La evaluación de la acción y de las políticas públicas

Rafael Bañon I Martínez

Compilador

Editorial Díaz de Santos

Madrid, 2003

Este material se utiliza con fines
exclusivamente didácticos

¿QUÉ TIENE DE ESPECÍFICO LA METODOLOGÍA DE EVALUACIÓN?

MARÍA BUSTELO RUESTA

1. Las funciones de la evaluación. 2. Tres características específicas de la evaluación. a) El carácter político de la evaluación. b) El enjuiciamiento sistemático. Su carácter práctico.- 3. La distinción de la evaluación con respecto a otras prácticas. a) Evaluación versus investigación. b) Evaluación de programas y políticas públicas versus observatorios de problemas públicos. c) Evaluación versus seguimiento o monitorización. -4. El reflejo de la especificidad en la metodología de evaluación. a) La importancia del contexto como organizador previo y motor de la evaluación. b) La centralidad de los criterios de evaluación. c) Una metodología que además de dirigir la recopilación y el análisis de la información, establezca la forma de enjuiciar y hacer recomendaciones. d) La importancia de la comunicación del proceso y de los resultados de la evaluación.-5. Conclusiones.-6. Bibliografía.

La evaluación de programas y políticas públicas representa un ámbito profesional y académico que ha ido adquiriendo en los últimos años un creciente interés y una especificidad propia. Sin embargo, en España estamos todavía lejos de contar con una práctica normalizada y con una verdadera cultura de la evaluación en nuestros organismos públicos. Este todavía incipiente campo de la evaluación de programas y políticas públicas necesita, por lo tanto, un mayor desarrollo académico y profesional. En muchas ocasiones, la falta de referencias ha significado la sustitución de la verdadera evaluación por otro tipo de prácticas de análisis que, creemos, es necesario distinguir. De esta manera, o bien se ha suplido esta falta de especialización en evaluación con una aplicación estricta y aséptica de la metodología de investigación social al uso, realizando estudios parciales que han sido poco sensibles al contexto y han resultado poco útiles para las organizaciones y las políticas evaluadas, o bien se han manejado las situaciones (aunque en la mayoría de los casos no intencionadamente) para hacer estudios *ad-hoc* que fueran favorables a una determinada coyuntura política, dando origen a situaciones con poca credibilidad externa que han desprestigiado la práctica de la evaluación.

En este capítulo queremos poner énfasis en lo que de específico tiene la evaluación de programas y políticas, diferenciándola de otras prácticas de análisis y estudio y estableciendo conclusiones sobre cómo se refleja dicha especificidad tanto en los conceptos básicos como en la práctica evaluativa. Insistir en su carácter específico es una forma de otorgarle una identidad propia a la evaluación de programas.

Para esto, dividimos el capítulo en cuatro partes. En la primera, definimos las tres funciones básicas de la evaluación. En la segunda desarrollamos tres características específicas de la evaluación: su carácter político, el enjuiciamiento sistemático y su carácter eminentemente práctico y aplicado que hace de la utilidad y de las recomendaciones unas piezas clave de la evaluación. En tercer lugar, se distingue la evaluación de programas de otras tres prácticas de análisis: la investigación, el seguimiento y la evaluación de los problemas públicos. Por último, se extraen conclusiones sobre cómo la especificidad de la evaluación acotada y definida a lo largo del capítulo se refleja en la metodología de la evaluación, es decir, en la forma en la que se evalúa.

1. LAS FUNCIONES DE LA EVALUACIÓN

La evaluación se define, en su acepción más simple, como la recopilación y análisis sistemático de información que nos permite emitir juicios –también sistemáticos– sobre el mérito y el valor del programa evaluado. Sin embargo, la evaluación no se hace en abstracto sino con una intención clara de mejorar los programas y políticas evaluadas, rendir cuentas y responsabilidades y/o ejemplificar o ilustrar acciones futuras. Es decir, la información que se pretende arroje la evaluación tiene vocación de ser utilizada para dichas funciones. Por lo tanto, se puede afirmar que el carácter aplicado y práctico de la evaluación, y que trataremos más adelante, hace necesaria la inclusión, en la propia definición, de su porqué o razón de ser, es decir, de sus funciones.

Stufflebeam y Shinkfield (1987: 23) señalan tres funciones según el uso que se quiera hacer de la evaluación: perfeccionamiento o mejora (*improvement*), recapitulación¹, rendimiento de cuentas o responsabilidad (*accountability*) y ejemplificación o iluminación para acciones futuras (*enlightenment*). En sus diferentes versiones, creemos que éstas son las tres funciones básicas de la evaluación.

¹ Aquí respetamos la traducción realizada en la versión castellana del término *accountability*, aunque nos parece más adecuada la expresión “rendimiento de cuentas” o incluso “responsabilidad”, como se hace en la traducción de este término en otras obras (Rossi y Freeman, 1989; Vedung, 1997).

La primera función, la de *mejora* del programa o política evaluada, significa que la evaluación se concibe como un artefacto metodológico que permite la retroalimentación y el aprendizaje sobre la propia práctica. En este sentido, la evaluación resulta imprescindible como instrumento de mejora y perfeccionamiento y como garante de la calidad de un servicio o programa. Lógicamente esta función refuerza la atención al contexto concreto y determinado de lo que se está evaluando, a las necesidades y naturaleza de las personas destinatarias de dicho servicio y de todos los agentes críticos en el mismo, especialmente el personal que lo atiende y la organización que lo promueve.

El *rendimiento de cuentas o responsabilidad* sería la segunda función de la evaluación. La evaluación debe servir como instrumento que permita, en los diferentes niveles de responsabilidad, rendir cuentas (en el sentido amplio del término) sobre la gestión y los resultados de un programa. Esto adquiere un mayor sentido cuando se refiere a programas y políticas públicas, ya que el rendimiento de cuentas debe formar parte indiscutible del sistema democrático. Este rendimiento de cuentas debe darse a varios niveles, incluyendo de forma especial a la ciudadanía: el/la ciudadano/a tiene derecho a saber no sólo en qué se están empleando los fondos públicos sino también con qué grado de idoneidad, eficacia y eficiencia se están asignando, gestionando y empleando dichos fondos.

Evert Vedung analiza con cierta profundidad esta función de rendimiento de cuentas o responsabilidad, cuestión que normalmente se obvia en la literatura al uso. En su análisis, distingue cuatro perspectivas distintas respecto a la responsabilidad (Vedung, 1997: 134-137). Desde la *perspectiva política*, la evaluación es un instrumento clave para que los representantes políticos comprueben que la administración está realmente ejecutando las tareas que se le han asignado. Desde una *perspectiva técnico-organizativa*, la evaluación es un instrumento para que los directivos de los organismos públicos tengan información sobre si los subordinados están cumpliendo correctamente su tarea. Aunque Vedung reconoce que en el discurso sobre la evaluación rara vez se escucha el punto de vista de los ciudadanos, éstos configurarían una tercera perspectiva. La ciudadanía necesita la evaluación para valorar cómo los políticos elegidos y sus agentes a diferentes niveles están desarrollando sus tareas, y para exigirles responsabilidades en las próximas elecciones. Por último, desde la *perspectiva del cliente* (categoría más limitada que la de los ciudadanos), la evaluación puede producir información sobre si los servicios prestados son satisfactorios y se están distribuyendo equitativamente. Estas cuatro perspectivas hacen que se den dobles lealtades en la cadena de representación pública (los representantes políticos son a la vez superiores y subordinados, la burocracia nacional es agente del gobierno y “jefe” de otros organismos públicos, etc.), siendo evaluados los que a su vez tienen que evaluar, lo que contribuye, según el autor, a la tibieza en el interés mostrado por la evaluación en la vida pública.

