graficar la teoría	34
	3.5. Determinación de los aspectos a evaluar	35
	3.6. Métodos y técnicas	37
	3.7. Análisis.....	38
	3.8. Ventajas de la evaluación de la teoría.....	42
	3.9. Debilidades y críticas	43
4.	CONCLUSIONES FINALES	44
5.	REFERENCIAS	47

1. Introducción

No hay una única forma de evaluar. En un primer estadio de la evaluación que se puede fijar en torno a los años 50 del siglo XX, quizás sí existía un amplio acuerdo sobre el método a utilizar, el experimental: los diferentes actores esperaban validar los programas en función del cumplimiento de los objetivos declarados. La evaluación aportaba confianza y optimismo en sus posibilidades. Pero los resultados de las evaluaciones no llegaban a las expectativas depositadas, había dificultades prácticas, críticas metodológicas, problemas técnicos y resultados no concluyentes; algo que alejaba el ideal de convertir la evaluación en el apoyo científico de la decisión política.

Después de cinco décadas de realismo en cuanto al uso y las posibilidades de la evaluación, han emergido diferentes métodos, reflexiones y aproximaciones sobre cómo hacer evaluación. Por tanto, existen diversos abordajes, con diversos autores y escuelas que piensan y trabajan de forma distinta. Stufflebeam (2001) los ha clasificado en al menos 22 grandes corrientes. Esto quiere decir que hacer evaluación también implica decidir sobre el procedimiento a utilizar entre una amplia gama de opciones.

En este contexto teórico llama la atención que en el sector de la cooperación para el desarrollo se dé la presencia hegemónica de un determinado método de evaluación, como es el **modelo preordenado, por criterios y estándares** propuesto por el Comité de Ayuda al Desarrollo (CAD) de la Organización para la Cooperación y el Desarrollo Económico (OCDE).

El presente texto pretende revisar el modelo en sus virtudes y defectos y proponer al menos otra forma alternativa de proceder, la **evaluación de la teoría del programa**. No deja de ser una limitación el proponer solamente una única alternativa sabiendo que existen otras aproximaciones lúcidas y útiles. Pero el objetivo del texto es mostrar que hay otros métodos posibles y no aspirar a convertirse en una recopilación de teorías y modelos, algo que por otro lado ya está escrito.

2. Evaluación orientada por criterios

Un criterio es una definición sobre la situación deseable u óptima (Bustelo, 2010) de un proyecto en una determinada parcela de la realidad. De forma resumida se puede decir que la evaluación por criterios funciona sometiendo a examen al programa y valorándolo con relación a ciertos criterios establecidos. Para poder emitir un juicio sobre el programa, se requiere establecer unos puntos críticos en cada uno de los criterios, a los que se les denomina estándares. Si se alcanza el estándar se considera logrado el criterio. Por ejemplo, en el criterio eficiencia se marca como estándar que los beneficios superen los recursos invertidos. Si la evaluación concluye que esto es así el programa se considera eficiente.

Para diseñar una evaluación por criterios se proponen las siguientes fases¹:

Fases de la evaluación orientada por criterios:

- 2.1. Elección entre modelos estandarizados o modelos *ad hoc*
- 2.2. Identificación de fuentes y actores clave
- 2.3. Extracción de las necesidades informativas
 - a) Fase divergente
 - b) Fase convergente
- 2.4. Operacionalización
 - a) Trabajo vertical. Definición de los criterios y preguntas
 - b) Trabajo horizontal. Definición de indicadores y fuentes

2.1. Elección entre modelos preordenados o modelos *ad hoc* sensibles a las necesidades de los actores implicados

La evaluación preordenada (también llamada estandarizada) es aquella donde las cuestiones a evaluar están previamente establecidas. Por el contrario, la evaluación *ad hoc* o sensible a las necesidades de los diferentes actores se construye identificando los valores, ya sean explícitos o subyacentes, de los diferentes agentes implicados. Es en el propio proceso de evaluación donde emergen los criterios o las preguntas que se consideran relevantes evaluar.

El éxito de los modelos preordenados probablemente radique en utilizar un mismo formato para evaluar las intervenciones, sea cual sea la evaluadora² y esté donde esté la acción. Unifican el lenguaje y disipan incertidumbres sobre cómo acometer el trabajo, sobre todo en las primeras fases ya que sólo hay que disponer los criterios preestablecidos y los mecanismos que proponga el método. Tal es el caso del modelo propuesto por el CAD de la OCDE (1995). El CAD plantea cinco criterios para evaluar las acciones de cooperación para el desarrollo: impacto, eficacia, eficiencia, pertinencia y sostenibilidad. Esta propuesta ha tenido una gran aceptación y ha sido asumida por los principales actores de la cooperación internacional. En España, la Secretaria de Estado para Cooperación Internacional del Ministerio de Asuntos Exteriores y de Cooperación la desarrolló y difundió a través de su manual *Metodología de Evaluación de la Cooperación Española* en el año 1998.

La idea de una estructura única de evaluación ha generado críticas. Lo considerado bueno u óptimo, ya sea en cooperación o en cualquier otro ámbito, no es universal, depende del contexto y de los actores implicados en esa valoración.

En el campo de la acción pública existen diversos actores con condiciones y posiciones heterogéneas, lo que provoca diferentes perspectivas sobre el programa o sobre lo que debe ser evaluado. Lo que valora como positivo un político sobre un sistema de salud puede que no sea relevante para el beneficiario y puede también

¹ Gran parte de lo que a continuación se va a narrar procede del texto *Operacionalización*, (Ligero, 2010 (en curso)). Materiales del magíster en Evaluación de Programas y Políticas Públicas, Universidad Carlos III de Madrid.

² Convención de género en la redacción. Dentro de las diferentes propuestas para no invisibilizar a ningún género, se ha optado en este texto por utilizar el femenino para referirse a los y las evaluadoras, el resto de los actores serán citados en masculino.

ser diferente a lo que esperan los trabajadores. Por tanto, los actores implicados en un programa pueden tener diferentes expectativas y deseabilidad hacia lo que ofrece una intervención. No se puede presuponer una igualdad de valores entre los diversos actores, ya que ni los contextos sociales ni las condiciones que soportan son homogéneas.

A esta misma conclusión han llegado las metodologías que se conocen bajo el nombre de calidad. En un primer momento, a mediados del siglo XX, se entendía la calidad como la tenencia de un conjunto de características de un bien o servicio que se ajustaban a una descripción predeterminada. Es lo que Bouckaert ha denominado *la perspectiva orientada por el productor* (Löffler: 23). En los años 50 y 60 el concepto de calidad cambia. Se reconoce la existencia de un mercado fragmentando con diversos clientes con sus respectivas necesidades y que además no tienen porqué coincidir. La calidad pasa a ser una función de las expectativas y percepciones individuales (Ibídem: 24). Por esta razón se ha denominado calidad subjetiva.

El concepto de calidad ha transitado desde una visión única, la del productor del servicio, hacia las percepciones que de ella tienen los diferentes clientes. Este cambio de enfoque no ha supuesto ninguna pérdida de operatividad o bloqueo metodológico, sino más bien una ganancia al producirse un análisis más realista del sistema y de los actores implicados.

En evaluación ha ocurrido lo mismo, se ha ido introduciendo una perspectiva que también podría llamarse subjetivista en la medida que atiende a sus diversos “clientes”. Cronbach argumenta que la evaluación que permite a un solo grupo definir las cuestiones a indagar tiende a no ser útil ya que la toma de decisiones en sociedades democráticas está caracterizada por el pluralismo y la evaluación debe tener en cuenta este aspecto (Chen: 61). Aceptar estas premisas supone preguntarse qué actores están implicados en los programas, quiénes tienen legitimidad para participar en un proceso de evaluación y cómo hacer que las evaluaciones respondan a las diferentes necesidades para que sean útiles.

La orientación hacia los actores se le ha llamado pluralismo o *responsiveness*³ (Stake, 2006), que podría ser traducido como evaluación sensible a las necesidades de los diversos actores implicados con el programa. Este enfoque recoge una corriente de pensamiento con diversos y significativos autores como Cronbach, Patton, Guba, Lincoln, Weiss, House y Monnier entre otros. Pero también ha calado en otras formas de evaluación más “objetivistas” que empiezan a tener cierta empatía hacia los requerimientos de los *stakeholders*⁴. Rossi y Chen defienden un enfoque que integre ambos métodos (Chen). El mismo CAD describe que se pueden descartar o añadir nuevos criterios adicionales dependiendo de las preguntas evaluativas y de los objetivos de la evaluación (CAD, 2010: 9). En el ámbito nacional, la DGPOLDE⁵ también reconoce que puede haber situaciones donde “se necesite recurrir a criterios complementarios”, incluso propone un listado de ellos (MAE, 2007: 57).

Aunque se opte por modelos preordenados de criterios, actualmente hay cierta flexibilidad para adaptarse a los diversos contextos y esquemas valorativos de los diversos actores.

³ En castellano, existe la voz mexicana responsivo: perteneciente o relativo a la respuesta (RAE, 2001).

⁴ Actor afectado o que afecta al programa o a la evaluación (Bustelo, 2010).

⁵ Dirección General de Planificación y Evaluación de Políticas de Desarrollo.

2.2. Identificación de fuentes y actores clave

Como se ha visto en el apartado anterior, en un sistema de acción, como es un programa, existen diferentes *stakeholders* que podrán tener diferentes esquemas y valores (Stake, 2006). La dificultad a la hora de diseñar una evaluación estriba en identificar las fuentes a las que se les otorga legitimidad para participar en el proceso, detectar sus sistemas de valores y necesidades informativas. Eric Monnier (1992) propone tener en cuenta tres subsistemas de actores:

- **Subsistema de la legitimación:** se compone de actores clave a los que denomina responsables legítimos de la decisión. Esta legitimidad puede ser otorgada de forma jurídica, política o científico-técnica.
- **Subsistema de acción:** son los actores que de una forma u otra implementan el programa. Por ejemplo, técnicas, dirección o colaboradoras.
- **Subsistema de reacción:** lo componen actores sociales heterogéneos, que reaccionan ante el programa, ya sea porque son los beneficiarios, porque no reciben los beneficios o porque se erigen en defensores de uno u otro grupo.

Se puede añadir a este listado un cuarto actor que es el propio equipo de evaluación. A las evaluadoras se les confiere legitimidad para participar aunque con ciertas limitaciones. Cronbach advierte que “el evaluador debe buscar cuestiones procedentes de la gente del lugar; (...) se excedería en su autoridad si todas las cuestiones del estudio fueran de su invención” (Ibídem: 163).

Otra fuente importante es la comunicación institucional (programación, memorias, publicaciones, producciones audiovisuales, registros, etc.). La documentación representa un determinado discurso de la institución, normalmente el de los actores con mayor poder. Otros actores sin acceso a las publicaciones pero con capacidad de influencia sobre la práctica pueden tener otras visiones diferentes no recogidas en la documentación. Por eso el estudio de la producción documental o audiovisual normalmente es el estudio del discurso institucional.

Por último, hay perspectivas externas que pueden tenerse en cuenta e incluirse en el proceso de diseño de evaluación a través del estudio de la bibliografía especializada, de las investigaciones, de los meta-análisis, de las meta-evaluaciones o de cualquier otra reflexión teórica. Es una forma de incluir un conocimiento social más amplio y de ajustar la evaluación a los criterios de la comunidad científica.

Esta propuesta de fuentes y actores no es normativa, hay otros autores y autoras que lo pueden entender de otra manera. Como por ejemplo la evaluación colaborativa de Liliana Rodríguez, dicho modelo define a priori quién participa y cuál debe ser su rol. En otros modelos como las aproximaciones “objetivistas”, o también llamadas “científicas”, no se recogen las perspectivas de los diferentes actores, incluso las rechazan por entender que es una fuente de sesgo. Scriven argumenta a favor de tener el menor contacto posible con el cliente por el riesgo a ser influenciado. Aunque la crítica que se hace a estos modelos es que responden a las perspectivas de las evaluadoras enmascarándolo de supuesta objetividad.

2.3. Extracción de las necesidades informativas

El término necesidades informativas hace mención a los requerimientos que cada actor tiene con relación a la evaluación. El objetivo de esta fase es detectar las necesidades de información de los actores. Se aplican técnicas de recogida de información y de análisis realizando un **primer trabajo de campo**⁶.

De esta tarea pueden emerger criterios de valor, preguntas, posibles sistemas de medición, indicadores o consejos útiles para desarrollar la evaluación, incluso predisposiciones hacia la evaluación. Para las preguntas de evaluación Cronbach realiza una propuesta para poder seleccionarlas. Diferencia entre una fase divergente y otra convergente (Stufflebeam, 1995).

2.3.1. Fase divergente: recopilación de las necesidades informativas

Se trata de recopilar todas las posibles necesidades de información. Para extraer dichas cuestiones se puede acudir a distintas fuentes, como puede ser la documentación del programa o preguntar a los diferentes actores implicados. La labor de la evaluadora es extraer las preguntas o bien relacionadas con los criterios estandarizados o con los propuestos por ellos mismos.

Tabla 1. Qué preguntas realizar

Modelo de criterios	Modelo preordenado	Identificación de preguntas relacionadas con los criterios de valor. Por ejemplo: ¿Qué le preocupa sobre la pertinencia del programa? ¿Qué cosas le preocupan sobre la sostenibilidad?
	Modelo ad hoc, sensible a los diferentes actores	Preguntas para extraer criterios: ¿Qué cuestiones le parecen más importantes evaluar? ¿Qué es lo que debe valorarse de este programa? Identificación de preguntas relacionadas con los criterios de valor mencionados por ellos mismos. Por ejemplo: De la calidad que ha comentado, ¿qué cuestiones le preocupan?

Aunque las preguntas están muy estructuradas se suele obtener una información variada en la que se entremezclan preguntas, criterios, indicadores, consejos varios, propuestas metodológicas, impresiones, valoraciones del programa o cualquier otro aspecto relacionado con la evaluación o el programa. Toda esta información hay que depurarla y ordenarla, seleccionando preguntas, también llamadas subcriterios de evaluación. Para filtrar las preguntas se pueden tener en cuenta los siguientes aspectos:

- Las preguntas deben implicar una valoración y no una mera indagación sobre algún aspecto del objeto.

⁶ El término primer trabajo de campo no implica necesariamente viajar a terreno. Se pueden obtener las necesidades informativas a través de diversas técnicas y vías aunque implicará un trabajo de obtención, recopilación y tratamiento de la información.

- Deben estar dirigidas al programa (o a su relación con las necesidades sociales) y no a sucesos o dinámicas sociales, ya que esto sería otro objeto de estudio más propio de la investigación.
- No deben ser indicadores sino que deben transmitir las cuestiones que preocupan a los actores implicados indistintamente de que especifique o no cómo se van a medir.

Teniendo en cuenta las salvedades hechas, la intención de esta fase es obtener un listado de preguntas cuanto más completo mejor, sin limitaciones. Los posibles sesgos no deben estar causados por la perspectiva parcial de ningún actor, o como advierte Cronbach “por la restrictiva visión del patrocinador, administrador o de quienes tengan la responsabilidad sobre el programa” (Stufflebeam, 1995: 152).

2.3.2. Fase convergente: selección de las necesidades informativas

No todas las preguntas identificadas tienen que ser contestadas por la evaluación; ya sea porque son poco importantes, porque despejan pocas dudas o porque no pueden ser asumidas por los costes que implicaría obtener los datos para responderlas. Según Cronbach se requiere de una fase previa que permita la selección de las preguntas en función de una serie de criterios, a esto le ha llamado fase convergente. El autor propone dos criterios a tener en cuenta.

- **Influencia:** entre las diversas preguntas puede haber algunas más importantes para la toma de decisiones o el funcionamiento del programa, éstas son las que se priorizarían en la selección.
- **Incertidumbre:** hay cuestiones sobre las que hay pocos conocimientos o estos no son confiables y por el contrario, hay otro tipo de cuestiones que aunque la evaluación pueda aportar conocimiento se tiene bastante certeza y seguridad sobre cuál va a ser la respuesta. La evaluación debería centrarse sobre las más inciertas.

Para ayudar a la selección se puede crear una tabla multicriterio para establecer una priorización de las preguntas (ver tabla 2).

Tabla 2. Multicriterios para priorizar las preguntas

	Gran influencia	Poca Influencia
Gran incertidumbre	Preguntas muy prioritarias	Preguntas prioritarias
Poca incertidumbre	Sólo información incidental	Sólo información incidental

El peligro de este ejercicio es que se dé más importancia a la perspectiva de unos actores sobre la de otros. La selección final de preguntas será la estructura de la evaluación, probablemente la parte más importante de todo el proceso de diseño. Todo lo demás es pensar en cómo responder a estas cuestiones identificadas.

2.4. Operacionalización

La operacionalización es un proceso deductivo que desde los criterios va estableciendo los subcriterios (preguntas), los indicadores y sus correspondientes fuentes de verificación. De forma secuencial y lógica se va

pasando por diferentes fases que desde el concepto más general y abstracto van aportando un mayor grado de concreción. Esta cadena finaliza en el diseño o búsqueda de las fuentes de extracción de información para cada uno de los indicadores.

La finalidad de la operacionalización es obtener una información, que de forma sistematizada y coherente, permita establecer en cada uno de los criterios un enjuiciamiento sobre el programa.

La operacionalización se concreta en dos fases (Bustelo, 2010):

- **Trabajo vertical:** explicitación de criterios. Definición de preguntas.
- **Trabajo horizontal:** formulación de indicadores. Búsqueda de fuentes de extracción de información.

2.4.1. Trabajo vertical: definición de los criterios y preguntas

Una vez que se obtienen los criterios, ya sean preestablecidos o extraídos de los diferentes stakeholders, es necesario concretarlos a través de preguntas (también llamadas subcriterios) para que puedan ser evaluables. Los criterios reflejan un estado deseable e ideal para el programa pero requieren al menos de un nivel más bajo y operativo que concrete ese ideal. Aunque los criterios estén muy bien acotados y definidos es inevitable que den cabida a múltiples significados.

Por ejemplo, la pertinencia puede estar claramente definida: “la adecuación del programa o servicio en todas sus dimensiones al contexto; entendiendo como contexto tanto las necesidades de la población, las políticas y ordenamientos jurídicos, las estructuras sociales y culturales, la teoría y el método de intervención e incluso hasta la satisfacción de los agentes implicados” (Ligero 2010). Pero aún partiendo de esta misma definición, el criterio se expresa de múltiples formas. En la Tabla 3 se muestran dos operacionalizaciones diferentes de pertinencia basadas en la misma definición.

Tabla 3. Ejemplos de dos operacionalizaciones de pertinencia

Programa de prevención de drogas en los colegios elaborados por técnicos locales	Programa de prevención de drogas en los colegios elaborados importado de otro país
¿La intervención está apoyada en una teoría sobre el cambio de los individuos?	¿El programa está adaptado a las características socioculturales de la población? (¿se entiende?, ¿se tienen planteamientos similares ante las drogas?)
¿Existe una congruencia lógica entre las actividades y los fines que se pretenden conseguir?	¿Las evidencias o teorías asumidas están validadas para diferentes contextos y países?
¿Es congruente con las necesidades manifestadas de la población?	¿Encaja bien con el engranaje educativo local? (coordinación, horas, asignaturas)
¿Es congruente con la normativa y legislación local y nacional sobre drogas?	¿Se han apropiado del proyecto?

