

La evaluación de la intervención social. Enfoques teóricos

The evaluation of the social intervention. Theoretical approaches

Miguel Santos González

Universidad Nacional de Educación a Distancia (UNED)

Resumen: Ante la necesidad y el creciente interés de la comunidad científica de disponer de una base teórica y metodológica para realizar la evaluación y la valoración de las intervenciones sociales, se intenta recopilar y sintetizar los diferentes enfoques teóricos en este estudio. Aunque, inevitablemente, esta evaluación sólo puede provenir de una investigación rigurosa, que derive de la aplicación sistemática del método científico, debido tanto a las limitaciones de la evaluación clásica y cuantitativa, como a las debilidades de la evaluación cualitativa, se ha generado un apasionado debate. Así, encontramos enfoques que entienden la evaluación como un juicio de valor, sinónimo de valoración o puntos de vista más restrictivos como la evaluación clásica y experimental. Otras perspectivas teóricas sitúan la prioridad en diferentes elementos: el conocimiento y la información, la gestión, los beneficiarios o la política. En la actualidad y como conclusión al debate generado, se detecta una tendencia a un enfoque teórico y metodológico integrador, pragmático y adaptado, en el que destacan algunos elementos clave para la práctica de la evaluación como proceso, valoración, relación causal, validez interna y externa, eficacia y eficiencia, toma de decisiones, formulación de políticas e intervenciones; con una finalidad: solucionar problemas sociales.

Palabras clave: Evaluación, Impacto, Intervención, Social, Teoría.

Abstract: Faced with the necessity and the growing interest in order to have a theoretical and methodological basis with the aim of carrying out the evaluation and the valuation of the social interventions by the scientific community, we are trying to collect and synthesize the different theoretical approaches in this study. Although, inevitably, this evaluation can only come from rigorous research, which derives from the systematic application of the scientific method due to both limitations of the classical and quantitative evaluation and weaknesses of the qualitative evaluation, a passionate debate has been generated. Thus, we find approaches which understand the evaluation as a value judgment, synonymous with valuation or more restrictive viewpoints like the classical and experimental evaluation. Other theoretical approaches have as priority different elements: knowledge and information, management, beneficiaries or policy. Nowadays and concluding the debate generated, it is detected a tendency to a theoretical and methodological approach which is integrating, pragmatic and adapted, and where some key elements for the practice of the evaluation highlight, such as,

process, valuation, causal relationship, internal and external validity, effectiveness and efficiency, decision-making, policy-making and interventions. All of them have the same purpose: solving social problems.

Key Words: Assessment, Impact, Intervention, Social, Theory.

Recibido: 15/09/2018 Revisado: 29/11/2018 Aceptado: 05/01/2018 Publicado: 10/01/2019

Referencia normalizada: Santos González, M. (2019). La evaluación de la intervención social. Enfoques teóricos. *Ehquidad International Welfare Policies and Social Work Journal*, 11, 55-90. doi: 10.15257/ehquidad.2019.0003

Correspondencia: Miguel Santos González. Facultad de Ciencias Políticas y Sociología. Universidad Nacional de Educación a Distancia (UNED). Correo electrónico: miguelsantos@movistar.es.

1. INTRODUCCIÓN

En este artículo se presenta un resumen del marco teórico de la Tesis Doctoral del autor (actualmente en proceso), titulada: “Evaluación relacional y del impacto social. Evaluación de la Fundación Lesmes de Burgos”, cuya finalidad es profundizar en esta área del conocimiento. Por lo tanto, este trabajo se plantea en el marco de la actividad investigadora dirigida a evaluar programas o intervenciones sociales. Dentro de las ciencias sociales, uno de los campos de estudio que ha despertado interés en la comunidad científica es la necesidad de disponer de una base teórica y metodológica para realizar la evaluación y la valoración de las intervenciones sociales. Resulta de vital importancia disponer de un sistema eficaz y económicamente viable, para obtener información de las consecuencias de las distintas políticas sociales desarrolladas. Cualquier intervención de carácter social, sea pública o de cualquier otra naturaleza, debe tener como resultado un beneficio para la sociedad, lo cual, tiene que ser inequívocamente demostrado. Esta evaluación sólo puede provenir de una investigación rigurosa, que derive de la aplicación sistemática del método científico. La obviedad y la contundencia de esta afirmación no debería abrir la puerta a ningún enfoque diferente, de hecho, debería zanjar cualquier discusión sobre el asunto.