Además de estas dos funciones, la evaluación puede cumplir una tercera que es la de dar luz o *iluminar acciones futuras*. Aunque el sentido último de la evaluación tiene que ver, esencialmente, con el programa concreto evaluado en un contexto, situación y momento determinado de forma que se contribuya a su mejora y se rindan cuentas sobre el mismo, las evaluaciones arrojan una información sistemática que contribuye al “acervo científico”, no sólo de ese programa, servicio, o política concreta sino también del abordaje general de determinados problemas públicos. Es claro que si se evalúan sistemáticamente los diferentes programas de prevención de drogodependencias, por ejemplo, dichas evaluaciones contribuirán, en su conjunto, a construir conocimiento sobre las diferentes formas de hacer prevención, cuáles son los contextos y grupos adecuados, quiénes son los agentes intervinientes necesarios, qué respuestas han dado los poderes públicos en diferentes momentos y por qué, etc. En definitiva, se sabrá más sobre cómo hacer prevención y esto servirá de orientación general a los poderes públicos sobre cómo abordar el problema de las drogodependencias. En este sentido, en la actualidad, en la comunidad científica internacional algunos evaluadores reclaman más experiencias de síntesis de evaluaciones, meta-análisis y metaevaluaciones para aprovechar todo el cúmulo de estudios de evaluación realizados en determinadas áreas en los últimos años (Cook, 1997). Desgraciadamente, esta petición no puede ser trasladada a muchos países como España, en donde lo primero que tiene que ocurrir es que se hagan suficientes estudios para poder ser sintetizados.

2. TRES CARACTERÍSTICAS ESPECÍFICAS DE LA EVALUACIÓN

Destacamos, en este apartado, tres características específicas de la evaluación: su carácter político, el enjuiciamiento sistemático y su carácter práctico y aplicado.

a) *El carácter político de la evaluación*

En los últimos años, se ha ido reconociendo, de forma paulatina y creciente, el carácter político de la evaluación. En algunos casos, se entiende que política y evaluación son dos realidades separadas aunque interactuantes. La evaluación debe responder ante el contexto político y, a su vez, ser una herramienta que permita controlarlo. Para otros autores, como Monnier (1995), la evaluación es un acto político en sí mismo.

Un buen indicador del reconocimiento del carácter político de la evaluación fue la publicación de la obra *The Politics of Program Evaluation* en 1987 y cuyo editor, Dennis J. Palumbo concluía en su capítulo introductorio, que la política y la evaluación están intrínsecamente interrelacionadas; tanto, que la efectividad de la evaluación depende de la incorporación de la realidad política a los procesos de evaluación, es decir, la única forma de tener un mayor impacto en las políticas y los programas es incorporando los factores políticos en los diseños de la evaluación y la difusión de sus resultados. Es más, precisamente es la ingenua falta de atención a las dimensiones políticas lo que puede contribuir a que la evaluación se utilice como una herramienta ideológica. Esta obra representa un hito en el campo de la evaluación, no sólo por el reconocimiento explícito de la relación política-evaluación, sino también porque *resta carácter peyorativo al término "política"*: no se trata sólo de partidismos entre partidos políticos sino, sobre todo, de "las interacciones de varios actores dentro y entre burocracias, clientes, grupos de interés, organizaciones privadas y legislaturas y sobre cómo se relacionan unos con otros desde diferentes posiciones de poder, influencia y autoridad" (Palumbo, 1987:18-19). Esta es la primera obra que tiene como propósito explícito aconsejar a los evaluadores la necesidad de comprender la naturaleza política de la evaluación y cómo incorporar la política a su quehacer evaluador.

Una de las primeras autoras que trata el tema de evaluación y política es Carol H. Weiss, quien escribe un artículo en 1973 que también será publicado catorce años más tarde en la obra de Palumbo anteriormente mencionada. En dicho artículo, titulado *Donde la política y la investigación evaluativa se encuentran*, la autora sostenía que la evaluación es una tarea racional pero que se lleva a cabo en un contexto político en el que hay que tener en cuenta tres consideraciones: la primera, que las políticas y programas que se evalúan son resultado de decisiones políticas. Segundo, la evaluación al llevarse a cabo para alimentar el proceso de toma de decisiones entra, inevitablemente, en la arena política. Tercero, la evaluación en sí misma adopta una postura política ya que, por su misma naturaleza, expone implícitamente cuestiones tales como la naturaleza problemática o necesidad de algunos programas, la legitimidad de las estrategias y objetivos, etc. Conocer estas resistencias y fuerzas políticas es necesario para realizar una evaluación útil. Años más tarde (Weiss, 1998), la autora sigue insistiendo en que la evaluación no se hace en abstracto, sino en un contexto determinado; más allá, ésta se realiza, principalmente, para facilitar la toma de decisiones sobre el programa o política evaluada. En definitiva, se reconoce el contexto político, incluso como determinante de una situación evaluativa a la que vamos a tener que responder metodológicamente de formas diversas. La evaluación como proceso sistemático reconoce, pero también ayuda a contrarrestar y controlar, ese clima político en el que está inmerso cualquier programa o política, especialmente si es público. De esta manera, para Weiss, el rigor de la evaluación sistemática es particularmente importante si los programas son complejos, las decisiones a tomar son importantes y caras y si se necesita información para convencer a otros de la validez de las conclusiones.

Sin embargo, otros autores, como Monnier (1995), van un paso más allá, y consideran que no sólo se trata de que la evaluación se desarrolla inevitablemente en un contexto político que hay que reconocer y tener en cuenta, sino que es una actividad política en sí misma y, por lo tanto, forma parte integrante de dicho contexto político. Es decir, evaluación y política no son dos realidades fácilmente separables. De esta manera, la evaluación, al margen del rigor metodológico, debe procurarse una legitimidad política, que supone el reconocimiento y consideración de los distintos actores sociales. Por lo tanto, la evaluación es un elemento generador e integrante de la realidad política.

De esta manera, el carácter político de la evaluación deja de ser un tema relacionado con la práctica evaluativa para ser parte crucial de la teoría de la evaluación. Buen exponente de ello es la conferencia dictada por Eleanor Chelimsky en el encuentro anual de 1997 de la AEA (Asociación Americana de Evaluación) en San Diego, titulada "El papel de la experiencia en la formulación de teorías sobre la práctica de la evaluación" (Chelimsky, 1998) y en la que considera que una de las cuestiones aprendidas con la experiencia, y que deben ser incorporadas a la teoría de la evaluación, es la inequívoca relación entre evaluación y política. Para ella, hasta entonces, se habían examinado evaluación y política separadamente, considerando la política como periférica o contextual a la evaluación y no como conectada a aquella de forma visceral: "Todavía, incluso hoy en día, entendemos la política meramente como el 'contexto' de una evaluación, como algo que se 'entromete' en el buen hacer, más que el motor que la conduce" (Chelimsky,

1998: 39). Sin embargo, se debe considerar la política como algo central en evaluación y los evaluadores deben ser más conscientes del dinamismo político que rodea cualquier situación evaluativa y reconocer su centralidad para la práctica profesional.