Cualquiera de las dos operacionalizaciones son válidas, en los dos casos se evalúa la pertinencia pero de forma distinta. Por lo tanto, en cada evaluación, siempre y cuando ésta no sea estandarizada, se está realizando una definición particular de cada criterio.

Algunos modelos preordenados parten de preguntas ya establecidas, como es el caso del TQM-EFQM (MAP), aunque es bastante usual dejar cierta flexibilidad en la operacionalización para permitir la adaptación del modelo de evaluación al contexto. Tal es el caso de la DGPOLDE, que aconseja utilizar un modelo por criterios pero propone que las preguntas reflejen los intereses de los actores implicados incluidos los de las evaluadoras (MAE, 2007: 61).

2.4.2. Trabajo horizontal: definición de indicadores y fuentes

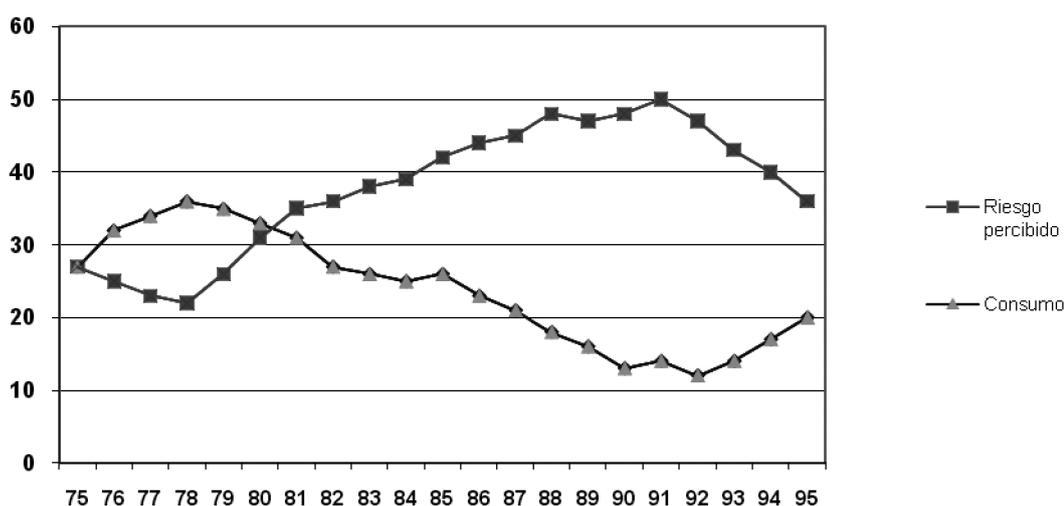
Las preguntas de evaluación son la definición del criterio, no tienen atender al hecho de que puedan ser medidas o no. El indicador, en cambio, es el recurso que permite medir el concepto definido, por lo general a través de aproximaciones. Mientras las preguntas se tienen que encargar de representar lo más completamente posible las diversas cuestiones que definen el criterio, el indicador debe ser una variable medible, observable y registrable que aporte una información que conteste a las preguntas.

“La generalidad de los conceptos constituyen variables ‘latentes’, ‘hipotéticas’, no directamente ‘observables’. Por lo que su concreción precisa de la traducción del concepto teórico a indicadores, a variables empíricas (observables o manifiestas), que midan las propiedades latentes enmarcadas en el concepto” (Cea: 125).

Los indicadores, por tanto, buscan acercarse lo más posible al concepto o al criterio a medir, aunque frecuentemente tienen que conformarse con reflejar la realidad de fenómenos paralelos, similares o parciales del concepto buscado aportando a la evaluación solamente una aproximación. Tal como describe Rutter (pág. 40) “los indicadores son rasgos que tienen relaciones directas con los procesos causales, pero que en sí mismos no forma parte de los mecanismos que están directamente asociados con la causalidad”.

Por ejemplo, un programa de prevención escolar tiene como objetivo reducir el consumo de cannabis. A la hora de medir el consumo se plantean problemas, si se pregunta directamente a los consumidores éstos pueden falsear la respuesta para evitar ser identificados. Si por el contrario se buscan pruebas médicas se necesita el consentimiento de las personas a las que se les va a practicar y los consumidores pueden buscar estrategias para no dar su consentimiento. Aunque el consumo de cannabis sea un concepto bastante concreto, es difícil obtener la información directamente. Pero existen estudios previos como el de Bachmann (2002) que muestran la alta correlación existente entre la percepción del riesgo y el consumo (ver gráfico 1).

Gráfico 1: Correlación entre consumo y riesgo en el consumo de marihuana.



Fuente: Elaboración propia a partir de los datos obtenidos en SAMHSA.

Aunque la percepción del riesgo no es el concepto en sí mismo que se quiere medir, se puede asumir que observando el riesgo, algo a lo que los jóvenes sí están dispuestos a contestar, podemos extraer conclusiones sobre el consumo. No es una relación directa pero sí es una correlación bastante alta que permite acercarnos al concepto. La percepción de riesgo es un instrumento, una variable medible que nos aproxima al concepto que se quiere evaluar, es decir, un indicador.

Los indicadores deben ser redactados como condiciones lógicas que si se cumplen contestan a la pregunta de evaluación. Se fija un estado, un punto o una condición a partir del cual se puede contestar afirmativa o negativamente a la pregunta formulada. En el caso del consumo de cannabis la redacción del indicador podría ser el que se indica en la tabla siguiente (ver tabla 4).

Tabla 4. Ejemplo de pregunta de evaluación e indicador

Pregunta de evaluación	Indicador
Disminución del consumo de marihuana entre los escolares.	Incremento con relación al semestre pasado de la percepción del riesgo en el consumo de marihuana/hachís.

La condición, que en el ejemplo citado es “incremento con relación al semestre pasado”, es el estándar. Un estándar es “una cantidad, un nivel o una manifestación de un criterio determinado que indica la diferencia entre dos niveles de mérito distintos” (Stake, 2006: 111). Los estándares pueden ser dicotómicos o establecer gradientes, ser flexibles o estrictos, relativos o absolutos, cualitativos o cuantitativos pero de alguna forma tendrán que explicitar en qué punto o momento la pregunta de evaluación es contestada.

2.5. Consideraciones sobre la evaluación por criterios

La secuencia lógica de una evaluación por criterios, la que se ha desarrollado en este capítulo, es la que propugna Fournier (Stake, 2006: 242).

- Establecimiento de criterios sobre el mérito del programa: ¿en qué dimensiones debe funcionar bien el evaluando?
- Definición de estándares: ¿cuál es el funcionamiento mínimamente bueno que debe alcanzar el evaluando?
- Medición del rendimiento y comparación con los estándares: ¿hasta qué punto funcionó bien el evaluando?
- Síntesis e integración de los datos en un juicio o valoración del mérito de la valía: ¿cuál es el mérito o la valía del evaluando?

Según esta forma de trabajar la evaluación, el producto final que se obtiene es una serie de juicios en las áreas definidas previamente (criterios). Es decir, si la información obtenida por los indicadores es viable y fiable, si se logra el nivel definido por los estándares, entonces las preguntas serán contestadas afirmativamente y, por consiguiente, el juicio sobre el programa será positivo. El programa será eficaz, eficiente, pertinente, viable, etcétera.

Hay dos aspectos de esta secuenciación para el enjuiciamiento que son interesantes remarcar. La primera es que el juicio se establece en los estándares de los indicadores, son realmente los criterios últimos de valoración. Los grandes criterios son sólo áreas de análisis dentro de las cuáles los estándares trabajan. Por lo tanto, la definición de los indicadores no es un mero trabajo técnico, sino que son los que determinan la valoración final del programa.

La segunda cuestión es que esta dinámica valorativa conduce a un enjuiciamiento sumativo sobre el programa (según la clasificación de Scriven). Por ejemplo, la evaluación de un proyecto de educación profesional básica apoyada en la estrategia de la alternancia (períodos en el centro educativo y períodos en el medio socio-profesional) en una zona rural andina⁷ fijó una serie de criterios: impacto, eficacia, eficiencia, cobertura, sostenibilidad, apropiación y calidad. Se establecieron unas preguntas y unos indicadores con sus respectivos estándares. Tras la realización del trabajo de campo y el análisis de los datos se obtuvieron las siguientes valoraciones:

- Detectaron varios impactos positivos (los cuales son relatados aunque no son concluyentes).
- Valoraron que la eficacia, la eficiencia, la cobertura y la calidad eran altas.
- La apropiación del proyecto se consideró media.
- La valoración de la sostenibilidad queda en suspenso por falta de información en el momento de hacer la evaluación.

Además de estos juicios, el informe de evaluación describe el procedimiento realizado, cómo se han obtenido los datos, la interpretación de los mismos para llegar a los juicios y propone recomendaciones y lecciones

⁷ Proyecto de la Fundación CODESPA.

aprendidas. El informe puede dar para más lecturas que sin lugar a dudas resultarán interesantes, pero realmente el producto final de este modelo es la valoración de los siete criterios expuesta en los cuatro guiones anteriores.

Lo que produce este tipo de evaluación, además de las recomendaciones, es enjuiciamiento. Algo que guarda una estrecha relación con el rendimiento de cuentas público, también llamado responsabilización.

2.6. Críticas a la evaluación orientada por criterios y estándares

Críticas referidas a los modelos de criterios preordenados (o estandarizados):

1. **Se considera que es una evaluación parcial:** como se ha visto anteriormente, existe un gran consenso en la literatura sobre que los diferentes grupos de interés implicados en la evaluación no tienen porqué compartir los mismos intereses, necesidades informativas o criterios sobre la calidad de un programa. El propio concepto de calidad varía dentro y entre los diferentes grupos de stakeholders (Greene; Stake). Por tanto, los modelos preordenados y los estándares no suelen reflejar de forma homogénea los intereses evaluativos o lo que consideran bueno los diferentes actores. Asumir una forma estándar para valorar una intervención es responder a la visión de un determinado punto de vista que puede ser que no coincida con la de los técnicos, los beneficiarios u otros actores.

Los enfoques integrados propuestos por Rossi y Chen (1990) intentan superar esta primera crítica sugiriendo una negociación entre los actores clave para obtener un diseño de evaluación de consenso y que satisfaga las diversas demandas. Pero Stake (Shadish y otros) advierte que la evaluadora no debe crear un consenso inexistente. Aunque se lograra un acuerdo, fallaría en su aproximación a las realidades diversas de una sociedad (Stake cit Greene: 164). La aplicación de estándares universales incluso puede entenderse como la plasmación de una determinada forma de autoridad (Abma cit Greene).

Los modelos preordenados son métodos parciales con una utilidad limitada para el conjunto de actores, que no tienen en cuenta las diferentes prioridades e imponen una determinada visión de la calidad y el valor de la intervención.

2. **Evaluación sesgada. Pretensión de objetividad:** muy relacionado con el punto anterior es que los modelos preordenados, en la medida en que son propuestas explícitas, generalizadas y asumidas por diferentes actores, corren el riesgo de ser tratadas como “objetivas” e independientes de los propios actores que las formulan. Por ejemplo, los cinco criterios del CAD representan un determinado sistema de valores, el de la OCDE, organización de los Estados más industrializados del mundo y principales donantes de la cooperación internacional. Es lógico pensar que para ellos el criterio de sostenibilidad es muy importante, ya que se fija en que la ayuda sea coyuntural y no estructural, es decir, no permanente.

Pero gran parte de las políticas (muchas de ellas objeto de cooperación internacional: asistencia sanitaria, asistencia social, mantenimiento del orden, seguridad y seguridad jurídica entre otras), no son por definición sostenibles ni pretenden serlo. Sus objetivos procuran satisfacer necesidades impostergables o impedir (o minimizar) perjuicios derivados de circunstancias adversas. En definitiva, mantener bajo control la desorganización que sobrevendría si no se lograrán. Son intervenciones indispensables, necesarias pero no sostenibles desde un punto de vista economicista (Hintze, 2005:147). El autor declara

que “las medidas simples y puntuales que cada tanto producen efectividades sostenibles como por arte de magia, parecen ser precisamente eso: arte de magia y como tales no parecen demasiado frecuentes en el mundo de la gestión de lo público”.

La inclusión del criterio de sostenibilidad en el análisis de la acción pública implica una particular visión de cómo deben ser y cómo deben funcionar las políticas y proyectos financiados, lo que quizás no encaje con la visión de los poderes públicos locales o de los beneficiarios. La perspectiva del CAD es legítima pero interesada, responde a su propia visión e intereses.

El peligro no está en que se propongan determinados criterios de evaluación, sino en que se entiendan estos como métodos objetivos y permanezcan en la sombra las intencionalidades que los motivan, algo que Abma ha definido como “autoría enmascarada” (Greene: 169).

Para evitar este sesgo metodológico existen propuestas, como por ejemplo la de Scriven, que invita a alejarse de los clientes, de su encargo y de la forma en que hay que hacer evaluación. Pero esto no hace más que sustituir los criterios del cliente por los de la evaluadora, algo que Macdonald ha denominado evaluadora autocrática.

Si volvemos la mirada hacia los autores apoyados en paradigmas constructivistas, encontramos que el valor de una intervención está depositado en el que mira y no reposa en la sustancia del objeto evaluado (Greene: 3). Incluso Scriven, uno de los mayores buscadores de objetividad en el campo de la evaluación, entiende que ésta se alcanza a través de un acuerdo intersubjetivo (Chen: 62).

El “pensamiento negociado” necesario para llegar a un valor compartido sobre la calidad y el valor requiere tiempo (Stake). Cuando los criterios están claramente definidos y estructurados y no permiten una reinterpretación o transformación de los puntos críticos de la calidad del programa por otros actores implicados, se está ante una imposición del valor en ocasiones camuflada de objetividad.

- 3. Los modelos de criterios tienden a centrarse casi exclusivamente en la dimensión de resultados:** en el modelo CAD el impacto y eficacia miden exclusivamente el cumplimiento de la jerarquía de los objetivos. La eficiencia analiza la relación de dichos resultados con los “inputs” del programa. La sostenibilidad quiere valorar el mantenimiento de la propia intervención y de sus resultados a largo plazo. Y, por último, la pertinencia según la entiende el CAD, se centra principalmente en la coherencia de los objetivos del proyecto con los objetivos de los donantes y del gobierno. Tal como se puede ver, en los cinco criterios la dimensión clave son los objetivos/resultados sobre la cual pivotan los juicios.

A las evaluaciones de resultados se les ha denominado de Caja Negra (Chen; Weiss). Se caracterizan por mirar solo a los resultados o como mucho a los inputs, sin preocuparse por el proceso intermedio de transformación.

La evaluación de resultados no se fija en la implementación ni en aspectos estructurales, por lo tanto “falla en identificar las deficiencias para futuras mejoras y desarrollos del programa” (Chen: 18). Debido a estas dificultades Stufflebeam las denomina cuasi-evaluaciones ya que emiten una valoración de un todo, el programa, valorando sólo una parte, los resultados.

4. **Problemas con los objetivos:** el término objetivo tiene diferentes acepciones que lo convierte en un concepto engañoso. Por un lado, la palabra se refiere a la meta, fin, misión, en definitiva a la “diana” que tiene que alcanzar el programa. Pero objetivo también significa algo ajeno a las percepciones, no expuesto a las subjetividades, es decir, sin sesgo. Los objetivos son propuestos por actores, que voluntaria o involuntariamente representan determinadas posiciones. Si se pide a los diferentes actores de una organización que definan los objetivos, no los que están en los documentos sino los que realmente quieren alcanzar, es probable que muchos de éstos no coincidan. Ralph Tyler, en su evaluación “Eight Year Study” sobre 30 “high schools”, encontró que en el primer año solo dos o tres centros estaban trabajando de acuerdo a lo descrito en el programa (Weiss: 9).

En la literatura se encuentran muchas otras referencias que advierten sobre la distancia entre los objetivos “oficiales” y los que al final se aplican. Stake habla de “objetivos declarados”, aludiendo a aquellos que están en el papel que no tienen porque coincidir con los que realmente trabaja la institución. Los objetivos, entiende Monnier, son la resultante de un proceso de negociación. Los objetivos declarados, los que están en la documentación o matrices de programación, suelen representar la visión de los actores con más poder. Incluso se llega a entender la política “como una declaración de intenciones por parte de actores que desean convencer o controlar a otros actores” (Deleau cit. Monnier: 105).

En ocasiones se superan las tensiones entre los diferentes actores implicados en una intervención, llegando a acuerdos sobre los objetivos superiores (general y específico) con formulaciones generales trasladando el disenso a los objetivos operativos (resultados esperados).

Estos condicionantes hacen que en un proceso de evaluación “los objetivos reales no se obtienen en documentos o en conversaciones, mejor ver a qué dedican su tiempo los actores implicados” (Weiss: 55) ya que se puede llegar al extremo de que la institución trabaje realmente en pro de unas metas que disten mucho de parecerse a las que tiene escritas; algo bastante alejado del concepto de metas insesgadas.

Criticas a todos los modelos de criterios y estándares:

5. **Problemas de los estándares:** hay que tener en cuenta que en el caso de los criterios de eficacia e impacto, los estándares son los objetivos declarados.

- **Reducción del valor del programa a un punto final.** El punto final o estándar simplifica el juicio pero también reduce a una sola medida la complejidad de una intervención. “Podemos medir la productividad de una persona que envasa mantequilla pero de nada nos sirve si para hacerlo se lame el pulgar antes de coger el papel de envasar” (Messick cit. Stake, 2006: 248).

Es un proceso de reducción y de pérdida de información sobre la variada y rica realidad de la intervención y rara vez mejora el conocimiento sobre lo que se hace.

- **Alcanzar el estándar.** Si no se alcanza el estándar fijado, no se puede emitir un juicio positivo sobre la intervención aunque el trabajo realizado haya sido encomiable. Si, por ejemplo, en un determinado indicador se fija un cumplimiento de 100 y se obtiene 98, el juicio que se emita sobre ese aspecto debe ser negativo aunque se haya logrando un 98%. En este caso dos puntos pesan más que los 98 restantes.

- o **Estándares desajustados:** si los estándares declarados son ambiciosos, el programa no será eficaz aunque el dispositivo sea excelente. Por el contrario, si los estándares son muy humildes siempre habrá una excelente versión de lo que ha ocurrido aunque el programa esté en mínimos e infrautilizado.
 - o **Legitimidad de quién fija los estándares:** teniendo en cuenta que sobre los estándares recae el peso del enjuiciamiento final, ¿quién o quienes están legitimados para fijar los estándares? En ocasiones se encubre como una mera decisión técnica lo que es de calado estratégico o político. Stein encuentra que frecuentemente los líderes políticos encubren el debate de cuestiones importantes sobre el valor del programa en supuestamente neutrales decisiones técnicas acerca de los indicadores y las medidas (House, 126).
6. **Visión túnel:** al fijarse la evaluación en unos determinados criterios pueden pasar desapercibidos otros elementos importantes de la intervención o sus resultados. Este efecto se ha definido como “visión túnel” (Leeuw y Furubo). La crítica no se ciñe sólo a una pérdida de la riqueza de evaluar, sino que evaluar sólo lo que está escrito puede ser una forma elegante de controlar a la evaluadora. Un centro de menores infractores puede tener definido claramente los aspectos que se pretenden alcanzar relacionados con la eficacia, el impacto o la pertinencia conformando una evaluación completa. Pero puede que no hayan tenido en cuenta y, por lo tanto no habrá indagado nada sobre casos de palizas y malos tratos. Scriven procura evitar el sesgo que produce tanto el cliente como el encargo, con su modelo de evaluación libre de objetivos, centrándose en los cambios en la realidad sin que esto tenga que estar estructurado por la declaración de objetivos.
7. **Enjuiciamiento no comprensión:** probablemente la crítica más importante al método de evaluación por criterios es que no conduce hacia la comprensión de lo que ocurre en el programa sino hacia una valoración sumativa.