Sin embargo, a pesar de que ante dicho irrefutable argumento existe un acuerdo general, debido, por un lado, a las limitaciones detectadas en la aplicación rígida de los diseños experimentales y, por otro lado, a la imposibilidad de realizar diseños puros en la mayoría de las ocasiones, se ha generado un apasionado debate. Desde la década de 1960, en el contexto de la gran influencia, no sólo en el campo de la evaluación de programas, sino en el de la investigación en general, de los trabajos de Campbell y Stanley (1963), sobre diseños experimentales, cuasiexperimentales y correlacionales, han ido surgiendo diferentes enfoques alternativos, orientados a una evaluación de carácter cualitativo. Como apunta Ballart (1996), ha sido la práctica profesional la que ha cuestionado los planteamientos teóricos.

Después de la revisión de las aportaciones de numerosos autores, llevada a cabo a lo largo de este trabajo (Ballart, 1996) y de un largo y profundo debate en la comunidad científica, en él analizado (Rossi & Freeman, 1985), se observa una tendencia a la integración de los diferentes enfoques teóricos y al pragmatismo, tanto teórico, como metodológico y aplicado. Vamos a profundizar en el estado de la cuestión, pero antes de entrar en la discusión sobre los diferentes enfoques teóricos actuales, postulados durante el último tercio del siglo XX, repasaremos muy someramente los antecedentes y el contexto histórico-científico en el que surge el debate, siguiendo a Anguera y Chacón (2008). Estos autores diferencian tres períodos:

- a) Antes de 1930. Basándose en el trabajo de otros autores como Freeman y Solomon (1981), Stufflebeam y Shinkfield (1987), Shadish, Cook y Leviton (1991), Fernández-Ballesteros (1992), Achúcarro y Guillén (1996) y Fernández-Ballesteros (1996); citan algunos antecedentes en China, en la antigua Roma o en Grecia, así como otros estudios realizados después del siglo XVII, que pueden ser considerados como ejemplos de evaluaciones de intervenciones y programas sociales. Hacen referencia a un estudio formal sobre un programa de ortografía realizado en América, a finales del siglo XIX, que, según Stufflebeam y Shinkfield (1987), debería ser reconocido como la primera evaluación social de la historia.

- b) Entre 1930 y 1973. Etapa denominada por Anguera y Chacón (2008) como la época de los grandes modelos, en la que, a partir de Madaus, Scriven y Stufflebeam (1983) apoyados en un trabajo de Smith y Tyler (1942) identifican a su vez tres periodos:
- 1.- Época tyleriana, durante la década de los cuarenta, caracterizada casi exclusivamente por la indagación sobre el cumplimiento de objetivos.
 - 2.- Época de la irresponsabilidad social, durante la cual, se destinan muchos recursos públicos a la intervención, por lo que se producen grandes avances, por ejemplo, en educación, pero no se observan progresos en materia de evaluación.
 - 3.- Época del realismo, hasta 1973, una etapa de cambios y progreso en el campo de la evaluación, priorizando la utilidad y la relevancia como criterios.
- c) Después de 1973. Anguera y Chacón (2008) se refieren a esta etapa como la de la configuración como disciplina científica, de la madurez y de la profesionalización de la evaluación de programas, avanzando hacia un enfoque integrador.