Algunas implicaciones de esta centralidad política son la necesidad de credibilidad para la práctica evaluativa, de *flexibilidad* en los procesos de evaluación y de atender a la temporalidad de los mismos. En primer lugar, si los resultados de la evaluación deben ser útiles y sobrevivir a los avatares políticos, la evaluación misma debe ser percibida como creíble e imparcial. En segundo lugar, la necesidad de flexibilidad requiere de una cierta provisionalidad durante el proceso de evaluación. Esto puede significar enfrentar una polémica inesperada cambiando el diseño de la evaluación, añadiendo nuevos consultores o nuevos componentes a la evaluación. En definitiva, cada situación evaluativa puede requerir una respuesta evaluativa única, adaptada a las circunstancias concretas. También implica ser más cautelosos a la hora de hacer valoraciones absolutas, enfatizando las observaciones lo más específicas, precisas y validadas posible.

La tercera implicación es la temporalidad o necesidad de oportunidad en el tiempo (*timeliness*), es decir, que la evaluación se haga en el momento oportuno y no se retrase. Cuando el clima político cambia, presentar los resultados de una evaluación demasiado tarde puede significar lo mismo que no presentarlos. Sin embargo, presentarlos a tiempo puede marcar la diferencia entre ser escuchados o no, que se utilicen los resultados o no, o que se aumente o disminuya la credibilidad del evaluador. En otras palabras, si se tarda mucho en contestar las preguntas de evaluación, puede que el debate que las generó ya no sea el mismo, que esas preguntas sean ya irrelevantes o que los destinatarios de la evaluación hayan desaparecido.

Se pueden mencionar otras implicaciones de este “reconocimiento político” de la evaluación. Así, hay que reconocer que los imperativos de los evaluadores y los de los gestores públicos son distintos y muchas veces entran en conflicto, que hay que entender el valor simbólico y de cambio de la información que ofrecen las evaluaciones, y que es necesario extender las evaluaciones más allá de los objetivos de un programa y comprender los procesos organizativos que subyacen a toda situación evaluativa. También implica que hay que tener en cuenta a todos los agentes críticos (*stakeholders*) –especialmente a los más indefensos– y asumir que existen diferentes, necesidades de información en los distintos puntos del ciclo e una política.

Finalmente, nos gustaría señalar que la evaluación sólo puede entenderse plenamente desde una *comprensión profunda de los procesos políticos y de las políticas públicas*, mucho más que desde una posición técnica estrictamente metodológica. La evaluación de programas y políticas públicas es algo más que una herramienta metodológica o, si se quiere, la “metodología” de evaluación debe entenderse en un sentido mucho más amplio del que normalmente se le asigna.

b) *El enjuiciamiento sistemático. Su carácter práctico*

Además de recopilar e interpretar sistemáticamente información, la evaluación cuenta con otros dos niveles de análisis –valoración o emisión de juicios y establecimiento de recomendaciones– que la hacen específica respecto a otros tipos de análisis e investigación.

En los primeros años de desarrollo de la disciplina de la evaluación, en los que dominaba además una concepción epistemológica netamente positivista y se concebía que la ciencia podía y debía estar libre de valores, había algunos autores que creían que esta orientación valorativa conducía a interpretaciones radicales y subjetivas y otorgaba un poder innecesario a los evaluadores. Sin embargo, a lo largo del tiempo, la mayoría de los autores clave en evaluación han dado a este elemento valorativo una importancia central en el desarrollo y articulación de las teorías de evaluación (Shadish, Cook y Levinton, 1991).

Ya en los años ochenta, este elemento valorativo es plenamente reconocido e incluso utilizado como elemento central en la definición de la evaluación. Stufflebeam y Shinkfield, por ejemplo, insisten en que la meta esencial de la evaluación “es determinar el valor de algo que se está enjuiciando. De acuerdo con esta definición, si un estudio no informa de cuán buena o mala es una cosa, no se trata de una evaluación” (Stufflebeam y Shinkfield, 1987: 20). Precisamente, es este elemento el que organiza su clasificación de modelos o enfoques de evaluación, ya que estos autores denominan “cuasievaluaciones” a los enfoques que no incluyen la valoración como parte de la evaluación y “evaluaciones verdaderas” a los que sí lo hacen.

El reconocimiento de este elemento valorativo implica su consideración como elemento sistemático en el proceso de evaluación. Esto significa que no sólo hay que procurar que la información que se recopile sea de calidad sino que se proporcione y clarifique una base lógica que justifique las perspectivas de valoración utilizadas para interpretar y valorar los resultados. Ya que existen distintos públicos y destinatarios de una evaluación, se pueden y deben emplear diferentes perspectivas de valoración. Esta necesidad de tener en cuenta las diferentes perspectivas de valoración de los actores críticos en el programa o

política evaluada, es una de las claves de la evaluación pluralista propuesta por Monnier. Precisamente, es el compromiso sobre la interpretación de los hechos y sus valoraciones entre observadores y observador, de una parte, y entre los diferentes protagonistas del programa, por otra, lo que hace que una evaluación sea legítima políticamente y, por lo tanto, útil socialmente (Monnier, 1990-91: 143-145).

Siguiendo a Stufflebeam y Shinkfield (1987: 21), además de los diferentes protagonistas, también habrá que tener en cuenta los conceptos de mérito y valor, además de los de viabilidad y equidad. Así las bases de valoración que deben ser tenidas en cuenta son las expectativas de los clientes; el mérito o excelencias del servicio, programa o política (¿hace bien lo que se supone que debe hacer?); hasta qué punto es necesario o adecuado a la comunidad a la que se dirige, es decir, si tiene validez; hasta qué punto es viable; si está pensado y es asequible para todos los miembros de la sociedad y si los niveles de logro o consecución de objetivos no son sustancialmente diferentes para los distintos grupos (equidad).

Estas bases de valoración se verán reflejadas en los criterios de evaluación que se establezcan en cada proceso evaluativo. Otros autores, como Dunn (1981: 343), realizan propuestas similares: efectividad, eficiencia, adecuación, equidad, capacidad de respuesta y propiedad (si merecen la pena los resultados obtenidos).

Además, las bases de valoración pueden ser relativamente diferentes, y tener sus matices específicos, dependiendo del ámbito al que pertenezca la política o programa evaluado. Así, por ejemplo, en la evaluación de las políticas de cooperación al desarrollo existe un consenso generalizado sobre la importancia de valorar a través de los criterios generales de *eficiencia* (hasta qué punto los resultados alcanzados derivan de una utilización eficiente de los recursos financieros, técnicos y humanos), *eficacia* (hasta qué punto se han alcanzado los objetivos propuestos), *impacto* (consecuencias previstas y no previstas para la sociedad, tanto negativas como positivas), *pertinencia* (hasta qué punto el programa evaluado se ajusta, en este caso, a los objetivos globales de desarrollo, a la política del donante y del receptor y a las necesidades y prioridades locales) y *viabilidad* (definida en este ámbito como hasta qué punto los cambios positivos logrados como consecuencia del programa se mantienen después de que éste haya finalizado) (NORAD, 1997; Ministerio de Asuntos Exteriores, 1998 y 2001; Cruz Roja, 2001).

En definitiva, el carácter sistemático del enjuiciamiento se obtiene por dos elementos: en primer lugar, el establecimiento previo de los criterios de evaluación en los cuales se contiene y van a verse reflejadas las perspectivas de valoración que van a utilizarse para enjuiciar los resultados; en segundo lugar, porque los juicios, además de en las perspectivas de valoración preestablecidas, se realizan basándose en una información recopilada y analizada sistemáticamente.

Su carácter práctico y aplicado: la preocupación por la utilización y el establecimiento de recomendaciones.

La preocupación por la utilización y el uso de la evaluación ha sido una constante en la literatura especializada. Dada la naturaleza específica de la evaluación, su éxito no sólo se debe medir por el grado en que ésta contribuye al aumento o mejora del conocimiento sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular.