El establecimiento de unos criterios de valor, con sus consiguientes estándares (ya sean más o menos estrictos o laxos) genera una cadena lógica que concluye en un juicio sumativo en cada una de las áreas definidas. Como se ha argumentado en el apartado “consideraciones sobre la evaluación por criterios”, se obtienen unas valoraciones y el programa se adjetiva como eficaz o ineficaz, pertinente o impertinente, etc. Obtener valoraciones de un programa tiene interés social y político; en sí mismo justifica una evaluación y contribuye directamente a la responsabilización o *accountability* de la acción pública.

La limitación radica en que esta estrategia no provee de explicaciones de cómo han ocurrido las cosas. Poniendo un ejemplo sencillo, se puede valorar el desempeño de una estudiante y concluir que no ha sido eficaz y ha suspendido el curso. Podemos agregarle más criterios y concluir que es pertinente su plan de estudios porque sus objetivos son coherentes con el centro académico y el método escogido está avalado por la teoría. Pero no nos permite saber más, no sabemos lo que está pasando con el desempeño, quizás la estudiante no sepa aplicar bien las técnicas de estudio, no esté motivada, sea miope, tenga una agenda muy cargada, sufra ansiedad o los profesores no sean buenos. Todas estas dimensiones, factores causales y aspectos relevantes han quedado reducidos a una valoración dual y sumativa.

Se puede buscar una interpretación más pormenorizada e intentar entender qué hay debajo del juicio, nadie lo va a impedir, pero la mecánica establecida en el método de criterios no está dirigida a ello sino a la emisión de juicios. La finalidad comprensiva de la evaluación queda relegada frente a la valoración y el enjuiciamiento y por tanto está más dirigida al rendimiento de cuentas.

El énfasis en el rendimiento de cuentas público (*accountability*) habla también de unas determinadas inquietudes sociales frente a otras. Hace 30 años Lee Cronbach observó que la demanda de *accountability* es un signo de patología social en el sistema político. Jennifer Greene apunta que lo considera una demanda regresiva e irrelevante para su trabajo como evaluadora (Greene: 1).

Lo que remarcan estos autores es que enjuiciar sin comprender obedece a una necesidad de catalogar, enjuiciar o “colocar programas en esquemas” más que a una motivación constructiva de mejora. La mecánica del modelo de criterios tiende a eso, a valorar sin la necesidad de pasar por la comprensión.

3. Evaluación de la teoría del programa

De forma resumida se puede definir la evaluación de la teoría del programa como un proceso que en primer lugar identifica cómo funciona la intervención y qué resultados pretende obtener construyendo un modelo teórico del programa, y, en segundo lugar, estructura todo el proceso de indagación a través de dicho modelo teórico. No solamente se va a evaluar si se han logrado los objetivos sino también se podrán identificar los mecanismos causales que han conducido a ellos.

En este capítulo se pretende describir cuáles son las fases y las reflexiones teóricas principales para el diseño y la realización de una evaluación bajo este prisma. Se incorporan al final unos epígrafes sobre ventajas y debilidades.

Primeras referencias teóricas

La evaluación de la teoría del programa tiene un amplio bagaje teórico. La idea básica ya se puede encontrar en los años 60 en Schuman cuando sugiere que la evaluación podría dirigirse hacia la cadena de objetivos de la intervención. La misma idea se repite en Bennet en el año 1975, quien utiliza el concepto de jerarquía de resultados (Rogers y otros). Progresivamente se va extendiendo la idea de que para evaluar es necesario el desarrollo de un marco teórico previo del programa.

En 1987 Bickman edita una publicación titulada *Using Program Theory in Evaluation* en la que define la teoría del programa como la construcción de un modelo razonable y sensible de cómo se supone que funciona la intervención. En la misma publicación Wholey señala que el propósito de la teoría del programa es identificar los recursos del programa, las actividades y los resultados previstos y especificar las cadenas de asunciones causales vinculando recursos, actividades, resultados intermedios y objetivos finales.

Desde 1972 Carol Weiss invita a identificar posibles modelos causales de los programas. Ella lo denomina teoría del cambio y lo describe como “la cadena de asunciones que explican cómo las actividades conducen paso por paso a los resultados esperados” (1998: 2).

Schon en 1997 utiliza el término teoría de la acción, muy parecido al utilizado por Carol Weiss y los autores Lenne, Cleland y Funnell, que la denominan lógica del programa (Rogers y otros).

Evaluación dirigida por la teoría (*Theory-driven Evaluation*) es el título dado por Chen para una de las monografías de referencia en este enfoque (1990). Lo define como la especificación de “lo que debe ser hecho para alcanzar

la meta deseada. Que otros importantes impactos pueden ser anticipados y como esos objetivos e impactos podrían ser generados” (Chen, 40). Davidson (2000) prefiere denominarla evaluación basada en la teoría, para evitar transmitir la idea de que la evaluación pueda estar condicionada a priori por una determinada teoría.

La palabra teoría despierta también otro tipo de recelos. Rossi la considera demasiado pretenciosa para describir los mecanismos internos que operan en un programa. La alternativa a teoría es el modelo lógico, en ambos casos se refieren a la cadena de asunciones causales que hay detrás de cada intervención. En este texto se utiliza de forma indistinta teoría o modelo lógico del programa.

Los postulados de la teoría del programa están incorporados en diversas perspectivas teóricas. Bob Stake se declara abiertamente cercano a esta forma de proceder. Rossi ve necesario la identificación de las relaciones lógicas en las intervenciones e incluso el experimentalista Tomas Cook se declara *fan* de la evaluación orientada por la teoría (2000: 32).

Fases para el diseño de una evaluación orientada por la teoría

Tal y como se ha definido anteriormente, la evaluación orientada por la teoría consiste en explicitar la teoría o modelo lógico del programa y las relaciones causales entre todos sus componentes y orientar la evaluación según dicho modelo.

Los pasos usuales en este tipo de evaluación son plasmar en un gráfico las relaciones causales que van desde los *inputs* a través de los procesos hasta provocar los resultados (Stufflemban, 2001). Posteriormente se identifican las dimensiones clave que deben ser medidas. Con la información que se obtiene se pueden poner a prueba las relaciones causales y extraer conclusiones sobre el modelo propuesto. Es decir, se trata de indagar sobre si los éxitos o fracasos del programa son debidos a las diferentes cadenas de trabajos y si los medios son adecuados para la implementación. Se proponen una serie de fases:

Fases de la evaluación orientada por la teoría:

3.1. Construcción, reconstrucción o identificación de la teoría del programa o modelo lógico.

3.1.1. Resultados.

3.1.2. Dimensión de procesos o mecanismos de implementación.

3.1.3. Dimensión estructural: los recursos, la organización y otros elementos estructurales.

3.1.4. La perspectiva sistémica.

3.2. Múltiples perspectivas sobre las teorías que subyacen a una intervención.

3.2.1. Fuentes para obtener las teorías que subyacen a una intervención.

3.3. Modelo lógico con perspectiva sistémica.

3.4. Graficar la teoría.

3.5. Determinación de los aspectos a evaluar.

3.6. Métodos y técnicas.

3.7. Análisis.

3.1. Construcción, reconstrucción o identificación de la teoría del programa o modelo lógico

En esta fase el objetivo es identificar, reconstruir o elaborar directamente el marco teórico que establece relaciones causales entre las diferentes dimensiones de una intervención. Se van a destacar tres dimensiones clave: resultados, procesos y elementos estructurales.

3.1.1. Resultados

Probablemente el aspecto más requerido en evaluación son los resultados. Los resultados son los cambios en la realidad producidos o atribuibles a la intervención. Hay tres formas principales de acotar los resultados:

- a. Utilizar resultados estandarizados.
- b. Fijarse en los resultados definidos por los objetivos del programa.
- c. Observar todos los posibles cambios dados en la realidad.

En el caso "a" los resultados que se evalúan son los mismos para diferentes intervenciones. Un buen ejemplo es el informe sobre el sistema educativo PISA. Se comparan los rendimientos académicos de los estudiantes de diferentes países ante pruebas iguales. En el sector de las ONG el informe de la Fundación Lealtad se apoya también en una lógica estandarizada. Todas las ONG son analizadas bajo los mismos baremos indistintamente de las peculiaridades de cada una.

El caso "b" es el modelo más extendido desde que Ralph Tyler en 1942 propuso la evaluación de los objetivos como medida del éxito del programa. Está presente en la propuesta de criterios del CAD (eficacia e impacto). Los objetivos son soluciones proyectadas, metas fijadas por el programa de antemano. Si los objetivos son un deseo, una intención, los resultados son los cambios constatados provocados por la intervención. Como se espera que los resultados sean la proyección de los objetivos, en ocasiones los términos objetivos y resultados se utilizan de forma indistinta, al menos en determinados sistemas de programación⁸, gestión o evaluación. Pero estrictamente los objetivos son meras intenciones y los resultados realidades, en la práctica no tienen porqué coincidir.

Lo primero, entonces, será identificar o programar los resultados pretendidos para poder evaluarlos. Los métodos de programación por objetivos⁹ funcionan básicamente proponiendo un objetivo específico y otros intermedios. Estos últimos son un medio necesario e indispensable para alcanzar el objetivo específico. La jerarquía de objetivos establece con claridad unas relaciones causales, si se logra el nivel inferior se tendrá que lograr el superior.

El soporte de programación suele ser un diagrama o una matriz donde se especifican los diferentes niveles de objetivos, desde el más genérico hasta el más concreto. Por ejemplo, en el caso de un programa que tiene

⁸ Por ejemplo, en el sistema de programación del Enfoque del Marco Lógico los objetivos intermedios se denominan directamente "resultados esperados".

⁹ Uno de los sistemas de programación por objetivos más extendidos es el Enfoque del Marco Lógico o el modelo ZOPP de la alemana GTZ. Se pueden tener en cuenta también los modelos de racionalización presupuestaria como el PPBS norteamericano, el RCB francés o su equivalente RFA de los países bajos (Monnier, 1992: 30).

como objetivo específico “la integración de los jóvenes” para que no se conforme un gueto en un determinado barrio de una gran ciudad, establece unos objetivos intermedios que son:

- 1) Los jóvenes van con gente que no es sólo de su misma nacionalidad o etnia.
- 2) Aprenden a aceptar las diferencias.
- 3) Incrementan su desarrollo personal.
- 4) Se responsabilizan con el entorno y la ciudadanía.

Gráfico 2: Objetivos del Programa Tiempo Joven.



Fuente: elaboración propia

Todos los objetivos que aparecen en el diagrama (ver gráfico 2) son cambios en la población o en el entorno. Todos son resultados aunque los del nivel inferior son meros instrumentos para lograr el superior. En esta intervención y en este proyecto el desarrollo personal, por ejemplo, solamente importa por su cualidad de ser medio para el fin pero en sí mismo carece de finalidad. Cada nivel inferior tiene que ser la condición suficiente y necesaria para que se cumpla el objetivo del nivel superior. De esta forma, se construye una cadena lógica, la modificación de una serie de “variables intermedias” (objetivos intermedios o resultados esperados) deben provocar cambios en el nivel superior (objetivo específico). El Enfoque de Marco Lógico es un ejemplo de esta lógica de programación.

El caso “c” trata de liberarse de las influencias y limitaciones de gerentes o financiadores u otros grupos de poder. Scriven propone observar todos los posibles cambios dados en la realidad. Se buscan directamente los resultados obtenidos por la intervención, sin conocer previamente los objetivos. Esta propuesta también conlleva críticas ya que la identificación de resultados es susceptible de estar sesgada por la propia mirada de quién los busca, en este caso de la evaluadora.

Sea cual sea la forma de acotar o identificar los objetivos se está haciendo un ejercicio de plasmación teórica, se define una relación causal entre diferentes niveles de objetivos. En el gráfico anterior (ver gráfico 2) se proponía una forma de conseguir la integración, pero puede haber otras formas de lograrlo. El modelo de la asimilación cree conseguirlo de esta manera (Giménez) (ver gráfico 3).

Gráfico 3. Asimilación.



Fuente: Elaboración propia.

El multiculturalismo es otra forma de buscar la integración, lo cree conseguir a través de informar sobre los derechos, desarrollar las lenguas y culturas de origen y adaptación de la normativa a las diferencias culturales (Giménez) (ver gráfico 4).

Gráfico 4: Multiculturalismo.



Fuente: Elaboración propia.

Es extraño encontrar un área de intervención humana donde no se pueda abordar un mismo problema desde diferentes estrategias. Ocurre con la promoción de emprendedores, con el incremento de la renta, con la mejora de la calidad de vida, con la mejora de las competencias técnicas y educativas, en atención sanitaria, en la reducción del crimen, en la inserción de excluidos, etc. En prevención de drogodependencias, por ejemplo, Petratis, Flay y Miller (Becoña: 76) encuentran 14 formas diferentes de prevenir el consumo (ver tabla 5).

Tabla 5. Las 14 teorías que consideran Petratis, Flay y Miller (1995) para explicar la experimentación con el uso de sustancias.

Teorías cognitivo-afectivas	1. Teoría de la acción razonada. 2. Teoría de la conducta planificada.
Teorías de aprendizaje social	3. Teoría del aprendizaje social. 4. Teoría del aprendizaje social/cognitiva social.
Teorías del apego social	5. Teoría del control social. 6. Modelo de desarrollo social.
Teorías donde las características interpersonales juegan un papel esencial	7. Modelo de ecología social. 8. Teoría del auto-desprecio. 9. Modelo de varias etapas de aprendizaje social.

	10. Teoría de la interacción familiar.
Teorías que integran constructos cognitivos-afectivos, de aprendizaje, compromiso y apego, en intrapersonales	11. Teoría de la conducta/problema. 12. Teoría de cluster de iguales. 13. Modelo de vulnerabilidad de Sher. 14. Modelo de Dominio.

De cada una de estas propuestas teóricas se podrían extraer al menos un programa de prevención con el mismo objetivo final y con diferentes objetivos intermedios. Por tanto, cuando se está frente un programa se está ante una determinada elección teórica entre una variedad de posibilidades. A esto es a lo que Carol Weiss llama teoría del programa (Weiss, 55), la asunción teórica por la que opta un programa para lograr el objetivo final de una determinada intervención.

Limitaciones de la dimensión objetivos/resultados:

La perspectiva de resultados presenta algunas limitaciones. No cabe duda de que conocer el grado de cumplimiento de los objetivos tiene un gran valor. Así, los que tienen que tomar las decisiones acerca del proyecto pueden saber si éste ha tenido éxito o no. Pero aparte de saber si el proyecto está o no está funcionando no se puede saber nada más.

En el caso de que se obtengan malos resultados, se desconoce si estos son producidos por una mala praxis, por una inadecuada formación del personal, por unos recursos insuficientes o por una mala gestión, entre otros posibles motivos. Al evaluar resultados se obtiene información sobre los resultados sin preocuparse por el proceso de transformación interno que los provoca. Este tipo de evaluación se denomina de caja negra (Chen, 18).

Con la evaluación de resultados no se pueden tomar decisiones de mejora porque no se sabe cuál es el problema. Esta es la crítica que hacen Weiss y Rien a este tipo de evaluación (Ballart, 329). Stufflebleman y Shinkfield definen las evaluaciones de objetivos como cuasi-evaluaciones ya que no llegan a comprender la totalidad de la evaluanda y emiten un juicio sobre el todo analizando sólo una parte.

Para superar estas críticas, las evaluaciones tienen que ampliar la perspectiva con la que abordan el objeto y empezar a mirar en el interior de la caja negra. Una de las propuestas para iluminar la caja es el análisis de los procesos.

3.1.2. Dimensión de procesos o mecanismos de implementación

Le dijo la cigarra a la hormiga: “amiga mía, todavía no entiendo cómo hago para pasar todos los inviernos hambre y frío mientras tú siempre tienes comida y calor”.

En la caja negra está todo lo que hace y ocurre en el programa; son acciones, decisiones, recursos, personas e interacciones. Se espera que la combinación de todos estos factores produzca los resultados previstos. Realmente los programas son esto (recursos, personas, interacciones, acciones, etc.) ya que los resultados son meramente las salidas, que incluso se puede llegar a conceptualizar como algo externo porque es lo que “sale” de los límites de la intervención.

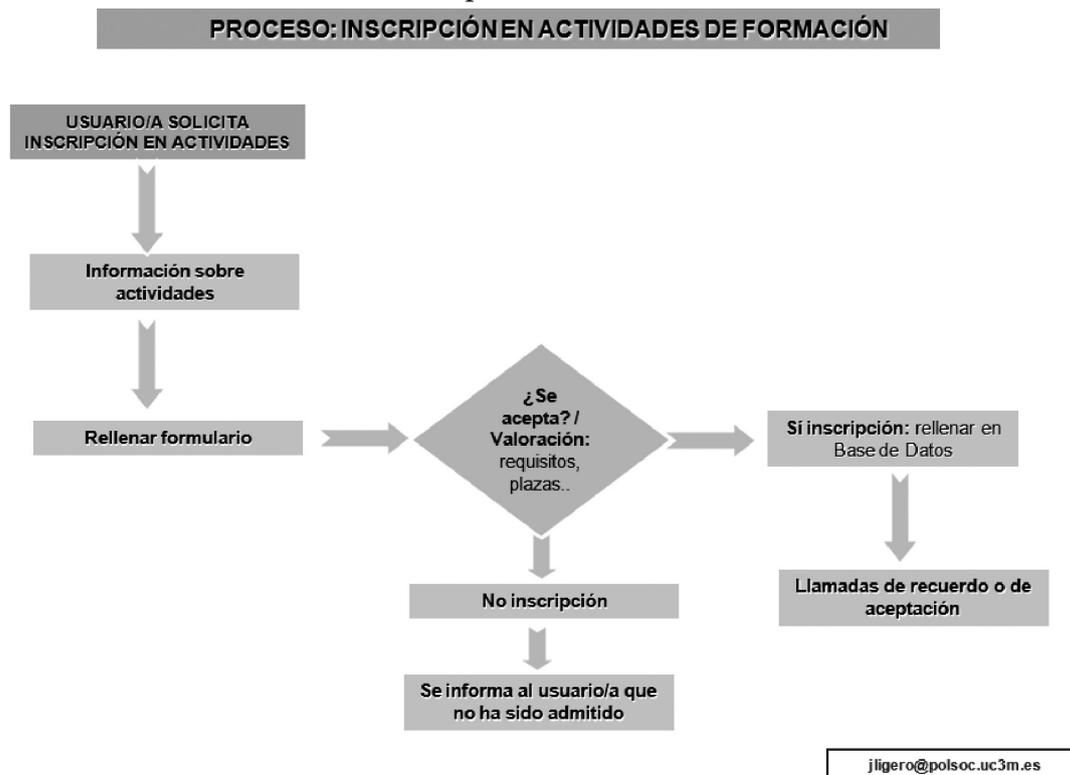
Si la perspectiva de resultados se fija en estas “salidas”, la de implementación se centra en los mecanismos internos de producción, en cómo suceden las cosas. Una de las formas más comunes para trabajar con todo lo que tiene que ver con la implementación es a través del análisis de los procesos.