2. ENFOQUES TEÓRICOS

Para tratar de agrupar y sintetizar los enfoques teóricos que han aparecido y centrado el debate científico sobre la evaluación de programas, durante la última parte del siglo pasado y que han llevado a la posición integradora actual, se ha seguido la selección de autores concretada por Ballart (1996), que a su vez se basa en la selección de Shadish, Cook y Leviton (1991). Según Ballart (1996), las aportaciones teóricas se deben fundamentalmente a autores norteamericanos, debido a que los trabajos europeos se han centrado en la realización y difusión de evaluaciones concretas. Se trata de una muestra significativa de los autores que mejor representan a las distintas perspectivas teóricas y metodológicas. Además, se ha completado dicha selección con aportaciones de otros autores, las cuales, a riesgo de simplificar en exceso, para favorecer el hilo conductor de la argumentación, se han tratado de clasificar. Es preciso aclarar, por tanto, que todos estos enfoques se entremezclan y comparten planteamientos, lo que dificulta

encasillar a la mayoría de los autores. Esta clasificación coincide con las clasificaciones propuestas en infinidad de trabajos que profundizan sobre los diferentes enfoques teóricos y metodológicos y que se hace más evidente al revisar sus fuentes bibliográficas. A modo de ejemplo: Anguera y Chacón (2008), Martínez-Mediano (1998), Bueno y Osuna (2012), De Miguel (2000), Bustelo (1999), Bouzas (2005), Hernández (2002) y Alvira (1991).

2.1. LA EVALUACIÓN COMO JUICIO DE VALOR: VALORACIÓN

Scriven (1991), destacado filósofo de la ciencia, entiende la evaluación como una ciencia cuyo objeto es determinar el mérito y el valor de lo evaluado. Su finalidad se orienta hacia la valoración. Fue el primero en plantear la distinción entre una evaluación formativa o de funcionamiento, dirigida a mejorar un programa y una evaluación sumativa o de resultados, planteada por Scriven (1967). Sin embargo, Scriven (1991) no plantea cómo medir esos resultados de forma operativa, de hecho, apuesta por una evaluación no contaminada por los objetivos del programa, libre de metas, aunque, eso sí, teniendo siempre muy presentes las cuestiones éticas, como establece el mismo Scriven (1973). Podemos afirmar que Scriven (1991) sienta las bases de la evaluación como disciplina científica, la investigación evaluativa.

A pesar de que Scriven (1991) rechaza la idea de que la evaluación se plantee al servicio de la toma de decisiones, ya estén éstas encaminadas a la gestión, ya se dirijan a los intereses de los beneficiarios, podemos encontrar importantes similitudes en algunos planteamientos teóricos. Stufflebeam y Schikfield (1987), en su modelo iluminativo, orientado a la toma de decisiones, consideran la evaluación como un proceso para proporcionar información descriptiva sobre el valor y el mérito de la ejecución de los resultados e impacto de una intervención social. Desarrollan un modelo de evaluación de carácter integral, global, que atiende a cuatro ámbitos: contexto, input, proceso y producto. Defienden una complementariedad entre los dos enfoques: cuantitativo y cualitativo. Stufflebeam y Schikfield (1987) distinguen evaluación proactiva, orientada a la toma de decisiones y retroactiva, enfocada a establecer responsabilidades.

Parlett y Hamilton (1972) afirman, siguiendo en esta línea de modelo iluminativo, que aplicar la evaluación tradicional a la evaluación de programas es un método inadecuado, porque es reducir la evaluación al puro control, que además está contaminado por influencias exógenas. Scriven (1991) hace referencia a estas influencias cuando habla de la necesidad de eliminar los inevitables sesgos.

2.2. LA EVALUACIÓN CLÁSICA – EXPERIMENTAL

La evaluación clásica es, como apunta Weiss (1975), a priori, muy sencilla: se trata de plantear unos objetivos, unas actividades encaminadas a conseguirlos y finalmente averiguar si se han logrado. La idea que subyace es el establecimiento de una relación causal entre intervención y resultados. De hecho, esta es la idea de fondo presente en todo el debate sobre la evaluación de programas. Este enfoque defiende que tal relación causal sólo se puede hallar aplicando el método científico-experimental, es decir, la investigación al servicio de la evaluación de la intervención social, como metodología.