Esta preocupación también ha llevado a la *identificación de elementos prácticos facilitadores de dicha utilización en los procesos de evaluación* y que aparecen de forma recurrente en la literatura. En general, los evaluadores deben tener en cuenta el *contexto* específico de la evaluación; identificar los *destinatarios* (audiencias) de la evaluación e implicarles (especialmente a clientes y agentes críticos – stakeholders–) en el diseño de la evaluación; mantener un *fuerte y frecuente contacto* y aportar resultados intermedios a lo largo del proceso de evaluación; *ajustar los informes* a las necesidades de los destinatarios, usando una variedad de formatos, incluyendo informes cortos y resúmenes, presentaciones orales y oportunidades para la interacción; presentar ilustraciones reales y concretas de los hallazgos; presentar los resultados con una *temporalidad adecuada* a una variedad de destinatarios; y usar un *lenguaje* simple y claro (Alkin, 1985; *Joint Committee*, 1994; Patton, 1985, 1997; Shadish, Cook y Levinton, 1991).

Algunos autores señalan la necesidad de enfocar la totalidad del proceso de evaluación (y no sólo las fases finales) a la maximización de su uso o, al menos, considerar la cuestión del uso desde el principio (Weiss, 1998: 324). El máximo exponente es Patton, quien elabora un modelo –la “evaluación orientada a la utilización”– que propugna que la evaluación debe hacerse “con y para usuarios específicos e intencionados para usos específicos e intencionados” (Patton, 1997: 23). Es decir, todo el proceso de evaluación, desde su diseño hasta su comunicación, está enfocado a responder las preguntas necesarias a determinados destinatarios identificados previamente. No sólo la utilización de la evaluación es su sentido último, sino que la concepción general de los procesos evaluativos debe estar subordinada a dicho fin.

Sin embargo, de forma paralela a esta preocupación por la utilización, la comunidad científica dedicada a la evaluación se ha ido haciendo más consciente y realista sobre cuáles pueden ser, *realmente*, los usos de las evaluaciones (entre otras cosas, en la toma de decisiones sobre un programa o política, la evaluación es sólo uno de los factores a tener en cuenta y la utilización no siempre es directa e inmediata). Hacerse conscientes de esta realidad ha derivado en una *visión más flexible de la utilización de los resultados* de la evaluación. La evaluación no sólo puede y debe servir para facilitar la toma de decisiones concreta respecto a un programa. También puede contribuir al aprendizaje organizativo y de los grupos humanos, y su utilidad puede ir más allá del programa o política concreta que evalúa, no sólo respecto al ámbito o campo que toque, sino también respecto a los propios procesos de evaluación. Cada proceso de evaluación puede y debe contribuir y ofrecer una información útil sobre cómo evaluar determinadas políticas en determinados contextos, qué enfoques, perspectivas, metodologías y técnicas han sido adecuadas, etc.

Nos resulta especialmente interesante insistir en conceptualizar como uso de la evaluación los efectos que se pueden producir *por el propio hecho de evaluar*. Se sabe que las personas que participan en un proceso de evaluación tienen mayores probabilidades de hacerse más conscientes del programa o política y sus objetivos. En definitiva, la evaluación es reflexión y revisión; el mero hecho de dar un espacio formalizado para ello tiene efectos positivos en sí mismos. Dependiendo del enfoque adoptado para la realización de la evaluación, se pueden obtener otros efectos colaterales. Por ejemplo, con los enfoques participativos se ayuda a las personas en tomo al programa evaluado a reflexionar sobre la práctica, a pensar de forma crítica y a preguntarse por qué el programa funciona como lo hace. Con una evaluación basada en la teoría, los profesionales deben analizar y hacerse conscientes de las premisas subyacentes a su trabajo, etc.

Por la propia naturaleza y carácter práctico y aplicado de la evaluación, ésta debe contar con un último nivel de análisis y exposición de los resultados que consiste en la traducción de los resultados de la evaluación en *recomendaciones*. Estas recomendaciones son sugerencias sobre posibles líneas de acción a seguir respecto al programa o política evaluada, cambios propuestos en el programa, o cosas a mantener en el mismo, o consejos para la mejora dirigidos a los patrocinadores, responsables y/o profesionales del programa. Evidentemente, estas recomendaciones deben seguir la secuencia lógica y sistemática de los niveles de análisis anteriores. Es decir, las recomendaciones de una evaluación deben estar basadas en las valoraciones realizadas que, a su vez, habrán estado basadas en la recopilación y análisis sistemático de la información.

Las recomendaciones son, además, “el vínculo formal entre una evaluación y su uso” (Torres, Preskill y Piontek, 1996: 105). Por lo tanto, el ofrecimiento de recomendaciones específicas y orientadas a la acción se considera como algo deseable y relacionado con las buenas evaluaciones. Patton, aunque insiste en que no sólo las recomendaciones sino todo el resto del proceso evaluativo es lo que provoca y estimula el uso de la evaluación, reconoce, sin embargo, que son, a menudo, la parte más visible de un informe de evaluación (Patton, 1997: 324).

En general, todos los autores prestan atención a este último paso de establecer recomendaciones.² Así, se han desarrollado una serie de orientaciones y sugerencias prácticas para su elaboración (Torres, Preskill y Piontek, 1996; Patton, 1997; Chelimsky, 1998). Todas estas sugerencias denotan la importancia concedida a su elaboración como “producto” de los procesos de evaluación. Una primera cuestión a tener en cuenta al desarrollar las recomendaciones es quiénes participan en su elaboración y de qué forma. En segundo lugar, hay que determinar el tipo de recomendaciones que se quieren realizar: pueden ser simples (se desarrolla la mejor opción), o se pueden desarrollar varias opciones, analizando las debilidades, amenazas, costes y beneficios de cada una, o se pueden dar opciones basadas en diferentes escenarios futuros con sugerencias para el seguimiento, o incluso diferentes recomendaciones para diferentes destinatarios. Otras cuestiones reseñables son que las recomendaciones deben seguir claramente y ser apoyadas por los hallazgos de la evaluación, que deben ser realistas, o mejor dicho, se deben enfocar hacia acciones bajo el control de los destinatarios de la evaluación, y que hay que dedicarle suficiente tiempo a su redacción, para trabajarlas con los destinatarios y para comprobar su claridad, comprensión, carácter práctico, utilidad y significado.

² No todos los autores están de acuerdo con la idea de que las recomendaciones son una parte ineludible de la evaluación. Para Scriven, es una falacia lógica pensar que juzgar el mérito y la valía conduce directamente a las recomendaciones; la principal función del evaluador es valorar y hacer juicios, no la de hacer recomendaciones (Scriven, 1993). Esta postura, que nace de un modelo de evaluación orientado al consumidor, ha sido criticada y/o ignorada por el resto de los autores. Sin embargo, el desprecio de Scriven a este último paso de la evaluación hay que entenderlo en el contexto de la importancia que este autor da al enjuiciamiento sistemático, siendo uno de los autores que más prolífica y seriamente ha tratado el tema.

3. LA DISTINCIÓN DE LA EVALUACIÓN CON RESPECTO A OTRAS PRÁCTICAS

Dada la inexistencia de una cultura de evaluación y de unos referentes claros en evaluación de programas y políticas públicas, se tiende a asignar el apelativo de evaluación a otras prácticas de análisis que realmente no lo son. Por lo tanto, parece imprescindible, al hablar de la especificidad de la evaluación, distinguirla de otras prácticas. Aquí queremos señalar las diferencias con la investigación en general, con los observatorios de problemas públicos y con el seguimiento y monitorización de los programas.

a) *Evaluación versus investigación*

La necesidad de caracterizar específicamente la evaluación, nos hace poner especial énfasis en su *diferenciación con la investigación* en el marco de las políticas públicas. En nuestra opinión, es necesario distinguir la evaluación de la investigación para otorgar a aquella un carácter e identidad propia. Las diferencias se pueden establecer según una serie de dimensiones (Bustelo, 1999), de las que destacaremos aquí algunas.