Un proceso es una secuencia de actividades que genera un valor para los beneficiarios respecto a una determinada situación inicial. A través de una cadena de trabajo se produce algo, ya sea un producto, un servicio, una orientación o un cambio en la población que antes no existía. Por ejemplo, una mesa ha sido generada por una secuencia de actividades que han transformado las materias primas (madera e ideas) en el producto final (mesa); lo que aporta un nuevo valor al usuario es poder escribir sobre ella o apoyar una taza de té. Si una secuencia de actividades no aporta nada no se está ante un proceso sino ante un mero compendio de acciones.

El término proceso lleva implícita la idea de sucesión, secuencia, eje temporal que organiza una serie de eventos hasta la consecución de un fin¹⁰. En la perspectiva de resultados no se requiere específicamente de una cronología, se programa lógicamente pensando en qué es lo que se necesita tener o conseguir para lograr una meta superior. En cambio, con la perspectiva de procesos tiene una importancia crucial el orden en el que suceden las cosas. Hay que especificar cómo aparecen las acciones y cómo se vinculan entre sí.

Se pueden graficar los procesos con un esquema denominado diagrama de flujo o flujograma, donde las diferentes acciones se visualizan como un flujo continuado. En el gráfico 6 se muestra un esquema sintético de uno de los procesos de un proyecto de formación (proceso de inscripción en las actividades).

Gráfico 6. Proceso: inscripción en actividades de formación.



Fuente: Elaboración propia.

¹⁰ Un proceso (del latín *processus*). Conjunto de fases sucesivas de un fenómeno natural o de una operación artificial (RAE, 2001).

Con los procesos, igual que con los resultados, ocurre que no hay una única forma de hacer las cosas. Habrá unos programas que planteen una forma de llegar a los fines y puede ser que otros lleguen a la misma meta de forma totalmente diferente. La existencia de diferentes opciones resalta la idea de que se está ante una determinada elección teórica; en este caso no de logros finales sino de métodos de actuación.

La elección escogida por el programa es lo que Weiss denomina teoría de la implementación (Weiss, 58) que se puede definir como la cadena de asunciones teóricas que vinculan la acción desarrollada por el programa o el servicio con el logro final de los resultados.

Programación por objetivos y por procesos

La principal diferencia entre resultados e implementación es que el primero describe cambios en la población o en el entorno, mientras que el segundo describe las cosas que hace el programa para conseguirlos. Por ejemplo, en un programa formativo el objetivo es que una población adquiera unos conocimientos. La implementación es todo aquello que se hace para lograrlo como puede ser dar clase. Dar clase no es lo mismo que adquirir formación, se pretende que una cause la otra pero son aspectos distintos. Esta es la diferencia básica entre implementación y resultados, lo que el programa hace y lo que el programa consigue.

Veamos otro ejemplo, en un máster de evaluación de programas los resultados son cambios en la población en sus dos niveles (ver gráfico 7).

Gráfico 7. Ejemplo de objetivos de un máster en Evaluación de Programas y Políticas Públicas.



Fuente: Elaboración propia.

Desde la perspectiva de resultados el gráfico en sí mismo es completo (se podría añadir un nivel más que especificara las actividades e incluso un cronograma). Pero esta forma de programar no está especialmente pensada para recoger todas las secuencias de actividades necesarias para el funcionamiento del máster.

En la siguiente tabla 6 se muestra un “mapa de procesos” en el que se enuncian los principales procesos de la organización.

Tabla 6. Ejemplo de mapa de procesos
(Mapa de procesos del Magíster en Evaluación de Programas y Políticas Públicas; elaborado a partir de la Memoria de Grado de Irene Rosales).

Procesos estratégicos	Definición posicionamiento estratégico. Definición operativa del programa del Magíster. Organización básica.
Procesos clave u operativos	1. Difusión del Magíster. 2. Proceso Educativo <ul style="list-style-type: none"> a) Selección de alumnos/as. b) Programa de sesiones. c) Docencia. d) Trabajo Intersesiones (trabajos de alumnos/as) e) Memoria de Grado. f) Prácticas en empresas g) Trato con el alumnado
Procesos de soporte o apoyo	<ul style="list-style-type: none"> - Selección de coordinación administrativa. - Secretaría: preinscripción y matrícula. - Administración: facturación. - Gerencia: economía y presupuestos. - Gestión de aprovisionamiento de viajes de profesores/as. - Gestión académica: calificación del alumnado (Actas); control de asistencia; evaluaciones de las sesiones. - Gestión de aulas. - Gestión de Convenios de Prácticas del alumnado. - Gestión de recursos Web.

Aunque los procesos se quisieran incluir como listados de actividades dentro de la matriz de objetivos, esto no sería posible porque muchos de ellos son transversales y atañen a varios objetivos a la vez. Además, tampoco permitirían incorporar las divisiones de flujos en dos o más caminos tan necesarias en los diagramación de procesos.

Tanto el enfoque de resultados como el de procesos aisladamente no cubren toda la realidad de una intervención, son complementarios. La programación conjunta de objetivos y procesos incrementa la congruencia del modelo teórico. La vinculación entre procesos y resultados es probablemente una de las contribuciones más importantes a la teoría de la evaluación (Greene; Shadish y otros) ya que busca la articulación causal entre lo que hace el programa y lo que consigue.

Con la especificación de la lógica de la implementación los resultados ya no brotan de una caja negra sino que son el producto de diferentes secuencias de actividades que estarán plasmadas en los diagramas de flujos.

La combinación de la teoría de la implementación y la teoría del programa conforman la teoría del cambio según Weiss. Rossi la denomina evaluación comprensiva (*comprehensive evaluation*¹¹) ya que pretende abarcar un espectro amplio de la realidad del programa. Con relación a la lógica que subyace identifica tres tipos relaciones causales:

¹¹ Se ha traducido *comprehensive* como comprensivo, ya que el término castellano “comprender” incluye la acepción de abrazar, ceñir, rodear por todas partes algo (RAE).

- 1) **Hipótesis causal:** son las relaciones entre los problemas a solucionar con las causas que los han creado.
- 2) **Hipótesis de la intervención:** establece la relación entre la acción del programa con los cambios en las causas.
- 3) **Hipótesis de la acción:** donde se vincula el programa con los cambios en el problema original (Fitzpatrick).

Críticas a la perspectiva de procesos

Se destacan cuatro bloques de críticas relacionadas con la programación, gestión o evaluación a través de procesos:

- La preocupación por ejecutar los procedimientos correctamente puede hacer perder el foco sobre los objetivos finales de toda la intervención. Intentando prevenir este problema, algunas iniciativas administrativas insisten en la necesidad de que los procesos tienen que estar orientados a los resultados, prueba de que esto no tiene porqué ser lo normal.
- A la gestión por procesos se le asocia otro problema, la pérdida de responsabilización de los profesionales. En un proceso se corre el riesgo de que cada actor se ocupe de su tarea asignada sin que se sientan responsables del producto final. A cada cual se le juzgaría por haber hecho bien su parte, aunque todo el programa fuera ineficaz nadie se sentiría directamente interpelado. Realmente esta crítica guarda similitud con el concepto de alienación hegeliano o marxista, donde lo producido es enajenado de su agente creador.
- Otro riesgo de la programación por procesos es que se caiga en un excesivo, pormenorizado e inoperante detalle de los procedimientos. Se pueden crear procesos cada vez más complejos que pierden la flexibilidad necesaria para responder a las diversas contingencias que van apareciendo. En esta crítica se ha apoyado uno de los principales argumentos a favor de la gestión por objetivos. La óptica objetivos/resultados da toda la libertad a quien implementa (precisamente porque no define nada sobre cómo actuar). No obstante, las propuestas de modernización de la administración pública son conscientes de la tendencia a la complicación de los procedimientos de gestión e intentan buscar la simplificación de la acción.
- Otro de los elementos a tener en cuenta en el análisis de procesos es que no son abstracciones que se ejecutan en el aire, se requiere de unos recursos para poder implementarlos: trabajadores, alianzas, medios, materiales, etc., que no necesariamente se recogen en los procesos. Todos los elementos necesarios para la puesta en funcionamiento de un servicio o programa es lo que se puede denominar elementos estructurales.

3.1.3. Dimensión estructural. Los recursos, la organización y otros elementos estructurales

Por estructura se puede entender “la organización relativamente estable de distintos tipos de recursos para alcanzar los fines del proyecto” (Cohen y Franco, 94). Bajo esta definición se pueden englobar múltiples aspectos de un programa pero todos ellos son elementos que se necesitan para que los procesos puedan funcionar. Se han escogido tres operacionalizaciones de la estructura que pueden resultar esclarecedoras para entender el concepto:

La primera de ellas son las claves enunciadas por Weiss para poder identificar las cuestiones relevantes de los “*inputs*, recursos y entorno del programa” (Weiss, 193), a saber:

- El auspicio organizacional.
- El presupuesto.
- Naturaleza del *staff*.
- La gestión.
- La localización.
- El plan de actividades.
- Los métodos de prestar el servicio.
- La filosofía o principios de operación.
- Los propósitos.

El segundo enfoque que se presenta es el modelo de excelencia TQM de EFQM¹² (MAP). En esta propuesta se definen una serie de elementos previos a la acción que se pueden considerar estructurales, reciben el nombre de agentes facilitadores (en lo que también estarían incluidos los procesos):

- Liderazgo: desarrollo de valores y sistemas para el logro del éxito. Los líderes son capaces de reorientar a la organización en situaciones de cambios y arrastrar tras ellos al resto de personas.
- Política y estrategia: implantación de la misión, visión y el desarrollo de una estrategia centrada en los grupos de interés, teniendo en cuenta el mercado y el sector donde operan.
- Personas: los integrantes de la organización, ya sean individuos o grupos.
- Alianzas y recursos: planificación y gestión de las alianzas externas con otras entidades, con los proveedores y los recursos internos; todo ello al servicio de la política y estrategia.

El tercer enfoque es una propuesta de evaluación institucional (Lusthaus, Anderson y Murphy) que señala como aspectos clave de la capacidad institucional los siguientes elementos:

- Estrategia de liderazgo.
- Recursos humanos.

¹² European Foundation for Quality Management.

- Otros recursos clave (infraestructura, tecnología y finanzas).
- Gestión del programa.
- Gestión de procesos.
- Alianza interinstitucionales.

Todos estos elementos tienen en común el ser condiciones previas y necesarias para la ejecución de los procesos. El peligro del análisis estructural es que se convierta en un repaso de componentes sin identificar aquellos elementos que realmente son relevantes para la marcha del proyecto. Al analizar la estructura debe buscarse la vinculación entre los elementos estructurales y la ejecución de los procesos lo que producirá, si la teoría es correcta, los resultados esperados por la organización.

Con la estructura ocurre como con los procesos y los resultados, se puede programar, analizar y evaluar de forma aislada e independiente a las demás dimensiones de la intervención. Pero el análisis se enriquece si se incorpora como una tercera dimensión a la evaluación de resultados y de procesos. La vinculación de las tres dimensiones responde a una visión mecanicista de la organización donde todos los elementos funcionan como piezas que conforman un gran sistema para la producción de resultados.

3.1.4. La perspectiva sistémica

La vinculación entre procesos y resultados o la integración de diversas dimensiones de un programa, lo que Rossi ha llamado evaluación comprensiva, es probablemente “el avance más evolucionado y sofisticado en evaluación hasta la fecha” (Greene, 164).

El enfoque sistémico entiende el programa o servicio como un conjunto de elementos articulados e interdependientes entre sí que conforman un sistema. Las relaciones medios-fines recorren el sistema desde su inicio hasta los resultados finales implicando a todos los elementos de la intervención. La variación en uno de los elementos afecta al resto de los componentes. Por ejemplo, la reducción de los recursos asignados a un programa tendrá repercusiones en la ejecución de los procesos y, por lo tanto, en los resultados finales obtenidos. Para el buen funcionamiento del programa se requiere del funcionamiento correcto de todos sus elementos.

La especificación de las dimensiones de una intervención y las relaciones entre sí se denomina teoría de la acción, del programa o también modelo lógico. Se han desarrollado variados sistemas para identificar los modelos lógicos de las intervenciones. Por citar algunas propuestas, en 1966 Donabedian desarrolló un concepto integral de la calidad hospitalaria. Considera que “está influida por diversos elementos relacionados armónicamente y los agrupa en tres niveles: estructura física de la organización, proceso desarrollado para prestar a la asistencia y el resultado obtenido” (Martínez del Olmo, 274). Stufflebeam y colaboradores diferenciaron entre contexto, inputs, procesos y productos (CIPP¹³). Con estas dimensiones pretenden abarcar toda la secuencia clave para el desarrollo de una intervención (Stufflebeam y Shinkfifeld). En 1967 Robert Stake definió un modelo que proponía recoger información sobre los antecedentes, las transacciones y los resultados del programa para poder comprender y establecer juicios.

¹³ Context, Inputs, Process and Products.

Una de las propuestas más referenciadas actualmente es la de la W.K. Kellogg Foundation, las dimensiones básicas de análisis propuestas se resumen en el gráfico 8.

Gráfico 8: El Modelo Lógico.



jliger@polsoc.uc3m.es

Fuente: elaboración propia a partir del gráfico incluido en W.K. Kellogg, figura 1: The Basic Logic Modelo.

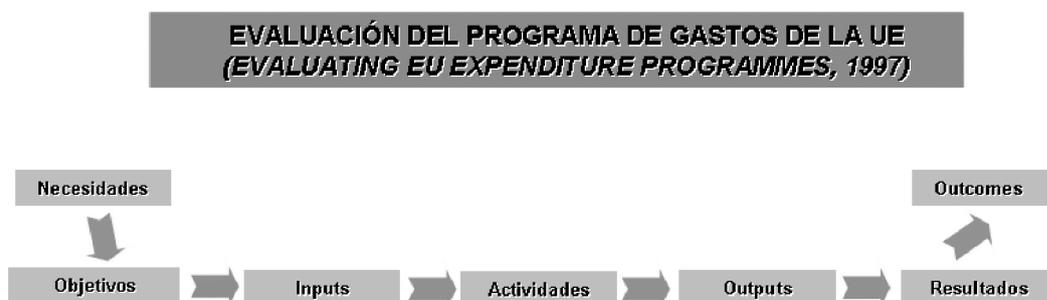
Otras propuestas teóricas como la de Klages y Haubner (Löffler, 31) discriminan entre resultados “subjeto” y “objetivo” manteniendo la cadena causal (ver gráfico 9).

Gráfico 9: Propuesta de Klages y Hauber.



Fuente: elaboración propia a partir de Löffler.

Gráfico 10: Adaptación de Jacques Plante.

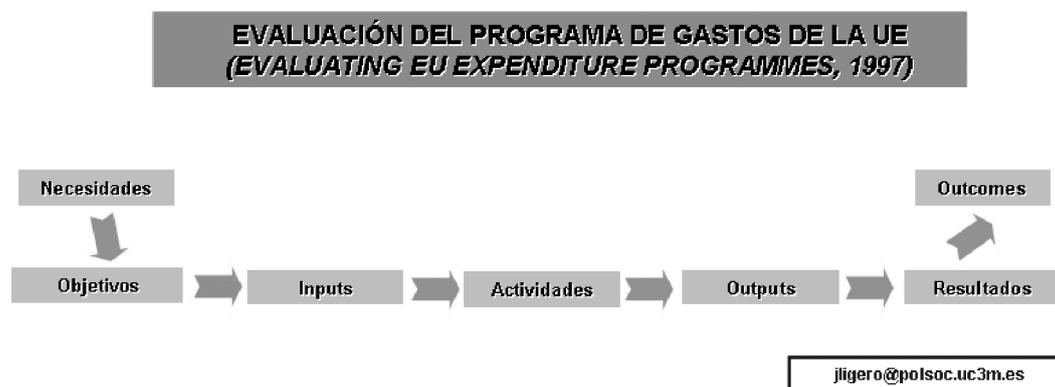


jliger@polsoc.uc3m.es

Fuente: elaboración propia a partir de Osuna y Márquez.

La Comisión Europea trabaja con varias propuestas de modelos lógicos, como con la anterior de Plante (ver gráfico 10), aunque en la guía *Evaluating EU Expenditure Programmes*, Nagarajan y Vanheukelen formulan el siguiente esquema (ver gráfico 11).

Gráfico 11. Evaluación del Programa de Gastos de la UE.

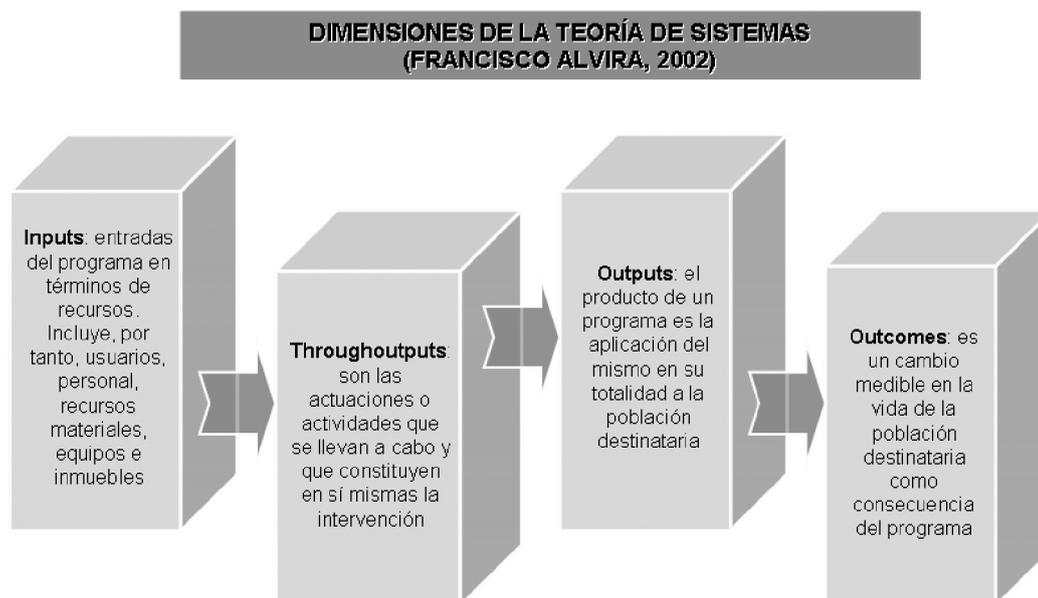


Fuente: elaboración propia a partir de Evaluating EU Expenditure Programmes.

El Observatorio Europeo de drogas y toxicomanías propone utilizar el modelo lógico propuesto por la Universidad de Wisconsin que diferencia entre situación, *inputs*, *outputs* y *outcomes* (www.uwex.edu).

En la literatura española se encuentran varias propuestas como la de Paloma López de Ceballos y otros en la que se diferencian entre finalidades, estructuras, métodos y entornos. Francisco Alvira expone cuatro dimensiones utilizando la terminología de la teoría de sistemas (ver gráfico 12).

Gráfico 12: Dimensiones de la Teoría de Sistemas.



Fuente: Elaboración propia a partir de Alvira (2002).