Campbell y Stanley (1963) desarrollan este principio de la relación causal en su teoría de la causación, basada en el concepto de validez. Es la aleatoriedad de los experimentos la que proporciona esta validez y la que permite establecer relaciones causa-efecto. Sin embargo, los estudios aleatorios no suelen ser factibles en la evaluación de programas sociales, por lo que se deben utilizar diseños cuasiexperimentales o correlacionales, con la consiguiente amenaza a la validez. Es aquí donde realizan una decisiva aportación a la investigación evaluativa, al proponer diseños cuasiexperimentales basados en medidas previas a la intervención y en comparaciones con grupos similares (no aleatorios). De esta forma, se pueden realizar atribuciones causales. Es destacable el trabajo de Campbell y Boruch (1975) dirigido al desarrollo de técnicas para corregir las limitaciones de los diseños cuasiexperimentales. En un apartado posterior de este trabajo, se profundiza en los diseños aplicables y en los conceptos de validez (interna y externa) y de fiabilidad.

Veamos ahora las aportaciones de algunos autores que adoptan este enfoque y que han contribuido significativamente a su desarrollo. Crane (1988) entiende la evaluación como una práctica investigadora interdisciplinar, en la que, simplificando, podemos identificar dos elementos: la medición y la valoración, siendo ésta la que genera la investigación por situarse en el centro sobre el que gira el proceso. Bingham y Felbinger (2002) y Langbein y Felbinger (2006) señalan que para calcular la eficacia de un programa es imprescindible comprobar que ha conseguido los objetivos planteados a su inicio y, además, comprobar cuál sería la situación de sus beneficiarios si no se hubiera llevado a cabo. Para ello, sólo puede aplicarse el método experimental. Por su parte, Riecken y Boruch (1974), firmes defensores de la aplicación de dicho método, reconocen que la utilización de técnicas cualitativas en el proceso de evaluación de la intervención social ha tenido como consecuencia una evolución y un enriquecimiento de la metodología utilizada. Bickman y Rog (2008) condicionan la evaluación y las técnicas usadas a su significación práctica (sustantiva), en el sentido de que se debe establecer una relación causal entre el programa y los resultados.

En esta línea, en clara contraposición con Scriven, pero en el mismo contexto temporal, Suchman (1968) defendía como planteamiento la aplicación estricta del método científico. La evaluación se queda en pura emisión de juicios de valor si no se basa en la investigación evaluativa, que, a través de evidencias científicas, demuestra la obtención de resultados. Así, Tyler (1986) resume la evaluación como la información del grado de consecución de los objetivos, comparando objetivos propuestos o metas y resultados, siempre con diseños experimentales o cuasiexperimentales. Shortell y Richardson (1978) centran su atención en la operacionalización de los objetivos, que permita formular indicadores para la medición y para la evaluación cuantificada. El evaluador, previamente, debe asesorar la formulación consensuada de los objetivos con los responsables de la intervención. Para algunos autores como Wolf (1979), la clave de la evaluación sólo puede ser la comprobación inequívoca de que, de hecho, se han conseguido los objetivos.

Sin embargo, esta perspectiva fue suavizándose paulatinamente. Shadish, Cook y Leviton (1991), autores que, con Campbell y Stanley (1963), han defendido el enfoque científico-experimental clásico en la evaluación, fue evolucionando hacia un planteamiento de múltiples metodologías, no exclusivamente experimentales. Shadish y Reichardt (1987) defienden el desarrollo de una Teoría de la Evaluación para evitar el elevado esfuerzo económico acarreado por una metodología múltiple. En el desarrollo de este planteamiento teórico, Shadish, Cook y Leviton (1991) identifican cinco componentes:

- Programación social.
- Construcción del conocimiento.
- Valores.
- Utilidad del conocimiento.
- Práctica evaluativa.