En los primeros años de desarrollo de la evaluación de programas en las décadas de los años sesenta y setenta, no se encuentran en la literatura especializada distinciones entre ambas actividades. Entonces, se concebía la evaluación como un tipo de investigación de carácter aplicado y que debía guiarse por los estándares de la misma. En realidad, se consideraba la evaluación como una hermana “menor” de la investigación, siendo a veces desprestigiada en la comunidad científica de entonces. La utilización de términos como “investigación evaluativa” (Rossi y Freeman, 1989; Weiss, 1975, etc.), aunque contribuían a dar una mayor “seriedad” y a obtener legitimidad para el nuevo campo en desarrollo, no ayudaban al esclarecimiento o diferenciación de los dos conceptos.

Respecto a su *naturaleza*, es evidente que la investigación y la evaluación son modos de indagación sistemática. Para Lincoln y Guba (1986: 547) se trata de tipos de indagación “disciplinadas”, es decir, producen algo que es confirmable públicamente y que permite explicitar la naturaleza de los datos, las fuentes y el contexto en el que fueron recogidos, así como los procesos de transformación de los datos en información (interpretaciones, conclusiones, extrapolaciones, recomendaciones). Por otro lado, ambos tipos de actividad pueden ser de carácter aplicado; pero la evaluación es siempre de carácter aplicado, nunca básico (es más, se puede afirmar que la evaluación es un tipo de indagación sistemática *especialmente* aplicada). Sin embargo, el hecho de que se dé una investigación aplicada (cuyo fin primario sea la solución de un problema o la respuesta a una necesidad concreta, más que el desarrollo de teoría en abstracto) no la convierte automáticamente en una actividad evaluativa, por lo que se apuntan aquí otras dimensiones que permiten establecer claras diferencias entre ambos conceptos.

La *finalidad* principal de la investigación es la construcción de conocimiento, mientras que la de la evaluación es la mejora de las políticas, el rendimiento de responsabilidades sobre las mismas y la iluminación sobre acciones futuras. Esto implica que la evaluación tiene sentido mientras sea útil para dichas finalidades y para esto debe tener capacidad de respuesta y una temporalidad ajustada. La evaluación es dependiente del *contexto* y tiene por objetivo la respuesta concreta a dicho contexto específico, que además es marcadamente político. Más allá, la evaluación siempre ocurre en un escenario de acción, es decir, lo prioritario es el programa o política evaluada y no su evaluación. También existen diferencias según el *objeto*, siendo el de la investigación los problemas sociales a solucionar y el de la evaluación la manera en que los poderes públicos abordan dichos problemas y si son efectivos o no en ello. Mientras que los *destinatarios* primarios de la investigación son grupos abstractos de la comunidad científica, los de la evaluación son grupos y personas concretas, agentes críticos (*stakeholders*) en torno al programa o política evaluada. Finalmente, respecto a los *agentes*, los investigadores y evaluadores requieren de capacidades de alguna forma diferenciadas: la capacidad investigadora requiere tener conocimientos básicos sobre metodología y exhaustivos del sector sobre el que se investiga. Los evaluadores necesitan de conocimientos metodológicos amplios, aunque el conocimiento sobre el sector sobre el que se evalúa no debe ser tan exhaustivo y puede irse desarrollando a lo largo del proceso. Sin embargo, el evaluador requiere de otras habilidades que le son necesarias, como son el *análisis* de políticas públicas y el conocimiento de los procesos políticos, el conocimiento y la capacidad de trato con personas y grupos, así como habilidades de comunicación.

b) Evaluación de programas y políticas públicas versus observatorios de problemas públicos

El conocimiento exhaustivo del problema que se pretende abordar con una política pública es un elemento necesario para su formulación, sobre todo, pero también para su ejecución, seguimiento y evaluación. Si, además, dichos problemas son complejos y dinámicos y su definición va evolucionando a lo largo del tiempo, ese conocimiento debe darse de una forma periódica y continuada, de forma que se pueda ir ajustando la intervención pública a la realidad cotidiana, respondiendo de una forma adecuada a las demandas y necesidades de la ciudadanía. Esta es la razón por la que han ido surgiendo, en diferentes ámbitos o sectores, los “observatorios” que suelen ser unidades o estructuras dedicadas al seguimiento exhaustivo y periódico de un determinado problema o conjunto de problemas (por ejemplo, las drogodependencias o la igualdad de género). En algunos casos, estos observatorios también recopilan y analizan información sobre las respuestas públicas a determinados problemas (por ejemplo, políticas públicas regionales o desarrollo local).

Sin embargo, la existencia de estos observatorios no obvia ni es equivalente a la evaluación de las políticas y las acciones públicas. Ambas funciones –el seguimiento e investigación sobre los problemas públicos y la evaluación de las acciones públicas que se realizan para paliarlos– son necesarias y complementarias, y no deben confundirse la una con la otra. El hecho de que se siga o monitorice la evolución, por ejemplo, de la situación social de las mujeres respecto a los hombres en una serie de indicadores, por mucho que esos indicadores tengan que ver de alguna manera con los resultados que las políticas pretenden conseguir, no significa, ni muchísimo menos, que se estén evaluando las políticas y las acciones públicas. Recordemos que las funciones de la evaluación son la mejora de las políticas, el rendimiento de cuentas y responsabilidades, y la iluminación para acciones futuras. En esta tercera función es en donde observatorios y evaluación coinciden más plenamente, aunque aquéllos la ejercen principalmente a través del análisis de la evolución del problema a estudiar, y ésta lo hace valorando la adecuación, relevancia y eficacia de las acciones públicas realizadas para paliarlo.

En definitiva, ambas actividades –observatorios y evaluación– se benefician y potencian mutuamente. Pero el hecho de que exista un observatorio, aunque resulta sumamente útil e interesante para la evaluación –proporciona una información sistemática y continua que puede ser muy importante para una parte de la evaluación de las políticas–, no implica que se esté haciendo evaluación. Por otro lado, la evaluación de las actuaciones, programas y políticas ofrece una información que puede contribuir sustancialmente a la identificación de nuevas variables e indicadores a tener en cuenta en los observatorios.

c) Evaluación versus seguimiento o monitorización

Otro de los conceptos con los que a menudo se confunde la evaluación, es el seguimiento o monitorización (monitoreo en otros países de habla hispana) de los programas o las acciones públicas. De hecho, ambos conceptos suelen ser citados conjuntamente, sin establecer ninguna distinción entre los mismos. En cierto modo, la frontera entre un concepto y otro no está claramente establecida y ambas actividades son claramente complementarias, aunque esto no quiere decir que sean lo mismo. Por un lado, es evidente que una buena parte de la información sistemática que debe utilizarse en una evaluación, es recogida en las tareas de seguimiento del programa a evaluar. Por otro, las acciones objeto del seguimiento también son objeto de la evaluación.