Todas estas propuestas son sólo una muestra de la profusión de estrategias para identificar los modelos lógicos. Lo curioso es que varían poco en la conceptualización de lo que debe tenerse en cuenta en un programa u organización. Todas las visiones tienen mucho en común, lo que evidencia la existencia de un paradigma dominante sistémico en el análisis de las organizaciones (Blas, 74). En la práctica, la gran mayoría de las instituciones se basan en un modelo causal genérico de inputs-process-outputs-outcomes (Rogers y otros, 6).

Se haya hecho explícito o no, cualquier programa está proponiendo un determinado modelo teórico de cómo cambiar la realidad. En la medida en que la lógica en la que se apoya una intervención es una teoría, pueden existir otras alternativas, y todas ellas pueden ser validadas. Tal como dice Rein, un programa es una teoría y la evaluación su test (Weiss, 55).

La evaluación orientada por el modelo lógico dirige las reflexiones de mejora hacia los pasos concretos en la cadena causal. Lo que permite conocer y consolidar cadenas causales exitosas, construir conocimiento sobre la intervención o cambiar aquellos supuestos tanto de la implementación como de los resultados que no están funcionando.

3.2. Múltiples perspectivas sobre las teorías que subyacen en una intervención

Todo este esquema funcionaría muy bien si existiera una única teoría del programa, tal como defiende el enfoque objetivista o también llamado aproximación científica (Chen). La idea básica de los “objetivistas” es que la teoría de la intervención es una y ajena a la visión y percepción de los diferentes actores implicados. Pero hay varios aspectos incómodos en esta aproximación, el primero es que la teoría del programa declarada (a través de la programación o la documentación oficial) en muchas ocasiones no suele estar establecida en términos racionales ni tampoco está uniformemente aceptada por los diferentes actores. Pueden existir otras teorías, algunas de ellas explícitas pero muchas implícitas (Weiss, 55). El segundo aspecto es que puede haber una distancia entre lo programado y lo real, o dicho de otra forma, entre lo que se dice y lo que se hace.

El hecho que haya una teoría explícita no implica que sea la versión objetiva. Como ya se ha comentado, la documentación suele representar la visión normativa. Monnier postula que para formular las hipótesis relativas a los efectos producidos por un programa, lo que denominamos una teoría de acción, es preciso adoptar un determinado punto de vista (Monnier, 142).

Por tanto, el concurso de los diferentes actores implicados es un elemento crucial en la construcción de teoría o teorías para que sean fieles a la situación real.

3.2.1. Fuentes para obtener pistas y sugerencias para construir la teoría

- **Realización de entrevistas a actores clave:** gerentes y grupos de interés, extrayendo las presunciones, esquemas y expectativas con relación al programa. En esto coinciden la gran parte de los autores.
- **Revisión de la documentación del programa:** Wholey recomienda tener en cuenta los documentos del programa, programación, legislación, ordenanzas internas, guías, justificaciones presupuestarias e informes de seguimiento o informes de logros (Chen, 66). En verdad, la documentación escrita se considera la plasmación de un determinado punto de vista, el institucional.
- **Observar el programa:** la observación puede ir dirigida a ver dónde realmente se gastan las energías (Weiss, 55; Lipsey y Pollar cit. Rogers y otros, 7). Se aconseja prestar más atención a lo ejecutado que a lo planeado.

- **Revisión de la literatura en investigación y evaluación:** la comprensión del fenómeno o de otras evaluaciones puede ayudar a entender la teoría.
- **Ideas provenientes del propio equipo evaluador:** diversos autores conceden legitimidad al equipo para proponer aspectos que haya identificado acerca de la teoría (Rogers y otros, 7).

3.3. Modelo lógico con perspectiva sistémica

La información recogida de las diferentes fuentes puede organizarse bajo algún modelo lógico. En el apartado “perspectiva sistémica” se han visto algunos ejemplos. Wholey propone utilizar un esquema con los siguientes elementos: recursos, actividades, resultados y establecer las relaciones entre ellos, a lo que llama “diseño del programa”.

Rossi y Chen proponen partir de los documentos del programa como la vía principal para detectar la teoría del programa confiando en los conocimientos y la experiencia de la evaluadora. Una vez elaborada esta primera aproximación la contrastan con los puntos de vista de los actores clave. Si se dan discrepancias propician una negociación entre los diferentes actores. La finalidad es lograr un consenso, “el entendimiento natural es mejor que lo obtenido por solo una parte” (Chen, 71).

Una forma de proceder es que las evaluadoras pueden construir uno o varios modelos lógicos tentativos que se discuten con los actores clave esperando encontrar confirmaciones o diferencias. A esta combinación de la propuesta “científica de la evaluadora” y las aportaciones “subjetivas de los actores” le han llamado enfoque integrado.

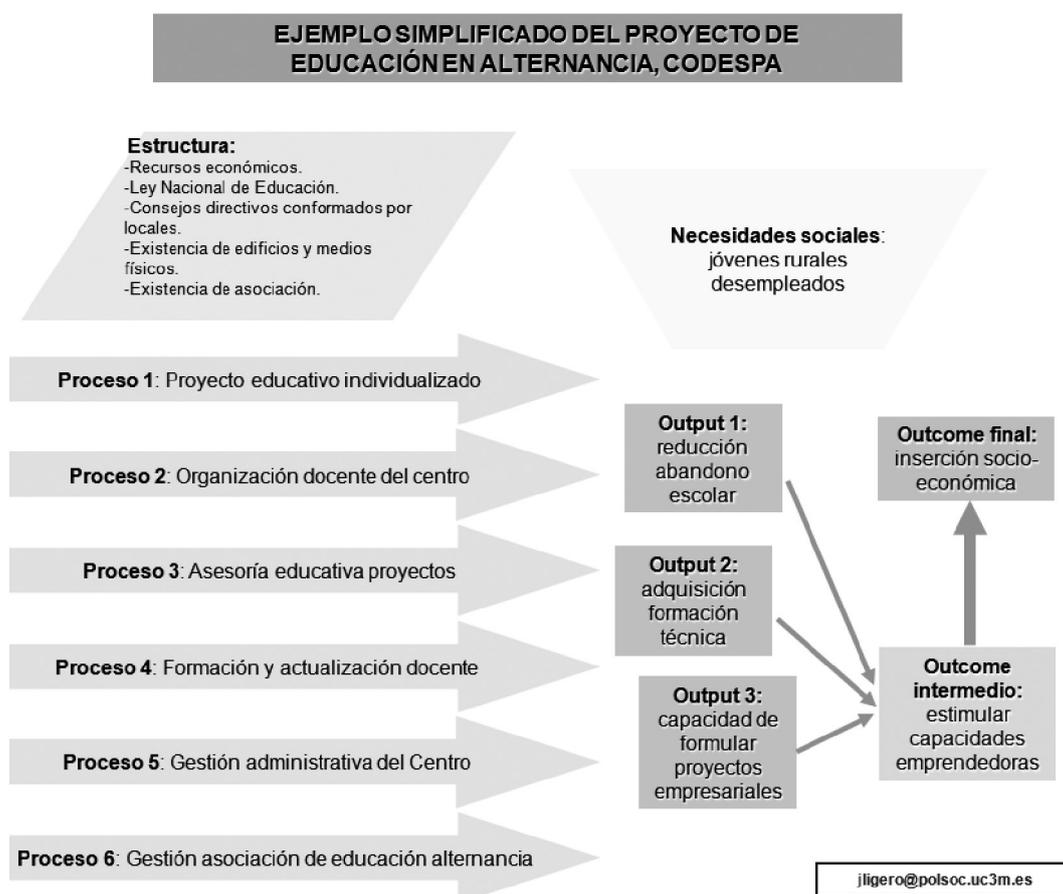
Cuando los diferentes actores están de acuerdo sobre un aspecto, no hay mayor problema, esa dimensión del modelo queda confirmada. Pero en el caso de que haya discrepancias hay que buscar consensos o al menos saber con claridad en qué no se está de acuerdo.

Wholey argumenta que el principal propósito de la construcción de la teoría es clarificar y desarrollar las teorías subyacentes de los diferentes actores implicados aunque éstas sean diametralmente opuestas. Las evaluadoras no deberían confiar en su propio conocimiento o experiencia para construir modelos. La teoría tiene que emerger de los datos más que de estructuras o hipótesis previas tanto de las evaluadoras como de otros actores (Chen, 67; Weiss, 76).

3.4. Graficar la teoría

Después de que las diferencias encontradas sean resueltas, pactadas o explicitadas, se completa formalmente una teoría o teorías del programa. Los diferentes componentes se presentan en un gráfico que muestre las cadenas causales (Rogers y otros, 7). Tal como citan estos mismos autores se pueden dibujar a través de una serie de “cajas” que representen las dimensiones de *inputs*, procesos, *outputs and outcomes* con flechas que las conecten entre sí. No es necesario especificar qué proceso lleva a qué *output*. Las relaciones entre los diferentes componentes serán exploradas en análisis posteriores. En el gráfico 13 se muestra un ejemplo simplificado de modelo lógico para un proyecto de educación en alternancia de la fundación CODESPA.

Gráfico 13. Ejemplo simplificado del proyecto Educación en Alternancia, CODESPA.



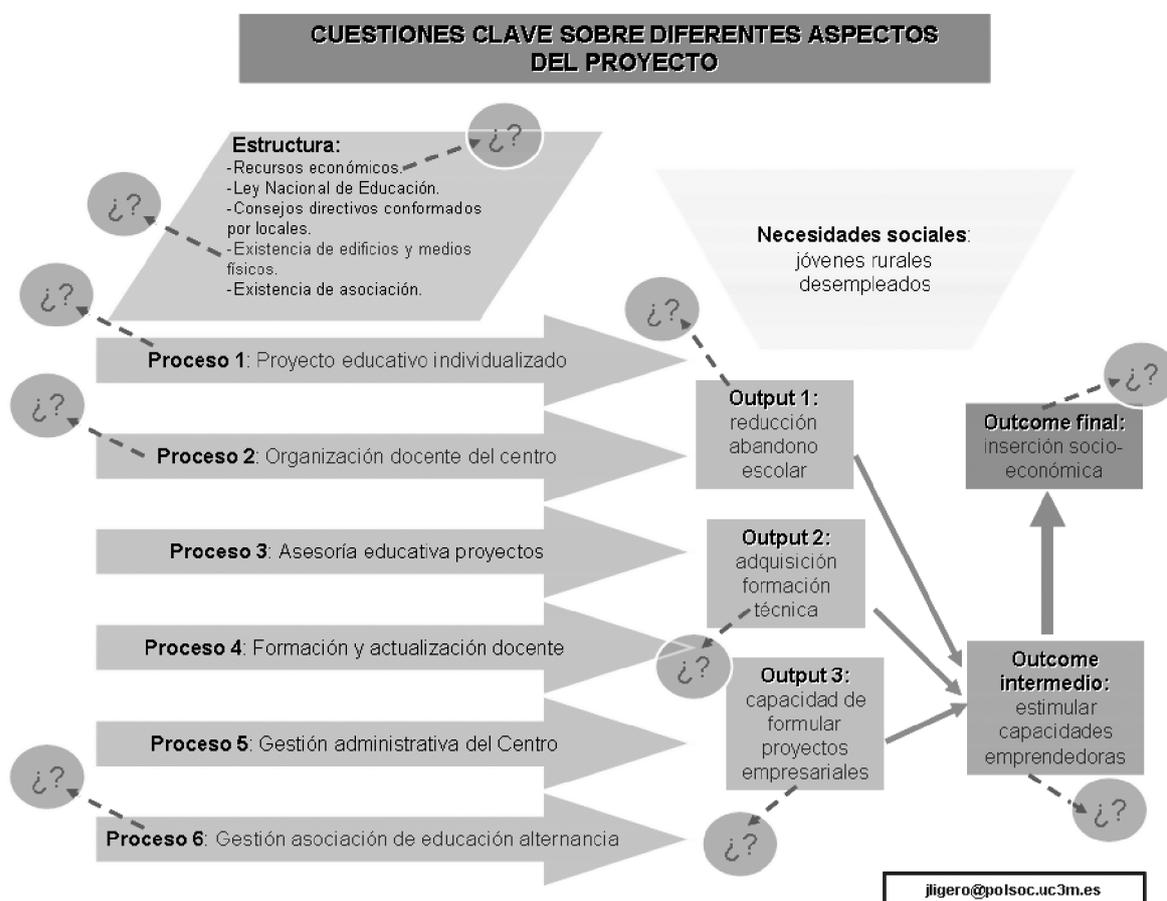
Fuente: Elaboración propia.

3.5. Determinación de los aspectos a evaluar

El modelo obtenido, la teoría, es usada como un marco para “determinar las variables que deben ser medidas en la evaluación” (Davidson, 18). Por tanto se pueden identificar de cada dimensión (resultados, procesos y estructura) los aspectos clave de la lógica que podrían ser tenidos en cuenta y medidos.

Las cuestiones deben cubrir toda la lógica de la intervención para después poder comprobar las relaciones causales que asume el programa. Si no se obtiene información de los aspectos clave se corre el riesgo de no poder analizar la congruencia de la teoría.

Gráfico 14: Cuestiones clave sobre diferentes aspectos del proyecto.



Fuente: Elaboración propia.

Carol Weiss (Weiss, 75) concreta una propuesta con cinco categorías de indagación:

- Cuestiones sobre los resultados finales:** para identificar los resultados una fuente de entrada es la “declaración de objetivos” pero existen otras fuentes (*staff*, clientes, etc.) además de procurar captar los objetivos no previstos. Davidson (Davidson, 22) aconseja la búsqueda libre de objetivos para procurar recoger todos los cambios producidos por el programa. No hay que olvidar que por muy bienintencionada que sea una acción puede no tener efectos o incluso que éstos sean contraproducentes.
- Cuestiones sobre los procesos:** identificar los aspectos más importantes de los procesos y fases clave. Además de saber si se están ejecutando de acuerdo a lo establecido.
- Indagaciones acerca de si los resultados son atribuibles al programa.**
- Indagación sobre si existen conexiones entre los procesos y los resultados.**
- Cuestiones que facilitan las explicaciones sobre diferentes aspectos:** no sólo se trata de saber qué es lo que ha pasado sino cómo y por qué. También si los acontecimientos y resultados han sido producidos por las relaciones causales de la teoría del cambio.

3.6. Métodos y técnicas

Igual que ocurriría en el modelo por criterios, las cuestiones tienen que ser contestadas a través de indicadores con sus respectivas fuentes de recogida de información (salvo en el caso de las técnicas cualitativas, que no requieren de indicadores). La evaluación basada en la teoría no tiene porque estar adscrita a ningún tipo de técnica, puede ser abordada a través de técnicas cualitativas y cuantitativas. La combinación de técnicas y métodos es algo ya totalmente admitido por la práctica y la teoría de investigación (Serrano y otros). Davidson (Davidson, 22) propone que las fuentes o personas a entrevistar no sean las mismas que las que han participado en los pasos previos de construcción del modelo.

Una vez que se tiene la información de las diferentes dimensiones la finalidad es estudiar las relaciones causa – efecto entre los *inputs*, los procesos y los resultados (Shadish y otros, 389). Es un caso específico de evaluación comprensiva tal como define Rossi, ya que busca articular e integrar las diferentes dimensiones del programa.

Uno de los métodos más extendidos para ver las relaciones entre las dimensiones es la modelización estadística a través de las ecuaciones lineales. Las correlaciones entre los diferentes elementos, controlando las posibles variables intervinientes, se utilizan como estimador de las relaciones causales. Es decir, la variación en los indicadores de la implementación se relaciona con la variación en los *outcomes* o las variaciones en los elementos estructurales con las variaciones con los procesos, asumiendo que si esto es así hay una relación causal entre dichas dimensiones.

Uno de los problemas de los sistemas de ecuaciones lineales es que en la práctica es una técnica muy exigente. Se necesita información cuantitativa de todas las variables que se quiere que estén en el sistema, además todas deben ser independientes entre sí y cumplir una serie de criterios estadísticos. En el caso que se consiga tener la información y construir modelos con validez interna, en muchas ocasiones las varianzas explicadas suelen ser muy bajas, por lo que pierde capacidad explicativa.

Se pueden construir modelos menos ambiciosos a través de medidas de asociación, como el estadístico chi-cuadrado y tablas de contingencia. La dificultad de esta opción es controlar múltiples variables intervinientes lo que incrementa el riesgo de sesgo.

Los experimentalistas entienden que la correlación o la asociación no son suficientes para establecer una atribución causal. Entre otras razones porque puede haber variables influyentes que se escapan del modelo y porque no hay validación contrafactual que pueda descartar otros modelos causales alternativos (Davidson, 23).

Thomas Cook critica la evaluación orientada por la teoría por su predisposición a utilizar modelizaciones causales y por que puede disuadir en la utilización de modelos experimentales. Aunque autores como Rubin o el mismo Cook creen que es posible evaluar basándose en la teoría pero a través de métodos experimentales (Cook, 34).

Existen otras aproximaciones metodológicas de corte naturalista, principalmente apoyados en técnicas cualitativas. Se pueden reconstruir las teorías a través del discurso, las historias de vida o presentación de casos individuales tal como lo hace Robert Stake. No se busca una generalización de tipo cuantitativo sino ayudar a los diferentes actores a comprender mejor que es lo que ocurre en su programa.

En cualquiera de los casos, el propósito de los métodos es reconstruir “un modelo causal apoyado en varias fuentes para entender cómo una intervención específica (o programa) está intentando lograr sus productos y resultados” (Stame, 2004).