En general, el seguimiento es una actividad relacionada con la gestión de los programas y acciones a evaluar, que se hace, además, de forma interna, permanente y continua, y que no implica enjuiciamiento sistemático. Una buena definición de seguimiento es la que ofrece la Agencia Noruega de Desarrollo: “Supervisión continua y periódica de la ejecución física de un proyecto para asegurar que los insumos, actividades, resultados y supuestos están desarrollándose de acuerdo con lo planificado” (NORAD, 1997: 133). Es decir, se trata de una actividad que sirve para ir comprobando que el programa va cumpliendo los objetivos, el calendario y los recursos previamente establecidos. Significa la generación de retroalimentación sobre el proceso de ejecución y los posibles problemas que durante el mismo puedan darse, así como la supervisión de la asignación de recursos y la producción de las actividades o servicios correspondientes. Sin embargo, la evaluación no sólo analiza, sino que también juzga globalmente, centrándose no sólo en el proceso de ejecución de las acciones, sino también en su diseño y en sus resultados, pudiendo ser, además, una actividad interna o externa. La evaluación significa una reflexión más global y profunda y que va más allá del seguimiento. Este, sin embargo, es una actividad cotidiana relacionada con la gestión interna de los programas.

De nuevo, la insistencia en la distinción entre estas dos actividades tiene su razón de ser en otorgarle una identidad propia a la evaluación de los programas. El hecho de que se realice un buen seguimiento de un programa no significa que se haga evaluación del mismo, aunque cuanto mejor sea el seguimiento más se facilitará una parte importante de la evaluación. En definitiva, el seguimiento de un programa es condición necesaria aunque no suficiente de su evaluación.

4. EL REFLEJO DE LA ESPECIFICIDAD EN LA METODOLOGÍA DE EVALUACIÓN

El sentido de la evaluación y sus características específicas se reflejan en los que, a nuestro parecer, son los cuatro elementos claves para la práctica de la evaluación y que deben concebirse y desarrollarse como parte del concepto de “metodología de evaluación”. Estos son los siguientes: la importancia del contexto como organizador previo y motor de la evaluación, la centralidad de los criterios de evaluación, una metodología que además de dirigir la recopilación y el análisis de la información, establezca la forma de enjuiciar y hacer recomendaciones y, finalmente, la importancia de la comunicación del proceso y de los resultados de la evaluación.

a) *La importancia del contexto como organizador previo y motor de la evaluación*

La evaluación se da en un contexto determinado, con un fuerte carácter político, y al cual debe responder. Por lo tanto, la evaluación es y debe ser contextual. Cada proceso de evaluación debe adaptarse a la situación concreta del programa o política a evaluar que puede ser muy distinta en cada caso concreto. Esta adaptación al contexto concreto es necesaria si se quiere que la evaluación sea relevante y útil para el programa o política concreta que se va a evaluar. Conviene aquí recordar que el sentido último de la evaluación consiste en su utilidad para la mejora y el rendimiento de responsabilidades del programa concreto que se evalúa. Por lo tanto, los métodos universales de evaluación, que pretenden aplicarse de igual forma independientemente del programa a evaluar, suelen resultar poco útiles.

Patton (1981: 25-26) identifica una serie de variables situacionales a tener en cuenta para realizar el diagnóstico de una situación evaluativa. Sólo con que varíe una de ellas de una situación a otra puede dar como resultado un contexto evaluativo distinto al que hay que responder de una forma específica. Entre estas variables, se pueden mencionar: el número de responsables y patrocinadores del programa (que pueden o no coincidir), la actitud del personal y beneficiarios ante la evaluación (de entusiasmo o resistencia), la claridad o ambigüedad sobre el fin y la función de la evaluación, la historia del programa, si éste ha sido evaluado o no y cuáles fueron los resultados de las otras evaluaciones, los patrones de interacción entre los diferentes agentes críticos del programa (de conflicto o cooperación), los conocimientos de evaluación de los profesionales del programa, los recursos disponibles para la evaluación, el tiempo que hay disponible para la misma, los datos existentes y la calidad de los sistemas de información y seguimiento, si la evaluación ha sido algo planteado de forma voluntaria o impuesta, etc.

La importancia de este elemento contextual va más allá del hecho de tenerlo en cuenta e incorporarlo en el diseño de los estudios y procesos de evaluación. En realidad, el contexto de una evaluación es de donde debe partir la evaluación –el organizador previo–, donde se ejecuta –y para su adecuada consecución, hay que adaptarse al mismo–, y la realidad concreta y específica a la que la evaluación debe responder. Por lo tanto, el diagnóstico de la situación evaluativa es el primer paso necesario de todo diseño evaluativo. Hacer una lectura ajustada de dicha situación permitirá contestarse a las preguntas de PARA QUÉ (razón de ser de la evaluación, objetivos y propósitos concretos de la misma atendiendo a las necesidades concretas de información) y PARA QUIENES (agentes críticos y destinatarios de la misma), determinar qué tipo de evaluación va a realizarse (formativa o sumativa de diseño, proceso y/o resultados, interna o externa, previa, intermedia y/o final), y adoptar el enfoque de evaluación más adecuado para dicha ocasión.

Una de las características de los contextos evaluativos es la presencia de diferentes personas y colectivos que tienen algo que ver con el programa o política que se pretende evaluar. Aunque el *cliente* o comanditario de la evaluación, es decir, quien encarga la evaluación y decide que hay que realizarla porque tiene unas determinadas necesidades de información, sea sólo uno, los destinatarios de la evaluación, es decir, a quién va dirigida y quiénes son los posibles usuarios de la evaluación, son muchos más. Esta realidad significa que existe un elenco amplio de potenciales *agentes críticos* (*stakeholders*) en la evaluación. Todos los agentes críticos en un programa³, por el hecho de serlo, serían susceptibles de tener interés en su

³ Se pueden establecer cuatro grandes categorías: 1) Patrocinadores, formuladores de políticas y responsables políticos; 2) Directores de los programas, gestores públicos y profesionales en general; 3) Destinatarios directos e indirectos (por

evaluación y, por lo tanto, ser destinatarios y posibles usuarios de los procesos evaluativos. En cualquier caso, se cree conveniente insistir en la pluralidad de actores potenciales de un proceso de evaluación y en su necesaria participación e inclusión en el mismo, ya que a mayor participación de diferentes actores en la evaluación, mayor probabilidad de uso de la misma. Por otro lado, la implicación de los principales actores es condición necesaria para un cierto pluralismo y para la integración de todas las informaciones, interpretaciones y juicios pronunciados en el marco del proceso evaluador. En general, la identificación de los destinatarios y agentes críticos (*stakeholders*) de la evaluación debe guiar también la planificación de la ejecución de la evaluación (se debe saber con quién se ha de contar y sus características y situación para planificar la recopilación y análisis de la información), así como el plan de difusión y comunicación de los resultados.

b) *La centralidad de los criterios de evaluación*

Los criterios de evaluación deben estar claramente definidos y, en los procesos de evaluación, se debe dejar un espacio prioritario para su establecimiento. Efectivamente, uno de los pasos cruciales en evaluación consiste en determinar qué es lo que se quiere evaluar, una vez se ha clarificado el por qué y el para qué queremos evaluarlo. Ya se ha mencionado que los criterios de evaluación son una pieza clave, ya que determinan y guían el estudio, orientando la recopilación y el análisis de la información; en definitiva, se trata de la identificación y definición de las variables que se van a valorar. Pero además, y de forma igualmente importante, en evaluación, se determinan también los criterios de valor, es decir, a través de la definición de los criterios se define la base lógica con la que posteriormente se van a emitir los juicios sobre el programa o política evaluada. En definitiva, los criterios de evaluación son los que definen la perspectiva de valoración a utilizar.

Por otro lado, la definición de los criterios de evaluación puede y debe servir para realizar un proceso de revisión del propio programa o política a evaluar. Este es un paso lógico cuando se cuestiona y reflexiona sobre qué se debe evaluar. En realidad, establecer los criterios de evaluación sirve para clarificar y reforzar los objetivos y para repasar o explicitar la “teoría de programas”, en otras palabras, las premisas, principios y mecanismos subyacentes al programa (por ejemplo, ¿por qué una campaña publicitaria sobre el reparto de las responsabilidades domésticas contribuye a un cambio de actitudes en la población?). Y esto en la doble vertiente apuntada por Weiss, lo que ella llama las “teorías de cambio” de un programa, que serían la combinación de la “teoría de programas” (¿Cuáles son los mecanismos por los que la transmisión de valores a través de los medios de comunicación provoca cambios de actitud?) y la “teoría de la implementación” del programa (¿Qué secuencia de distribución debe darse o cómo debe hacerse una campaña publicitaria para que una campaña publicitaria sensibilice a la sociedad?) (Weiss, 1998: 57-58).

Sin embargo, el análisis de algunas experiencias de evaluación⁴ nos hace concluir que, en general, pensar sobre qué se quiere evaluar –criterios de evaluación– es un paso clave que muchas veces se suele obviar o se suele dar por hecho, pareciendo que ya están preestablecidos unos criterios para la evaluación que son generales, preexistentes y válidos para cualquier situación, pero que además no se suelen explicitar de una forma clara. Una primera cuestión detectada, pues, es que no parece haber una reflexión seria previa a los procesos de evaluación sobre realmente qué se quiere evaluar. Es más, en muchas ocasiones, el QUÉ evaluar (criterios de evaluación) se define a través del CÓMO evaluarlo (metodología y técnicas utilizadas), dando prioridad a esto último. La razón de ser de esta centralidad de las técnicas sobre los criterios creemos que se halla en la falta de referencias propias en evaluación de programas, y que normalmente se han suplido con las coordenadas de otras prácticas de investigación social que se centran especialmente en la metodología sobre cómo estudiar y medir los fenómenos sociales. Pero aquí la clave insistimos en que se encuentra en saber qué se quiere evaluar y por qué (a través de una definición clara de los criterios o preguntas a evaluar), y luego decidir cuál es la mejor manera de contestarse a las preguntas que, de forma consciente, se hayan establecido previamente.

c) *Una metodología que además de dirigir la recopilación y el análisis de la información, establezca la forma de enjuiciar y hacer recomendaciones*

ejemplo, padres de niños en programas educativos), usuarios potenciales del servicio, programa o política evaluada y el público o comunidad en general; 4) Otros: profesionales y responsables políticos de programas similares, expertos del sector o política que se evalúa, científicos sociales en general. La elección de unos u otros agentes críticos como destinatarios principales de la evaluación depende del enfoque que se adopte.

⁴ Además de una apreciación más general, aquí nos basamos en la metaevaluación realizada sobre once experiencias de evaluación de planes de igualdad entre mujeres y hombres en el nivel central y el autonómico (Bustelo, 2001).

El hecho de que las evaluaciones se hayan centrado especialmente en la recopilación y análisis de la información producida por los instrumentos elegidos, también ha provocado que los pasos posteriores (valoraciones y recomendaciones) prácticamente no se hayan resuelto. El déficit de elaboración práctica de la información sistemáticamente recopilada es algo muy común. Por “elaboración práctica” nos referimos a que en evaluación, además de recopilar y analizar información sistemáticamente, se deben emitir juicios sobre la misma, a la luz de los criterios de evaluación previamente elaborados (enjuiciamiento sistemático), y este proceso debe concluir en una serie de orientaciones hacia la acción (establecimiento de recomendaciones), para la mejora, la responsabilidad y el mejor conocimiento de las políticas evaluadas. Utilizamos la denominación de “elaboración práctica”, ya que son estos dos últimos pasos los que confieren el necesario carácter aplicado de la evaluación. En definitiva, las evaluaciones se hacen con la clara y directa intención de ser utilizadas, bien sea para tomar decisiones y modificar aspectos concretos del programa a evaluar, bien sea para contribuir al aprendizaje de las organizaciones que tienen como una de sus misiones la función de evaluar sus actuaciones públicas. Por lo tanto, parece evidente que entre los “productos” de los procesos de evaluación, deben encontrarse también las valoraciones que sistemáticamente se hayan realizado y un listado de recomendaciones que “traduzca” en términos prácticos lo que se ha concluido a través de dicho proceso o estudio.

De esta manera, parece necesario incluir, en el concepto de “metodología de evaluación” no sólo la forma en la que se deben recopilar y analizar la información, sino también la forma en la que enjuicia sistemáticamente y en la que se establecen las recomendaciones, cuestiones que, tradicionalmente, no se consideraban parte de dicha metodología. Además, estas dos cuestiones mencionadas –enjuiciamiento sistemático y establecimiento de recomendaciones– que son los dos componentes de esa “elaboración práctica” a la que nos referimos, son precisamente dos elementos específicos de la evaluación de programas y políticas y que la diferencian de otros tipos de análisis o indagación. El hecho de que son dos pasos que se hacen y se deben hacer sistemáticamente, les confiere un “carácter metodológico” pleno, debiendo estar contemplados en la forma de hacer evaluaciones.

Probablemente, la tradicional no inclusión de esta “elaboración práctica” en el concepto de “metodología de evaluación” es lo que ha generado procesos incompletos de evaluación. En la metaevaluación realizada, mencionada anteriormente (Bustelo, 2001), se concluye que existe un clarísimo déficit en este sentido. En algunas de las experiencias se establecen algunas conclusiones valorativas, aunque estas tienden a ser pequeñas y parciales, es decir, no se realizan prácticamente valoraciones globales de los planes evaluados. En el caso de las recomendaciones es unánime y masivamente deficitario: ninguno de los informes de las experiencias analizadas contiene un apartado específico de recomendaciones o sugerencias, a pesar de la importancia que a este elemento se le da en la teoría de la evaluación.

Esta circunstancia del déficit de “elaboración práctica” es una de las que más influye, según nuestro punto de vista, en la falta de utilidad y de impacto de las evaluaciones. En un contexto en el que no existe cultura de evaluación, probablemente sea muy difícil que los potenciales destinatarios de dicha evaluación sean capaces de hacer dicha “traducción o elaboración” práctica por su cuenta, no por falta de capacidad, sino por falta de costumbre, método y, seguramente, también tiempo. Pretender que los responsables políticos y técnicos de los programas y acciones públicas hagan por sí solos dicha “traducción” puede resultar ingenuo. Esto se agrava más aún si se piensa que otros potenciales destinatarios de la evaluación son la sociedad civil organizada y la propia ciudadanía.

d) *La importancia de la comunicación del proceso y de los resultados de la evaluación*

La buena comunicación, tanto de los procesos como de los resultados de la evaluación, es una condición necesaria para la posterior utilidad y utilización de la evaluación. Además, dada la naturaleza específica de la evaluación, su éxito no sólo se debe medir por el grado en que ésta contribuye al aumento o mejora del conocimiento, sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular. Por lo tanto, una respuesta natural a la preocupación por la utilización de la evaluación y por dar respuesta a las necesidades específicas de información, ha sido la atención prestada a los *procesos de comunicación de los resultados* de las evaluaciones, incluyendo la importancia de la elaboración de un buen *informe*, legible y claro (y su difusión y seguimiento posterior), y de otras formas alternativas de comunicación (como por ejemplo, sesiones de trabajo, presentaciones orales, conferencias de prensa, comunicaciones cortas como faxes, memorandos o correos electrónicos, vídeos, posters, boletines y resúmenes ejecutivos).