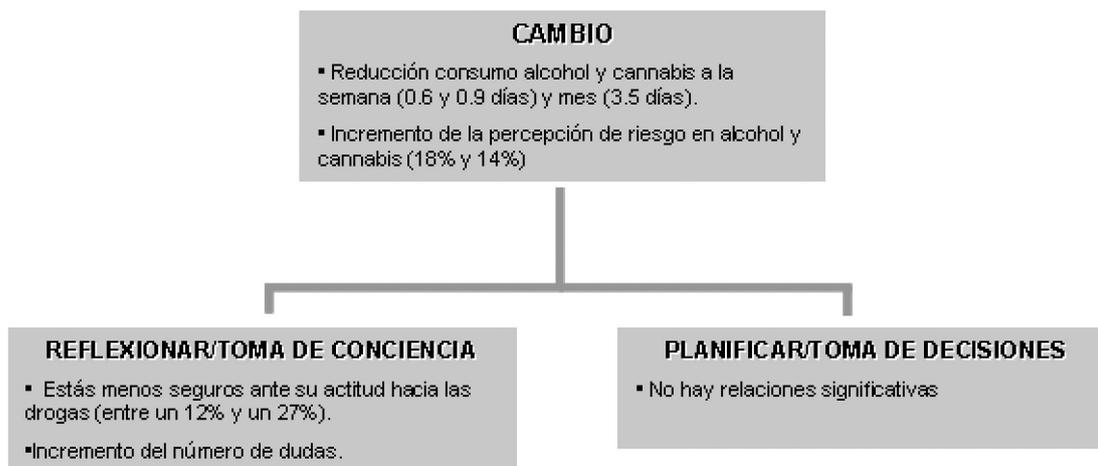
3.7. Análisis

El objetivo de la evaluación de la teoría es comprobar el modelo lógico establecido. Una vez que se tiene la información, se establece una secuencia de análisis que indague sobre si el programa funciona o no y a qué es debido (Rogers y otros, 10). La estructura del análisis guarda relación con la estructura de la indagación y de las preguntas. La secuencia puede ser la siguiente:

a) Análisis de la jerarquía de objetivos

Se estudia el grado de alcance de los objetivos y el grado de correlación existente entre los diferentes niveles o pisos de objetivos. En un programa de prevención de consumo de drogas (Ligero y otros) teníamos dos pisos de objetivos. Obtuvimos resultados para cada uno de los objetivos de forma aislada, tal como se muestra en el gráfico 15:

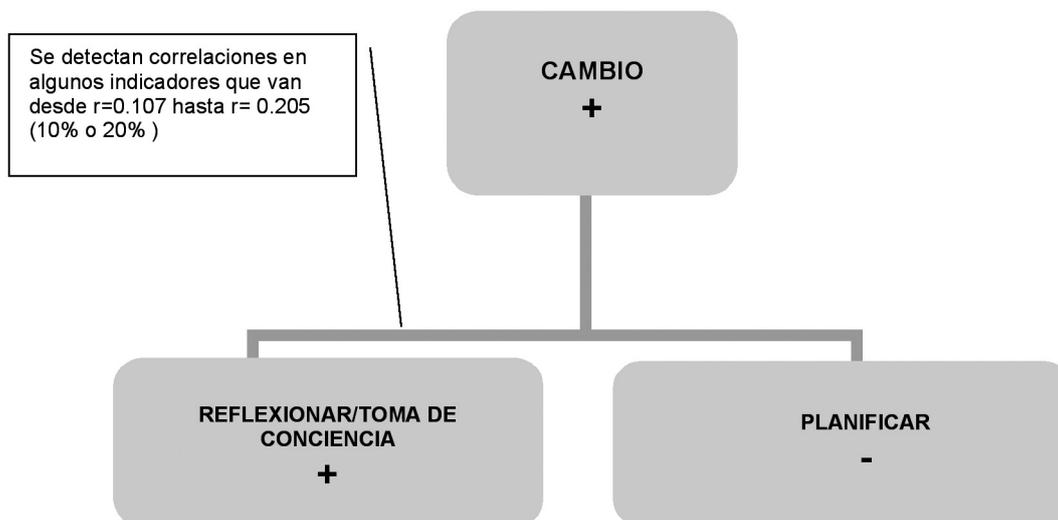
Gráfico 15. Análisis de la Teoría del Programa de un programa de prevención de consumo de drogas (Ligero y otros).



En uno de los objetivos intermedios “Planificar/Toma de decisiones” no se detectaron cambios significativos. En el resto de los objetivos sí.

El siguiente paso fue analizar las relaciones entre los dos niveles de objetivos. En el ejemplo que se está utilizando, se analizó la asociación entre reflexionar/toma de conciencia y cambio con relación al consumo. Resumimos los resultados en el siguiente gráfico 16.

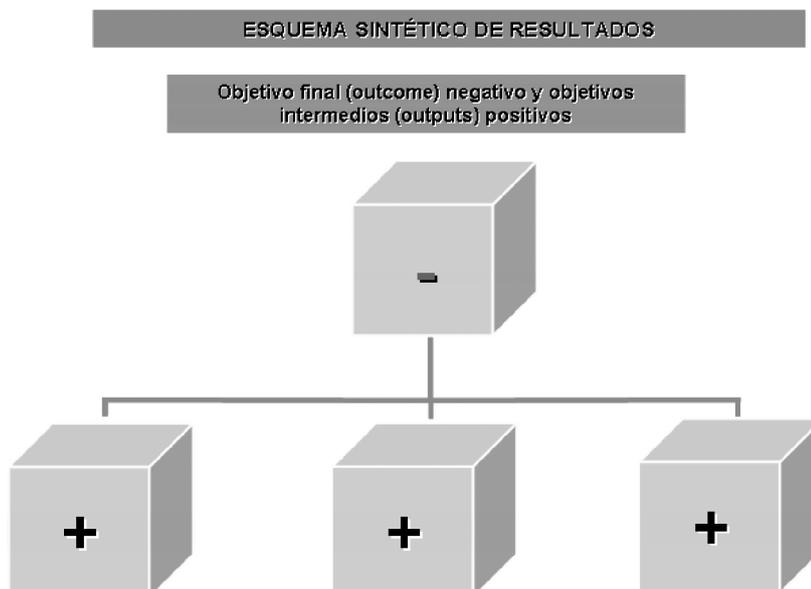
Gráfico 16. Análisis divariados entre indicadores del análisis de un programa de prevención de consumo de drogas (Ligero y otros).



Es decir, se detecta una relación entre el objetivo intermedio “reflexión” y el objetivo final “cambio”: una variación en el objetivo intermedio produce una reducción del consumo de alcohol y cannabis al mes entre un 10% y un 20%¹⁴. En este caso hay indicios que permiten establecer una relación lógica entre el objetivo intermedio y el final.

Pero puede haber otras posibles situaciones. En el gráfico 17 se resumen los resultados de los indicadores, el símbolo negativo es que no se ha alcanzado y el positivo que sí.

Gráfico 18. Esquema sintético de resultados.



Fuente: elaboración propia.

¹⁴ Se realizaron correlaciones entre diversos indicadores controlando algunas variables externas.

Los objetivos intermedios han sido alcanzados pero no hay cambios en el objetivo final. En esta teoría del programa, tal como está planteada, no hay una relación causal (sí hay un control efectivo de otras posibles variables intervinientes). Puede ser que con el concurso de otros objetivos intermedios se logre el cambio, pero tal como está planteado no tiene sentido continuar ejecutando el programa, ya que no hay conexión causal entre los dos pisos.

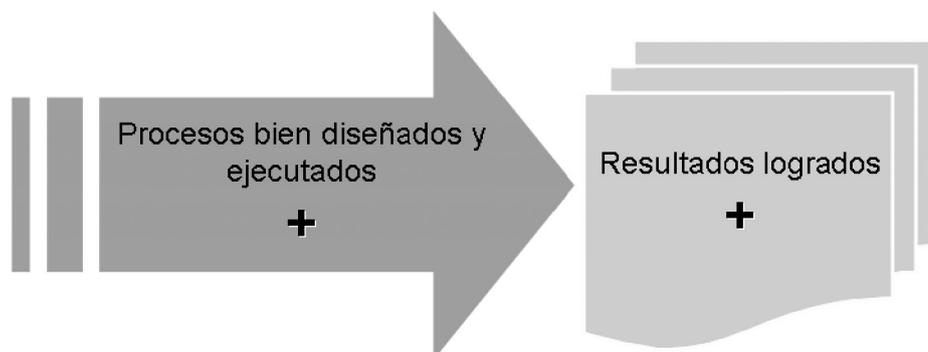
Estos dos ejemplos ilustran sobre las posibilidades del análisis de la jerarquía de objetivos. Se puede testar y validar la teoría en la que se apoya el programa e incluso conocer la fuerza de la relación. Es decir, se podría llegar a saber no sólo cuanto influye el programa en la realidad sino cuánto influye cada uno de los resultados esperados en el logro del objetivo específico.

b) Análisis del cumplimiento de procesos y la relación con los resultados

Lo que se pretende básicamente en este apartado es conocer, tal como argumenta Cronbach, “qué procesos contribuyen a los resultados esperados” (Ballart, 340). Las cosas que hace el programa, los procesos, teóricamente deben conducir al logro de los resultados. Si no se consiguen los resultados (y se han propuesto métodos para desechar que no sea debido a causas externas), hay que pensar que los procesos no se han ejecutado correctamente o que han sido mal diseñados.

La explicitación de las dos dimensiones (resultados y procesos) permite avanzar en el análisis de la congruencia teórica de la intervención. Por ejemplo, en el supuesto de que los procesos se hayan ejecutado correctamente (e insisto, en que se consigan controlar las variables externas) se deberían lograr los resultados previstos si la teoría de la implementación es correcta (ver gráfico 19).

Gráfico 19. Procesos y resultados positivos.

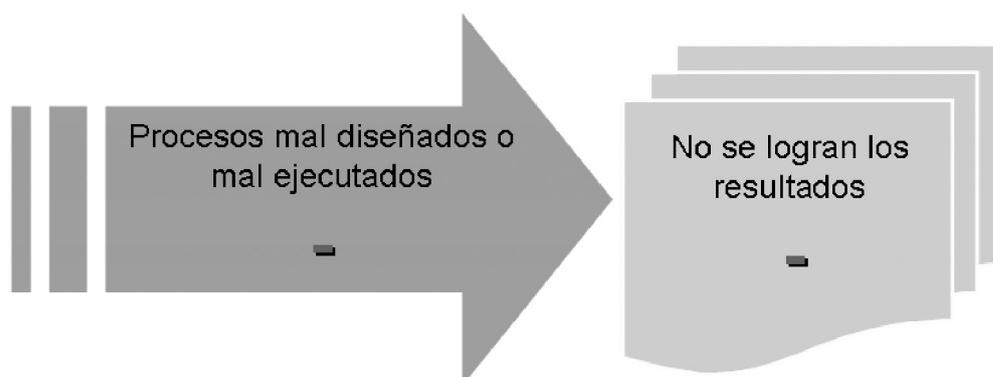


jligero@polsoc.uc3m.es

Fuente: elaboración propia.

Y a la inversa, si los procesos están mal diseñados o se han ejecutado indebidamente es lógico pensar que no se lograrán los resultados (ver gráfico 20).

Gráfico 20. Procesos y resultados negativos.



Fuente: elaboración propia.

Tanto Bickman, Lipsey como Weiss destacan que este tipo de análisis son útiles para distinguir entre los fallos de la teoría del programa (jerarquía de objetivos) o fallos en la implementación (procesos) (Rogers: 8). Vinculando la información de esta manera se podrá saber si el fallo está en la lógica de resultados, en la lógica de procesos o en la implementación. Pero no sólo de una forma genérica, si la información ha sido exhaustiva, se podrá definir en qué fase o fases de los procesos se está fallando.

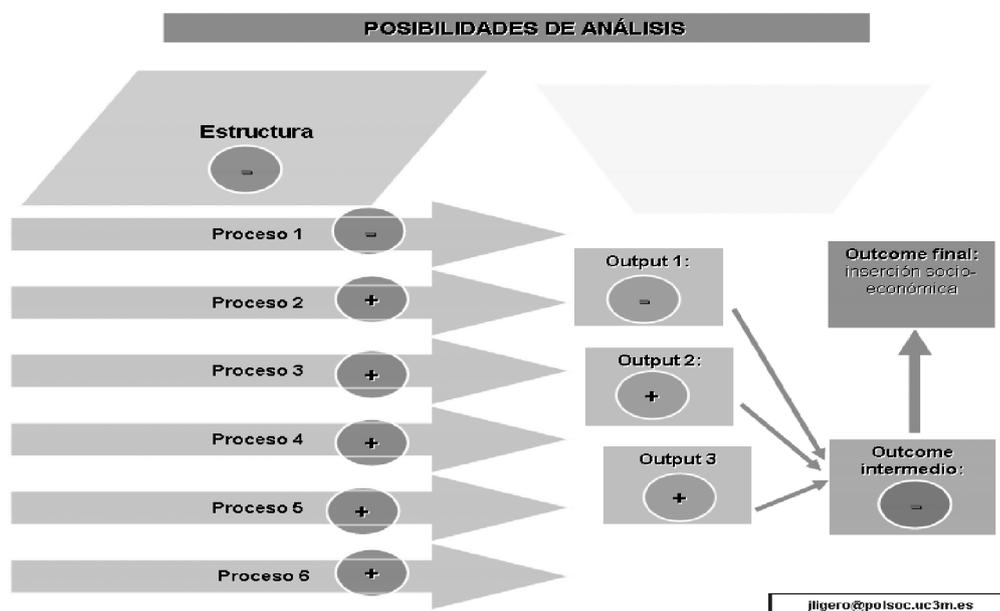
c) Análisis de los elementos estructurales

Por último, si un proceso no se ejecuta de acuerdo a lo definido o tiene problemas de implementación, las causas hay que buscarlas en los elementos estructurales, que son los medios necesarios para poner en funcionamiento la intervención. Pudiera ser que los resultados no se consigan porque los procesos tampoco se han ejecutado correctamente ya que no contaba con los medios adecuados. Si se ha previsto en el diseño de la evaluación, se podrá saber que factores estructurales provocan la mala implementación.

d) Análisis de la cadena causal

Lo que se intenta poner de manifiesto con los análisis anteriores es conocer la cadena causal de medios y fines que recorre el programa para saber dónde ha fallado éste o cuales son las claves del éxito (Weiss 1998, 66). Una vez que se conocen las relaciones causales, se sabe dónde aplicar correcciones o reforzar actuaciones. Es un modelo centrado en la comprensión de lo que se hace y sienta unas bases concretas para orientar la mejora.

Gráfico 21. Posibilidades de análisis: positivos y negativos (elaboración propia).



3.8. Ventajas de la evaluación de la teoría

- El ejercicio de **explicitar la teoría o las teorías** en sí mismo es positivo ya que obliga a pensar o repensar la lógica de la intervención. Es un ejercicio que incrementa la racionalización de la intervención. Incluso si la “racionalización” del programa no es la correcta, la explicitación ayuda a la intervención. Weick recoge la historia de una unidad de reconocimiento perdida tres días en los Alpes en una tormenta de nieve. Finalmente encontraron el camino de vuelta al campamento y se salvaron con la ayuda de un mapa. Un mapa que más tarde descubrieron que era de los Pirineos, no de los Alpes. Tal como dice el autor, “tener un preciso mapa de la zona puede ser menos importante que tener algún mapa que traiga orden al mundo y empuje a la actuación” (Rogers, 54).
- Contribuye a **reducir el disenso y las contradicciones** entre los diferentes actores implicados con relación al programa. Se favorece la búsqueda de denominadores comunes sobre la acción. En el caso de que no resuelvan las discrepancias se contribuye a explicitar el disenso.
- El desarrollo de la evaluación basada en la teoría provoca un **incremento de la comprensión** de los mecanismos por los cuales el programa produce sus efectos (Davidson).
- **Validar el marco teórico:** En el caso de que los datos sean positivos y la teoría congruente, se está ante una lógica de intervención válida al menos en el contexto donde se ha evaluado (aunque hay que tener en cuenta las críticas de los experimentalistas ante este tipo de atribución). Se puede sistematizar la experiencia y difundir tanto en ámbitos académicos como de acción, aportando una nueva información en el proceso de construcción de conocimiento. Otros actores podrán utilizar ese modelo cuando intenten solucionar esos problemas en otros lugares.
- **Modificar/ajustar el marco teórico:** De las conclusiones de la evaluación se extraen indicios razonables que justifican la modificación tanto de la lógica de los objetivos, como de los procesos o de los elementos estructurales. Se puede saber dónde la cadena o las cadenas causales no están funcionando o se han roto.

Por lo tanto, la evaluación de la teoría se dirige con precisión hacia los aspectos mejorables e identifica de forma concreta lo que hay que cambiar de los resultados, de la implementación o de los medios.

- **Cambiar la lógica de la intervención:** Una de las posibles conclusiones es que el marco teórico del programa no funcione. Muchos programas tienen impacto cero debido a unas presunciones ineficaces. La evaluación puede ayudar a legitimar y fundamentar un cambio de orientación parcial o total.
- **Toma de conciencia:** El que los diferentes actores implicados incrementen su conocimiento sobre el programa, conozcan más, cambien la percepción y, en definitiva, tomen conciencia sobre lo que ocurre es un fin en sí mismo. No sólo de algunos enfoques de evaluación sino también de otras metodologías como la Investigación Acción – Participativa. Se parte del principio de que los cambios tienen que ser endógenos y que no hay cambio sin un convencimiento previo de que esto es necesario. Por eso los procesos participativos buscan que los propios actores lideren el proceso de evaluación o investigación, para que a partir de una toma de conciencia real sobre la situación se busquen acciones de consolidación o de reforma. Patton y Saville afirman que se contribuye mejor al desarrollo de la organización cuando ayudan a que el estudio se convierta en una auto-evaluación (Stake, 85). La teoría del programa, por su ordenación lógica de la práctica, ayuda a los diferentes actores a comprender y dotar de sentido a los datos procurando incitar la toma de conciencia.

3.9. Debilidades y críticas

- Existe una confianza excesiva sobre las posibilidades reales de validar la teoría de la acción (Davidson). Existen dudas sobre que los métodos estadísticos permitan sacar conclusiones sobre la atribución causal. Tomas Cook critica que la modelización permita desechar otras teorías alternativas.
- La gran influencia del paradigma constructivista en la elaboración de teorías y el peso otorgado a los diferentes puntos de vista de los actores implicados, puede ser acusado de modelo “**subjetivista**” con la consiguiente pérdida de credibilidad. De hecho, Rossi declara que el staff u otros actores no son perspicaces para identificar la cadena causal, siendo mejor que lo hagan los científicos, entendiendo por ello a las evaluadoras.
- La crítica opuesta es mantenida por Stufflebeam, que por el contrario cree que el hecho de que la evaluadora extraiga o negocie con la teoría la **sesga y usurpa la responsabilidad del diseño del staff**.
- Las **discrepancias entre los actores** a la hora de definir la teoría puede suponer un bloqueo del proceso de evaluación o un desgaste y pérdida de energías en procesos de negociación que pueden enmascarar conflictos latentes de la organización previos a la evaluación.
- La **generalización de los hallazgos** es cuestionada porque está muy **condicionada por el contexto**. La validez de una teoría en un lugar puede no serlo en otro. No obstante, determinadas aproximaciones defendidas por Cronbach, no conceden mucha importancia a este tipo de validez, es más importante la contribución que logra el proceso de evaluación en la toma de conciencia de los actores implicados.

4. Conclusiones finales

Utilizar un método u otro de evaluación no es indiferente, tiene consecuencias que incluso pueden llegar a resultados antagónicos. Las dos propuestas presentadas en este texto, la evaluación por criterios y la evaluación de la teoría del programa, están respaldadas institucionalmente, por la teoría y validadas en la práctica pero tienden a generar diferentes productos y están dirigidas a diferentes propósitos evaluativos.

María Bustelo (2010) recopila tres propósitos para la realización de evaluaciones. Una adaptación de ellos se muestra en el siguiente esquema (ver tabla 7):

Tabla 7. Propósitos de la evaluación.

Propósitos	Concepto e ideas
A) Ilustración, comprensión (<i>enlightenment</i>)	Ilustrar, aclarar, dar luz, comprender qué está pasando y cómo funciona el programa. La evaluación suele producir un cambio, dota de nuevos significados sobre la realidad del programa y de su intervención.
B) La Retroalimentación o mejora (<i>improvement</i>)	Es una manera de recibir “feedback” sobre las acciones, intervenciones o programas que se llevan a cabo, un modo de mejorar y progresar. La evaluación puede ser un proceso o de aprendizaje de la propia evaluación.
C) El control de responsabilidades y rendimiento de cuentas (<i>accountability</i>)	“Rendir cuentas” en la acepción más flexible del término, no se restringe exclusivamente a los aspectos económicos sino también a la responsabilidad social y política, especialmente si las actividades llevadas a cabo se financian con fondos públicos ya que el contribuyente tiene derecho a saber en qué y cómo se está empleando el dinero público.