Un buen exponente de esta preocupación por el elemento comunicativo son obras recientes como la de Torres, Preskill y Piontek (1996), la cual se centra en fundamentar la importancia de dicho elemento en

los procesos evaluativos y en ofrecer orientaciones prácticas para la comunicación y difusión de los resultados. Una consecuencia de esto es el aumento de responsabilidades en el papel del evaluador: la responsabilidad del evaluador no acaba con la entrega del informe, sino que también tiene responsabilidades de comunicación, difusión y seguimiento de dicho informe. De hecho, tiene todo el sentido esta preocupación con los procesos de comunicación ya que, según demuestra una encuesta realizada a evaluadores profesionales de la Asociación Americana de Evaluación (Torres, Preskill y Piontek, 1997), dichos profesionales estaban sólo moderadamente satisfechos con sus esfuerzos y resultados en la información y comunicación de resultados. Los dos impedimentos más mencionados para una buena información y comunicación tenían que ver con la falta de tiempo invertida en esta tarea, en primer lugar, y la complejidad política y organizativa en los procesos de evaluación, por otro. Las razones mencionadas son: falta de claridad entre los agentes críticos sobre sus necesidades de información, falta de respuesta ante los esfuerzos por comunicar e informar, cambio de cliente o destinatario de la evaluación en mitad del proceso, situaciones muy cargadas políticamente, resistencia ante los hallazgos negativos, características de individuos particulares e interpretación errónea de los hallazgos.

Sin embargo, y a pesar de la importancia concedida a la comunicación, en las experiencias de evaluación analizadas en España, los procesos de comunicación de los resultados son normalmente deficitarios. Los informes no suelen ser completos, no suelen estar accesibles y su difusión suele ser bastante restringida. En este sentido destacamos la variabilidad encontrada respecto al acceso a los informes de evaluación⁵. El carácter público de los informes de evaluación no está claro en muchos casos. Esto nos sugiere una línea de investigación sobre la difusión y acceso a la información producida por las evaluaciones. ¿Qué criterios se están siguiendo en las administraciones públicas españolas? ¿Y cuáles se deberían seguir?

5. CONCLUSIONES

En este capítulo se han destacado las características específicas de la evaluación y cómo éstas deben reflejarse en la metodología al uso. Para esto, se han analizado las funciones de evaluación y se ha apuntado la utilidad para dichas funciones como el sentido último de la evaluación. Así mismo, se han destacado tres características específicas de la evaluación y que la definen diferencialmente con otro tipo de prácticas: su carácter político, de enjuiciamiento sistemático y su carácter práctico y aplicado, el cual se traduce en una especial preocupación por la utilización de sus hallazgos. Además, se establecen las diferencias de la evaluación con tres tipos de prácticas de análisis con las que se suele confundir, como son la investigación, los observatorios y el seguimiento. De esta manera, y una vez delimitadas las características específicas de la evaluación, se concluye y discute cuál es y debe ser el reflejo de dichas características en la metodología de evaluación. Se señalan, por lo tanto, cuatro elementos clave de dicha especificidad: la importancia del contexto, la centralidad de los criterios de evaluación, la necesidad de incluir en el concepto de metodología la forma de emitir juicios y establecer recomendaciones y, finalmente, la importancia de la comunicación del proceso y de los resultados de la evaluación.

⁵ En la metaevaluación ya mencionada, el acceso a los informes de evaluación fue muy heterogéneo según la experiencia de la que se tratara, habiéndonos encontrado evaluaciones en las que fue muy difícil, e incluso imposible, acceder a los informes de evaluación realizados (Bustelo, 2001).

BIBLIOGRAFÍA

- Alkin, M. C. (1985). *A guide for evaluation decision makers*. Beverly Hills, CA, Sage.
- Bustelo Ruesta, M. (1998). Deontología de la evaluación: el modelo de los códigos éticos anglosajones. *Gestión y Análisis de Políticas Públicas 11-12* (Enero/Agosto 1998): 141-156.
- (1999). Diferencias entre evaluación e investigación: una distinción necesaria para la identidad de la evaluación de programas. *Revista Española de Desarrollo y Cooperación 4* (Primavera verano 1999): 9-29.
- (2001). *La evaluación de las políticas públicas de igualdad de género de los gobiernos central y autonómicos en España: 1995-1999*. Tesis doctoral. Madrid, Universidad Complutense.
- Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation 19* (1): 35-55.
- Cook, T. D. (1997). Lessons learned in evaluation over the post 25 Years. En: E. Chelimsky, W. R. Shadish, (Ed.). *Evaluation for the 21 st Century. A handbook*. Thousand Oaks, Sage. 30-52.
- Cruz Roja Española (2001). *Cooperación internacional. Evaluación de proyectos. Manual del alumno/a*. Madrid, Cruz Roja Española.
- Dunn, W. N. (1981). *Public policy analysis: An Introduction*. Englewood Cliffs, Prentice-Hall. Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards. 2nd ed. How to assess evaluations of educational programs*. Thousand Oaks, Sage.
- Lincoln, Y. y Guba, E. (1986). Research, evaluation, and policy analysis: Heuristics for disciplined Inquiry. *Policy Studies Review 5* (February 86) (3): 546-565.
- Ministerio de Asuntos Exteriores (1998). *Metodología de evaluación de la Cooperación Española*. Madrid, Ministerio de Asuntos Exteriores. SECIPI.
- (2001). *Metodología de evaluación de la Cooperación Española II*. Madrid, Ministerio de Asuntos Exteriores. SECIPI.
- Monnier, E. (1990-91). Objetivos y destinatarios de las evaluaciones. *Documentación Administrativa 224-225*: 131-152.
- (1995). *Evaluación de la acción de los poderes públicos*. Madrid, Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda.
- NORAD (1997). *Evaluación de proyectos de ayuda al desarrollo. Manual para evaluadores y gestores*. Madrid, IUDC-UCM y CEDEAL.
- Palumbo, D. J., (Ed.). (1987). *The politics of program evaluation*. Newbury Park, Sage.
- Patton, M. Q. (1981). *Creative Evaluation*. Beverly Hills, Sage.
- (1985). *Culture and evaluation*. San Francisco, Jossey-Bass.
- (1997). *Utilization focused evaluation. The new century text. 3rd. ed.* Thousand Oaks, Sage.
- Rossi, P y Freeman, H. (1989). *Evaluación. Un enfoque sistemático para programas sociales*. México, Trillas.
- Scriven, M. (1993). *Hard-won lessons in program evaluation*. San Francisco, Jossey-Bass.
- Shadish, W., Cook, T. y Levinton, L. (1991). *Foundations of program evaluation. Theories of practice*. Newbury Park, Sage.
- Stufflebeam, D. y Shinkfield, A. (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid, Paidós/MEC.
- Torres, R., Preskill, H. y Piontek (1996). *Evaluation strategies for communicating and reporting*. Thousand Oaks, Sage.
- Torres, R., Preskill, H. y Piontek (1997). Communicating and reporting: Practices and concerns of internal and external evaluations. *Evaluation Practice 18*(2): 105-125.
- Vedung, E. (1997). *Evaluación de políticas públicas y programas*. Madrid, Ministerio de Trabajo y Asuntos Sociales. Instituto Nacional de Servicios Sociales.
- Weiss, C. H. (1987). Where politics and evaluation research meet". En: D. J. Palumbo, (Ed.). *The politics of program evaluation*. Newbury Park, Sage. 47-69.
- (1998). *Evaluation. Methods for studying programs and policies*. Second edition. Upper Saddle River (NJ), Prentice Hall.