La evaluación por criterios produce juicios de valor sobre el programa en cada uno de los criterios definidos. No requiere de una comprensión previa para emitir los juicios, sobre todo si el diseño está realizado de antemano como suele ser usual en la cooperación. La comprensión se puede dar, pero el modelo no está orientado hacia esa finalidad. Su funcionalidad se encuentra en el *accountability*, en el rendimiento de cuentas público.

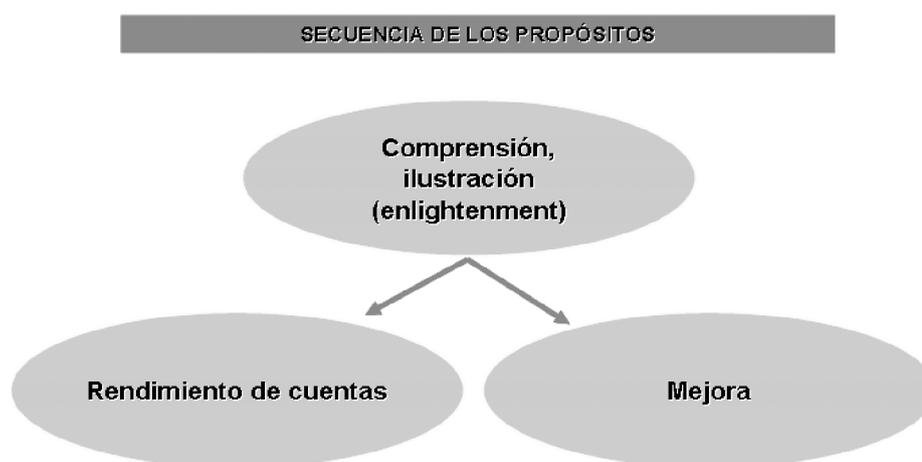
La evaluación de la teoría del programa tiene como finalidad comprender las cadenas o secuencias causales que provocan los resultados. Su intención es saber por qué han ocurrido las cosas. Cuando se sabe el porqué es cuando se puede mejorar, proponiendo recomendaciones enfocadas y concretas que atañen a la totalidad de la intervención. Si se quiere enjuiciar después de la comprensión se puede, pero el modelo no está dirigido hacia el juicio, no lo necesita, ya que su finalidad principal es la ilustración (*enlightenment*) y la comprensión.

En mi experiencia profesional he podido ver como el juicio perdía su importancia ante la fuerza y la significación de la comprensión. Si a los diferentes actores relacionados con el programa les anima una intención constructiva, lo que demandan es la comprensión del fenómeno para poder solucionarlo. Para explicar esto me gusta recurrir a un ejemplo de un estudiante: sus padres pueden recibir informes sobre si el niño es buen o mal estudiante, sobre sus capacidades, aptitudes, actitud y colaboración en el colegio u otros muchos criterios, pero estarán mucho más receptivos a un informe donde les expliquen concretamente las causas del mal rendimiento. Quizás la razón sea que el estudiante no sabe estudiar, nadie le ha enseñado, e invierte con tesón todas sus energías en acciones inútiles. Después de saber esto el enjuiciamiento en función de criterios pierde importancia. Ya se conocen las causas y se pueden buscar acciones de mejora, indistintamente de que se le haya calificado como buena o mala en cualquiera de los criterios.

Lo mismo ocurre con los proyectos de desarrollo, enjuiciar a través de criterios está bien, pero si nos anima una intención de mejorar lo que se está evaluando, puede ser más útiles sistemas de evaluación centrados en la comprensión de lo que ocurre. Cuando se comprende y se toma conciencia de las dificultades que se tienen también se sabe como mejorar. El rendimiento de cuentas público puede concluir diciendo que es necesario mejorar pero no saber exactamente cómo hacerlo.

La mejora de las intervenciones en busca de la calidad está más cómoda en las propuestas de evaluación comprensivas que en las centradas en el *accountability*. Eso me ha llevado a proponer una secuenciación entre los propósitos de evaluación, en la que aparece primero la comprensión y después el enjuiciamiento y la mejora.

Gráfico 22. Secuencia de los propósitos.



Fuente: elaboración propia.

Cada uno de estos propósitos tiene sentido en sí mismo, habrá que saber qué es lo que se busca con cada evaluación en concreto. Pero si lo que se quiere es la mejora será más importante pensar en sistemas centrados en la comprensión como la evaluación de la teoría del programa para después poder deducir recomendaciones concretas. Si por el contrario lo que se busca es producir un rendimiento de cuentas público, un sistema adecuado es la evaluación por criterios.

Existe una literatura propia sobre el escaso uso y utilidad de las evaluaciones y también he escuchado a responsables políticos pensar en “obligar” a usar las evaluaciones. Creo que todo esto son síntomas de una distancia de la evaluación con las necesidades reales de clientes y otros *stakeholders*. Si lo que se quiere es mejorar con sistemas centrados en el *accountability* no es extraño que no se usen las recomendaciones, ya que no se ha generado el producto que es previo al cambio, la comprensión. No hay cambio sostenible si los propios actores implicados no han comprendido lo que ocurre y tomado conciencia de la necesidad de cambiar.

Igual que ocurriría en la definición de criterios a priori, los diferentes actores en el sector de la cooperación pueden tener intereses diferenciados y buscar los métodos de evaluación más adecuados a ellos. La OCDE y el CAD pueden proponer un modelo que busque el *accountability*, ya que sus prioridades y demanda de evaluación puede ser esa, lo cual, por otro lado, es legítimo. Pero los beneficiarios, técnicos y ONG pueden tener otros intereses y demandas hacia la evaluación que también son legítimas y que pueden no verse reflejadas del todo en el modelo de criterios.

La hipertrofia del modelo de criterios en la cooperación responde probablemente a su dependencia característica de los financiadores. Se ha desarrollado de tal manera que se está ante una propuesta metodológica “institucional”. Es decir, la institución o instituciones definen que método hay que utilizar, cuando normalmente esto suele ser atribución del campo técnico o académico. En otros sectores no es así, se sorprenden cuando se asoman a la cooperación. Es como si el Ministerio de Sanidad exigiera solamente evaluar con determinados criterios de eficiencia o eficacia, con técnicas experimentales o el Ministerio de Educación sólo con el modelo CIPP de Stufflebeam. Recibirían merecidas críticas por limitar sus posibilidades, el ajuste a las necesidades de los diferentes actores y se sospecharía de las razones que les llevan a atreverse a tal imposición.

Teóricamente no hay ninguna presión para utilizar el método de criterios. Según el CAD (2010) sus propuestas intentan proveer de una guía hacia las buenas prácticas. Además se pueden encontrar en sus textos una apertura en la que caben otros tipos de evaluación y concretamente la evaluación orientada por la teoría. Así lo expone Nicoletta Stame (2004) entendiendo que la evaluación de la teoría del programa forma parte de los métodos del CAD. En el punto 2.9 del documento citado del CAD, se dice que “la metodología se elabora de acuerdo con el enfoque elegido. (...) La metodología seleccionada debe responder a las preguntas de la evaluación utilizando evidencias creíbles” (CAD, 2010: 10).

Pero tal como lo define Osset y Touza¹⁵ esto es contradictorio con lo que aparece antes en el punto 2.8. En él se especifica que en evaluación se aplican los criterios adaptados por el CAD referentes a la evaluación de la ayuda al desarrollo (CAD, 2010: 9). Aunque las consignas del CAD no sean vinculantes, no hay razón metodológica y técnica para no aplicar los consejos de una organización de la que España participa.

El peso de la institución (OCDE) y de la propuesta del CAD asfixian la necesaria flexibilidad metodológica y convierten cualquier otra propuesta que se salga de los criterios en arriesgada aunque puede estar mayoritariamente aceptada por la comunidad científica y sean “solo unos pocos quienes fijan estándares por adelantado” (Stake 2006, 243).

En cualquier caso los dos métodos son compatibles, se pueden aplicar ambos a la vez. Con esto se puede responder un contexto metodológico rígido y exigente y a la necesidad de sacar más partido a la evaluación. Aunque si uno puede elegir, sería mejor primero comprender qué es lo que ocurre en el programa antes de que los juicios sumativos nos dificulten el proceso de mejora de los programas.

¹⁵ Comentarios realizados por Iván Touza a los estándares del CAD. Grupo de trabajo: Métodos de Evaluación. Magíster en Evaluación de Programas y Políticas Públicas, UCM.

5. Referencias

Alvira, F. (2002). *Manual para la elaboración y evaluación de programas de prevención de abuso de drogas*. Madrid: Comunidad de Madrid, Agencia antidroga.

Bachman, J.G., O'Malley, P., Schulenberg, J.E., Johnston, L.D., Bryant, A.L. y Merline, A.C. (2002). *The Decline of Substance Use in Young Adulthood : Changes in Social Activities, Roles and Beliefs*, U.S.A, Lawrence Erlbaum Associates.

Blas, F.A. (2005). Evolución histórica del estudio del comportamiento organizacional en Gil, F y Alcocer, C.M. *Introducción a la psicología de las organizaciones*. Madrid: Alianza Editorial.

Ballart, X. (1992). *¿Cómo evaluar programas y servicios públicos?* Madrid: MAP.

Becoña, E. (1999). *Bases Teóricas que Sustentan los Programas de Prevención de Drogas*, PNSD.

Bustelo, M (2001). BUSTELO RUESTA, María. 2001. La evaluación de las políticas públicas de igualdad de género de los gobiernos central y autonómicos en España: 1995-1999. Madrid: UCM. Tesis Doctoral. Publicación electrónica.

¿Qué tiene de específico la metodología de evaluación? En Bañon, R. (Eds), *La evaluación de la acción de las políticas públicas*. Madrid: Díaz de Santos.

(2006). The potential role of standards and guidelines in the development of an evaluation culture in Spain *Evaluation*, 12,; 437-453.

(2010). *Materiales Magíster en evaluación* Madrid: UCM.

CAD (1995). *Principios del CAD para una ayuda eficaz. Manual de la ayuda al desarrollo*. Madrid: Mundi-Prensa/OCDE.

Estándares de Calidad para la Evaluación del Desarrollo (2010). OCDE.

Cea D' Ancona, M.A. (1996). *Metodología Cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Síntesis.

Chen H-T. (1990). *Theory-Driven Evaluations*. California: Sage.

Cohen, E. y Franco, R. (1993). *Evaluación de Proyectos Sociales*. Madrid: Siglo.

Cook, T. (2000). The False Choice between Theory – Based Evaluation and Experimentation in Rogers, P.J., Hacsí, T.A., Petrosino, A. & Huebner, T.A (Eds.) *Program Theory in Evaluation: Challenges and Opportunities*. *New Directions for Evaluations*, 87.

Davidson, E.J. (2000). Ascertaining Causality in theory-Based Evaluation in Rogers, P.J., Hacsí, T.A., Petrosino, A. & Huebner, T.A (Eds.) *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Evaluations, 87.

Fitzpatrick, J. (2009). *Ponencia Magíster Evaluación de Programas*. Magíster Evaluación de Programas, UCM.

Giménez, C. (2007). Ponencia presentada en el curso *Retos y Perspectivas de la inmigración: las segundas generaciones*. U. Carlos III de Madrid.

Greene, J. C. (1999). The inequality of performance measurements. *Evaluation* 5: 160-172.

Hintze, J. (2005) Instrumentos de evaluación de la gestión del valor público, en Echebarría y otros *Responsabilización y evaluación de la gestión pública*. Caracas: CLAD.

House, E. (2006). Democracy and Evaluation. *Evaluation* 12 (1):119-127.

Leeuw, F. L. and Furubo, J.-E. (2008). Evaluation Systems. What are they and why study them? *Evaluation* 14,157-169.

Ligero, J.A (2001). Propuesta de un Nuevo Criterio de Evaluación, la Cobertura. *Revista Española de Desarrollo y Cooperación*, nº 8, primavera-verano.

(2010). Comprender la Evaluanda Madrid: Magíster Evaluación, UCM.

Ligero, J.A., Franco, O., García, O. y Pariente, A. (2010). *Informe de evaluación del proyecto Sales Hoy?*. Madrid: Cruz Roja Española, Comunidad de Madrid.

Löffler, E. (1996). *La Modernización del Sector Público desde una Perspectiva Comparativa. Conceptos y Métodos para Evaluar y Premiar. La Calidad en el Sector Público en los Países de la OCDE*. Madrid: Documentos INAP.

López de Ceballos, P. (1989). Un método para la investigación-Acción Participativa. Madrid: Editorial Popular.

Lusthaus, C., Anderson, G. & Murphy, E. (1995). *Institutional Assessment: A Framework for Strengthening Organizational Capacity*. Ottawa: International Development Research Centre.

MAE (Ministerio de Asuntos Exteriores) (2007) Manual de Gestión de Evaluaciones de la Cooperación Española. Madrid: MAE, Secretaría de Estado de Cooperación, DGPOLDE.

(1998) *Metodología de Evaluación de la Cooperación Española*. Madrid: Ministerio de Asuntos Exteriores, Secretaria de Estado para la Cooperación Internacional y para Iberoamérica.

MAP (Ministerio de Administración Pública) (2004). Modelo EFQM de excelencia. Caso práctico para la Administración Pública. Madrid: MAP.

Martínez del Olmo, M.T. (2000). Evaluación de la calidad de los servicios sanitarios. Trinidad, A. *Evaluación y Calidad en las Organizaciones Públicas*. Madrid: INAP.

Monnier, E. (1992). *Evaluación de la acción de los poderes públicos*. Madrid: Instituto de Estudios Fiscales.

Nagarajan N. y Vanheukelen, M. (1997). *Evaluating EU Expenditure Programmes*. European Commission.

Osset, A. (2010). *Diseño de una evaluación del sistema de gestión de evaluaciones de convenios firmados entre las ONGD y AECID*. Memoria de grado Magíster Evaluación de Programas y Políticas Públicas. Madrid: CSEG, UCM.

Rodríguez, L. (2008). *Evaluaciones Colaborativas*. Llumina Press.

Rogers, P.J. (2000). Causal Models in Program Theory Evaluation en Rogers, P.J., Hacsí, T.A., Petrosino, A. & Huebner, T.A (Eds.) *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Evaluations, 87.

Rogers, P.J., Hacsí, T.A., Petrosino, A. & Huebner, T.A (2000). Program Theory Evaluation: Practice, Promise, and Problems en Rogers, P.J., Hacsí, T.A., Petrosino, A. & Huebner, T.A (Eds.) *Program Theory in Evaluation: Challenges and Opportunities*. New Directions for Evaluations, 87.

Rutter, M., Giller, H. y Hagell, A. (2000). *La conducta antisocial de los jóvenes*. Madrid: Cambridge University Press.

Serrano, A., Blanco, F., Ligeró, J.A., Alvira, F., Escobar, M. y Sáenz, A. (2009). *La investigación multimétodo*. UCM.

Shadish, W., Cook, T. & Leviton, L. (1991). *Foundations of Program Evaluation*. Ca: Sage.

Stake, R.E. (1967). The countenance of educational evaluation. *Teachers College Record*, Volume 68, No. 7, April.

(2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Grao.

Stame, N. (2004). Theory-based Evaluation and Varieties of Complexity. *Evaluation*, 10, 58-76.

La evaluación de impacto en VI Seminario de Experiencias de Evaluación (2009). Ponencia. Madrid: Centro Superior de Estudios de Gestión y Evaluación, UCM (www.magisterevaluacion.es).

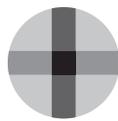
Stufflebeam, D. (2001). *Evaluation Models*. San Francisco: New Directions in Program Evaluation.

Stufflebeam, D. y Shinkfield, A. (1995). *Evaluación sistemática. Guía teórica y práctica*. Barcelona: Paidós.

Varela, F. y Santomé, J.M. (2002). La sostenibilidad en los proyectos y programas de cooperación para el desarrollo. *Dossier Técnico nº 1*. Experiencias de Desarrollo. Revista de Cooperación Internacional. Madrid: Eptisa Internacional.

Weiss, C. (1998). *Evaluation*. New Jersey: Prentice – Hall.

W.K. Kellogg Foundation (2004). *Logic Model Development Guide*. Michigan: W.K.Kellogg Foundation.



CEU

*Instituto Universitario
de Estudios Europeos*

Universidad San Pablo



Boletín de Suscripción

Deseo recibir los próximos números de los Documentos de Trabajo del Centro de Estudios de Cooperación al Desarrollo (CECOD):

Nombre y Apellidos

.....

Dirección

Población C.P. País

Teléfono Correo electrónico

Usted tiene derecho a acceder a la información que le concierne, recopilada en nuestro fichero de clientes, y cancelarla o rectificarla en el caso de ser errónea. A través del Centro de Estudios de Cooperación al Desarrollo (CECOD) podrá recibir información de su interés. Si no desea recibirla, le rogamos que nos lo haga saber mediante comunicación escrita con todos sus datos.

Centro de Estudios de Cooperación al Desarrollo (CECOD)

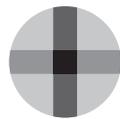
Julián Romea 22, 28003 Madrid

Teléfono: 91 514 04 22 / Fax: 91 514 04 28

info@cecod.org

www.cecod.org

www.cecod-fi.net



CEU

*Instituto Universitario
de Estudios Europeos*

Universidad San Pablo

Boletín de Solicitud de números atrasados

Deseo recibir los próximos números de los Documentos de Trabajo de la Serie “CECOD” del Centro de Estudios de Cooperación al Desarrollo:

Nombre y Apellidos

.....

Dirección

Población C.P. País

Teléfono Correo electrónico

Nº Título

.....

.....

.....

.....

Usted tiene derecho a acceder a la información que le concierne, recopilada en nuestro fichero de clientes, y cancelarla o rectificarla en el caso de ser errónea. A través del Centro de Estudios de Cooperación al Desarrollo (CECOD) podrá recibir información de su interés. Si no desea recibirla, le rogamos que nos lo haga saber mediante comunicación escrita con todos sus datos.

Centro de Estudios de Cooperación al Desarrollo (CECOD)

Julián Romea 22, 28003 Madrid

Teléfono: 91 514 04 22 / Fax: 91 514 04 28

info@cecod.org

www.cecod.org

www.cecod-fi.net

Números Publicados

Serie Unión Europea

- Nº 1 / 2000** “La política monetaria única de la Unión Europea”
Rafael Pampillón Olmedo
- Nº 2 / 2000** “Nacionalismo e integración”
Leonardo Caruana de las Cagigas y Eduardo González Calleja
- Nº 1 / 2001** “Standard and Harmonize: Tax Arbitrage”
Nohemi Boal Velasco y Mariano González Sánchez
- Nº 2 / 2001** “Alemania y la ampliación al este: convergencias y divergencias”
José María Beneyto Pérez
- Nº 3 / 2001** “Towards a common European diplomacy? Analysis of the European Parliament resolution on establishing a common diplomacy (A5-0210/2000)”
Belén Becerril Atienza y Gerardo Galeote Quecedo
- Nº 4 / 2001** “La Política de Inmigración en la Unión Europea”
Patricia Argerey Vilar
- Nº 1 / 2002** “ALCA: Adiós al modelo de integración europea?”
Mario Jaramillo Contreras
- Nº 2 / 2002** “La crisis de Oriente Medio: Palestina”
Leonardo Caruana de las Cagigas
- Nº 3 / 2002** “El establecimiento de una delimitación más precisa de las competencias entre la Unión Europea y los Estados miembros”
José María Beneyto y Claus Giering
- Nº 4 / 2002** “La sociedad anónima europea”
Manuel García Riestra
- Nº 5 / 2002** “Jerarquía y tipología normativa, procesos legislativos y separación de poderes en la Unión Europea: hacia un modelo más claro y transparente”
Alberto Gil Ibáñez
- Nº 6 / 2002** “Análisis de situación y opciones respecto a la posición de las Regiones en el ámbito de la UE. Especial atención al Comité de las Regiones”
Alberto Gil Ibáñez
- Nº 7 / 2002** “Die Festlegung einer genaueren Abgrenzung der Kompetenzen zwischen der Europäischen Union und den Mitgliedstaaten”
José María Beneyto y Claus Giering
- Nº 1 / 2003** “Un español en Europa. Una aproximación a Juan Luis Vives”
José Peña González
- Nº 2 / 2003** “El mercado del arte y los obstáculos fiscales ¿Una asignatura pendiente en la Unión Europea?”
Pablo Siegrist Ridruejo

- Nº 1 / 2004** “Evolución en el ámbito del pensamiento de las relaciones España-Europa”
José Peña González
- Nº 2 / 2004** “La sociedad europea: un régimen fragmentario con intención armonizadora”
Alfonso Martínez Echevarría y García de Dueñas
- Nº 3 / 2004** “Tres operaciones PESD: Bosnia i Herzegovina, Macedonia y República Democrática de Congo”
Berta Carrión Ramírez
- Nº 4 / 2004** “Turquía: El largo camino hacia Europa”
Delia Contreras
- Nº 5 / 2004** “En el horizonte de la tutela judicial efectiva, el TJCE supera la interpretación restrictiva de la legitimación activa mediante el uso de la cuestión prejudicial y la excepción de ilegalidad”
Alfonso Rincón García Loygorri
- Nº 1 / 2005** “The Biret cases: what effects do WTO dispute settlement rulings have in EU law?”
Adrian Emch
- Nº 2 / 2005** “Las ofertas públicas de adquisición de títulos desde la perspectiva comunitaria en el marco de la creación de un espacio financiero integrado”
José María Beneyto y José Puente
- Nº 3 / 2005** “Las regiones ultraperiféricas de la UE: evolución de las mismas como consecuencia de las políticas específicas aplicadas. Canarias como ejemplo”
Carlota González Láynez
- Nº 24 / 2006** “El Imperio Otomano: ¿por tercera vez a las puertas de Viena?”
Alejandra Arana
- Nº 25 / 2006** “Bioterrorismo: la amenaza latente”
Ignacio Ibáñez Ferrándiz
- Nº 26 / 2006** “Inmigración y redefinición de la identidad europea”
Diego Acosta Arcarazo
- Nº 27 / 2007** “Procesos de integración en Sudamérica. Un proyecto más ambicioso: la comunidad sudamericana de naciones”
Raquel Turienzo Carracedo
- Nº 28 / 2007** “El poder del derecho en el orden internacional. Estudio crítico de la aplicación de la norma democrática por el Consejo de Seguridad y la Unión Europea”
Gaspar Atienza Becerril
- Nº 29 / 2008** “Iraqi Kurdistan: Past, Present and Future. A look at the history, the contemporary situation and the future for the Kurdish parts of Iraq”
Egil Thorsås
- Nº 30 / 2008** “Los desafíos de la creciente presencia de China en el continente africano”
Marisa Caroço Amaro
- Nº 31 / 2009** “La cooperación al desarrollo: un traje a medida para cada contexto. Las prioridades para la promoción de la buena gobernanza en terceros países: la Unión Europea, los Estados Unidos y la Organización de las Naciones Unidas”
Anne Van Nistelrooij

- Nº 32 / 2009** “Desafíos y oportunidades en las relaciones entre la Unión Europea y Turquía”
Manuela Gambino
- Nº 33 / 2009** “Las relaciones transatlánticas tras la crisis financiera internacional: oportunidades para la Presidencia Española”
Román Escolano
- Nº 34 / 2010** “Los derechos fundamentales en los tratados europeos. Evolución y situación actual”
Silvia Ortiz Herrera
- Nº 35 / 2010** “La Unión Europea ante los retos de la democratización en Cuba”
Delia Contreras
- Nº 36 / 2010** “La asociación estratégica UE- Brasil. Retórica y pragmatismo en las relaciones Euro-Brasileñas”
(Vol 1 y 2)
Ana Isabel Rodríguez Iglesias
- Nº 37 / 2011** “China’s foreign policy: A European perspective”
Fernando Delage y Gracia Abad
- Nº 38 / 2011** “China’s Priorities and Strategy in China-EU Relations”
Chen Zhimin, Dai Bingran, Pan Zhongqi y Dingchun
- Nº 39 / 2011** “Motor or Brake for European Policies? Germany’s new role in the EU after the Lisbon-Judgment of its Federal Constitutional Court”
Ingolf Pernice
- Nº 40 / 2011** “Back to Square One - the Past, Present and Future of the Simmenthal Mandate”
Siniša Rodin
- Nº 41 / 2011** “Lisbon before the Courts: Comparative Perspectives”
Mattias Wendel
- Nº 42 / 2011** “The Spanish Constitutional Court, European Law and the constitutional traditions common to the Member States (Art. 6.3 TUE). Lisbon and beyond”
Antonio López – Pina
- Nº 43 / 2011** “Women in the Islamic Republic of Iran: The Paradox of less Rights and more Opportunities”
Désirée Emilie Simonetti
- Nº 44 / 2011** “China and the Global Political Economy”
Weiping Huang & Xinning Song
- Nº 45 / 2011** “Multilateralism and Soft Diplomacy”
Juliet Lodge and Angela Carpenter
- Nº 46 / 2011** “FDI and Business Networks: The EU-China Foreign Direct Investment Relationship”
Jeremy Clegg and Hinrich Voss
- Nº 47 / 2011** “China within the emerging Asian multilateralism and regionalism as perceived through a comparison with the European Neighbourhood Policy”
Maria-Eugenia Bardaro & Frederik Ponjaert
- Nº 48 / 2011** “Multilateralism and Global Governance”
Mario Telò

- Nº 49 / 2011** “Bilateral Trade Relations and Business Cooperation”
Enrique Fanjul
- Nº 50 / 2011** “Political Dialogue in EU-China Relations”
José María Beneyto, Alicia Sorroza, Inmaculada Hurtado y Justo Corti
- Nº 51 / 2011** “La Política Energética Exterior de la Unión Europea: Entre dependencia, seguridad de abastecimiento, mercado y geopolítica”
Marco Villa

Serie Política de la Competencia

- Nº 1 / 2001** “El control de concentraciones en España: un nuevo marco legislativo para las empresas”
José María Beneyto
- Nº 2 / 2001** “Análisis de los efectos económicos y sobre la competencia de la concentración Endesa-Iberdrola”
Luis Atienza, Javier de Quinto y Richard Watt
- Nº 3 / 2001** “Empresas en Participación concentrativas y artículo 81 del Tratado CE: Dos años de aplicación del artículo 2(4) del Reglamento CE de control de las operaciones de concentración”
Jerónimo Maíllo González-Orús
- Nº 1 / 2002** “Cinco años de aplicación de la Comunicación de 1996 relativa a la no imposición de multas o a la reducción de su importe en los asuntos relacionados con los acuerdos entre empresas”
Miguel Ángel Peña Castellot
- Nº 2 / 2002** “Leniency: la política de exoneración del pago de multas en derecho de la competencia”
Santiago Illundaín Fontoya
- Nº 3 / 2002** “Dominancia vs. disminución sustancial de la competencia ¿cuál es el criterio más apropiado?: aspectos jurídicos”
Mercedes García Pérez
- Nº 4 / 2002** “Test de dominancia vs. test de reducción de la competencia: aspectos económicos”
Juan Briones Alonso
- Nº 5 / 2002** “Telecomunicaciones en España: situación actual y perspectivas”
Bernardo Pérez de León Ponce
- Nº 6 / 2002** “El nuevo marco regulatorio europeo de las telecomunicaciones”
Jerónimo González González y Beatriz Sanz Fernández-Vega
- Nº 1 / 2003** “Some Simple Graphical Interpretations of the Herfindahl-Hirshman Index and their Implications”
Richard Watt y Javier De Quinto
- Nº 2 / 2003** “La Acción de Oro o las privatizaciones en un Mercado Único”
Pablo Siegrist Ridruejo, Jesús Lavalle Merchán, Emilia Gargallo González
- Nº 3 / 2003** “El control comunitario de concentraciones de empresas y la invocación de intereses nacionales. Crítica del artículo 21.3 del Reglamento 4064/89”
Pablo Berenguer O’Shea y Vanessa Pérez Lamas
- Nº 1 / 2004** “Los puntos de conexión en la Ley 1/2002 de 21 de febrero de coordinación de las competencias del Estado y las Comunidades Autónomas en materia de defensa de la competencia”
Lucana Estévez Mendoza
- Nº 2 / 2004** “Los impuestos autonómicos sobre los grandes establecimientos comerciales como ayuda de Estado ilícita ex art. 87 TCE”
Francisco Marcos
- Nº 1 / 2005** “Servicios de Interés General y Artículo 86 del Tratado CE: Una Visión Evolutiva”
Jerónimo Maíllo González-Orús

- Nº 2 / 2005** “La evaluación de los registros de morosos por el Tribunal de Defensa de la Competencia”
Alfonso Rincón García Loygorri
- Nº 3 / 2005** “El código de conducta en materia de fiscalidad de las empresas y su relación con el régimen comunitario de ayudas de Estado”
Alfonso Lamadrid de Pablo
- Nº 18 / 2006** “Régimen sancionador y clemencia: comentarios al título quinto del anteproyecto de la ley de defensa de la competencia”
Miguel Ángel Peña Castellot
- Nº 19 / 2006** “Un nuevo marco institucional en la defensa de la competencia en España”
Carlos Padrós Reig
- Nº 20 / 2006** “Las ayudas públicas y la actividad normativa de los poderes públicos en el anteproyecto de ley de defensa de la competencia de 2006”
Juan Arpio Santacruz
- Nº 21 / 2006** “La intervención del Gobierno en el control de concentraciones económicas”
Albert Sánchez Graells
- Nº 22 / 2006** “La descentralización administrativa de la aplicación del Derecho de la competencia en España”
José Antonio Rodríguez Miguez
- Nº 23 / 2007** “Aplicación por los jueces nacionales de la legislación en materia de competencia en el Proyecto de Ley”
Juan Manuel Fernández López
- Nº 24 / 2007** “El tratamiento de las restricciones públicas a la competencia”
Francisco Marcos Fernández
- Nº 25 / 2008** “Merger Control in the Pharmaceutical Sector and the Innovation Market Assessment. European Analysis in Practice and differences with the American Approach”
Teresa Lorca Morales
- Nº 26 / 2008** “Separación de actividades en el sector eléctrico”
Joaquín M^a Nebreda Pérez
- Nº 27 / 2008** “Arbitraje y Defensa de la Competencia”
Antonio Creus Carreras y Josep Maria Julià Insenser
- Nº 28 / 2008** “El procedimiento de control de concentraciones y la supervisión por organismos reguladores de las Ofertas Públicas de Adquisición”
Francisco Marcos Fernández
- Nº 29 / 2009** “Intervención pública en momentos de crisis: el derecho de ayudas de Estado aplicado a la intervención pública directa en las empresas”
Pedro Callol y Jorge Manzarbeitia
- Nº 30 / 2011** “Understanding China’s Competition Law & Policy: merger control as a case study”
Jerónimo Maillo

Serie Economía Europea

- Nº 1 / 2001** “Impacto económico de la inmigración de los Países de Europa Central y Oriental a la Unión Europea”
M^a del Mar Herrador Morales
- Nº 1 / 2002** “Análisis de la financiación de los Fondos Estructurales en el ámbito de la política regional de la Unión Europea durante el período 1994-1999”
Cristina Isabel Dopacio
- Nº 2 / 2002** “On capital structure in the small and medium enterprises: the spanish case”
Francisco Sogorb Mira
- Nº 3 / 2002** “European Union foreign direct investment flows to Mercosur economies: an analysis of the country-of-origin determinants”
Martha Carro Fernández
- Nº 1 / 2004** “¿Es necesario reformar el Pacto de Estabilidad y Crecimiento?”
Ana Cristina Mingorance
- Nº 2 / 2004** “Perspectivas financieras 2007-2013: las nuevas prioridades de la Unión Europea y sus implicaciones en la política regional”
Cristina Serrano Leal, Begoña Montoro de Zulueta y Enrique Viguera Rubio
- Nº 3 / 2004** “Stabilisation Policy in EMU: The Case for More Active Fiscal Policy”
María Jesús Arroyo Fernández y Jorge Uxó González
- Nº 1 / 2005** “La negociación de las perspectivas financieras 2007-2013: Una historia de encuentros y desencuentros”
Cristina Serrano Leal
- Nº 9 / 2006** “La cuestión agrícola en las negociaciones comerciales multilaterales”
Ana Fernández-Ardavín Martínez y M^a Ángeles Rodríguez Santos
- Nº 10 / 2007** “El modelo de desarrollo finlandés y su posible adaptación a los países del Este”
Zane Butina
- Nº 11 / 2008** “La estrategia de Lisboa como respuesta de la UE a los retos de la globalización y al envejecimiento de su población”
Miguel Moltó Calvo

Serie del Centro de Estudios de Cooperación al Desarrollo

- Nº 1 / 2003** “Papel de la UE en las recientes cumbres internacionales”
Mónica Goded Salto
- Nº 1 / 2004** “La asociación Euro-Mediterránea: Un instrumento al servicio de la paz y la prosperidad”
Jesús Antonio Núñez Villaverde
- Nº 2 / 2004** “La retroalimentación en los sistemas de evaluación. Experiencias en la cooperación al desarrollo”
José María Larrú Ramos
- Nº 3 / 2004** “Migraciones y desarrollo: propuestas institucionales y experiencias prácticas”
Carlos Giménez, Alberto Acosta, Jaime Atienza, Gemma Aubarell, Xabier Aragall
- Nº 4 / 2004** “Responsabilidad social corporativa y PYMES”
Amparo Merino de Diego
- Nº 1 / 2005** “La relación ONG-Empresa en el marco de la responsabilidad social de la empresa”
Carmen Valor y Amparo Merino
- Nº 1 / 2008** “Dos modalidades de evaluación: evaluaciones de impacto aleatorias y evaluaciones participativas”
José María Larrú Ramos y Jorge Lugrís Llerandi
- Nº 2 / 2008** “A system not fit for purpose?”
Sven Grimm
- Nº 3 / 2008** “El fortalecimiento institucional de la sociedad civil: principal desafío de la cooperación internacional”
Ramón E. Daubón
- Nº 4 / 2009** “La relación entre las instituciones y el desarrollo económico de las naciones”
Pablo Bandeira
- Nº 5 / 2009** “El desarrollo institucional en el contexto de la ineficacia de la ayuda oficial: valoración crítica y propuestas de acción”
Pablo Bandeira
- Nº 6 / 2009** “El fortalecimiento de capacidades y el apoyo al desarrollo desde las bases: la experiencia de la RedEAmérica”
Rodrigo Villar
- Nº 7 / 2009** “Mind the gap: Addressing the “Delivery Challenge” in EC Development Cooperation”
Jean Bossuyt
- Nº 8 / 2009** “De la reforma política en nuevas democracias: aspectos sistémicos e institucionales y calidad de la democracia”
Manuel Alcántara Sáez y Fátima García Díez
- Nº 9 / 2009** Algunas limitaciones metodológicas para analizar la gobernabilidad
Miguel Fernández Trillo-Figueroa

- Nº 10 / 2009** “Fortalecimiento de la sociedad civil para la acción pública y la gobernanza democrática en contextos de desarrollo”
Gonzalo Delamaza
- Nº 11 / 2010** “La gestión de la información en organizaciones de desarrollo Vol. 1 y Vol. 2”
Rodríguez - Ariza Carlos
- Nº 12 / 2010** “¿ Más es mejor?”
Larru, José María
- Nº 13 / 2010** “Civil society capacity building: An approach in Uganda”
Groenendijk, Kees
- Nº 14 / 2010** “El futuro de la cooperación europea al desarrollo: ¿Buscar soluciones globales o volver a un nicho para 2020?”
Sven Grinm y Erik Lundsgaarde

Serie Arbitraje Internacional y Resolución Alternativa de Controversias

- Nº 1 / 2007** “Towards a new paradigm in international arbitration. The Town Elder model revisited”
David W. Rivkin
- Nº 2 / 2008** “Los árbitros y el poder para dictar condenas no pecuniarias”
David Ramos Muñoz
- Nº 3 / 2008** “La lucha contra las prerrogativas estatales en el arbitraje comercial internacional”
José Fernando Merino Merchán
- Nº 4 / 2009** “Due process and public policy in the international enforcement of class arbitration awards”
Stacie I. Strong
- Nº 5 / 2009** “The permanent court of arbitration and the uncitral arbitration rules: current interaction and future prospectives”
Sarah Grimmer

Resumen: No hay una única forma de evaluar, existen diferentes métodos, reflexiones y aproximaciones sobre cómo hacer evaluación. Esto quiere decir que hacer evaluación también implica decidir sobre el procedimiento a utilizar entre una amplia gama de opciones.

Ante esta diversidad de propuestas teóricas, llama la atención que en el sector de la cooperación para el desarrollo se dé la presencia hegemónica de un determinado método de evaluación, como es el modelo de criterios propuesto por el Comité de Ayuda al Desarrollo (CAD) de la Organización para la Cooperación y el Desarrollo Económico (OCDE).

El presente texto pretende revisar el modelo en sus virtudes y defectos y proponer al menos otra forma alternativa de proceder, la evaluación de la teoría del programa. No deja de ser una limitación el proponer solamente una única alternativa sabiendo que existen otras aproximaciones lúcidas y útiles. Pero el objetivo del texto es mostrar que hay otros métodos posibles y no aspirar a convertirse en una recopilación de teorías y modelos, algo que por otro lado ya está escrito.

Palabras clave: Evaluación, cooperación al desarrollo, métodos, evaluación por criterios y estándares, evaluación de la teoría del programa, teoría del cambio, modelo lógico, uso y utilidad de la evaluación.

Abstract: There is no unique way to evaluate. There are many different methods, concepts and approaches. Therefore, conducting evaluations requires deciding which procedure, among many, to follow.

In the context of a diversity of theories, it is surprising to see, in the development cooperation area, the preponderance of the criteria model proposed by the Development Assistance Committee (DAC) of the Organization for Economic Co-operation and Development (OECD).

This paper reviews the criteria model, its strengths and weaknesses and proposes another alternative method: theory driven evaluations. It is limited to offer only one alternative when there are many clever and useful approaches. However, the paper aims to prove that there are other approaches beyond a mere recompilation of theories and models.

Keywords: Evaluation, development cooperation, methods, evaluation criteria and standards, theory driven evaluation, change theory, logic model, use of evaluation.

Patrocina:



Colabora:



Centro de Estudios de Cooperación al Desarrollo (CECOD)
Julián Romea 22, 28003 Madrid
Teléfono: 91 514 04 22 / Fax: 91 514 04 28
info@cecod.org
www.cecod.org
www.cecod-fi.net