2.3. LA PRIORIDAD ES EL CONOCIMIENTO Y LA INFORMACIÓN

Weiss (1975), desde el campo de la sociología, que en su planteamiento inicial defendía la aplicación de los métodos experimentales, fue una de las primeras en detectar sus limitaciones. Su argumento parte de la observación de la escasa utilización de los resultados de las evaluaciones clásicas, en especial en la toma de decisiones políticas, además de constatar que estas evaluaciones, a menudo, hallaban escasa eficacia en los resultados de los programas. Weiss (1975) apunta que, entre otras razones, esto puede deberse a que se realizan evaluaciones inmediatas, sin dejar transcurrir el tiempo necesario. Por su parte, Dunn (1994), apunta la necesidad de focalizar los esfuerzos en definir bien los problemas, puesto que constata que cuando las evaluaciones ofrecen resultados insuficientes se debe más a una incorrecta definición de la necesidad que a una intervención inadecuada.

Otro de los graves problemas que encuentra Weiss (1998) es el de los conflictos entre los diferentes grupos de interés o *stakeholders*, siendo uno de los más frecuentes el conflicto entre los intereses de los evaluadores y los de los responsables de las intervenciones, los cuales, en su opinión, deben

prevalecer. Otra de las grandes limitaciones que encuentra Weiss (1998) en la aplicación de métodos experimentales es la aplicación de tales diseños en la realidad, fuera del laboratorio, las dificultades para crear grupos de control y para poder comparar los resultados entre programas.

Debido a la detección de estas limitaciones, la propia autora termina recomendando la utilización de métodos cualitativos. Weiss y Bukuvalas (1980) apuntan la desconexión entre los resultados de las evaluaciones y la formulación de las políticas sociales. Su propuesta es que la evaluación no sea un instrumento para las políticas sociales, sino una fuente de información y de conocimiento, de ahí que subrayen la importancia de la difusión de las evaluaciones y de su efecto de sensibilización. Por este motivo, Nioche (1982) pone el acento, cuando se trata del análisis de políticas públicas, en el carácter dialogante, participativo y democrático del evaluador, frente a la rigidez del enfoque tecnocrático.

Weiss, aunque introduce el elemento del contexto político, no lo sitúa en el centro de atención, sin embargo, otros autores, compartiendo su enfoque, realizan planteamientos muy cercanos a la línea de Cronbach y Snow (1977). Por ejemplo, Palumbo (1987) subraya el carácter político de la evaluación, y, sin entenderlo como algo negativo, destaca la importancia 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, p. 18-19). Si la evaluación no tiene en cuenta la realidad política y si, además, no se difunde, pierde toda eficacia. Monnier (1995) todavía da un paso más en esta dirección y considera la evaluación como un acto político en sí mismo. Evaluación y política no son elementos separables y esto no está reñido con la aplicación de un método riguroso. Chelimsky (2006) apunta la relevancia del momento de la evaluación, en relación con el contexto político, concluyendo que el incumplimiento de la temporalidad oportuna puede dejar obsoleta la evaluación, en especial, cuando las circunstancias políticas son inestables.

2.4. LO CENTRAL ES LA GESTIÓN

Wholey (1979), profesional del campo de las matemáticas, en la misma línea que Weiss, condiciona la evaluación a la utilidad de sus productos, sin embargo, sitúa la evaluación no orientada a las políticas, sino al servicio de la gestión de las mismas y de los diferentes programas e intervenciones. Horst, Nay, Scanlon y Wholey (1974) encuentran en sus trabajos que, desafortunadamente, esto, en la práctica, no sucede o lo hace en escasa medida. Las evaluaciones no se utilizan lo suficiente para mejorar la gestión de los programas. Wholey, Abramson y Bellavita (1986) entienden la gestión como las decisiones sobre la aplicación de los recursos, pero siempre, focalizados en unos objetivos especificados, previamente formulados: modelo de gestión orientada a resultados. Defienden la participación directa del evaluador en la gestión del programa para poder mejorarlo. Teniendo en cuenta que la evaluación es, a menudo, extremadamente costosa, Abramson, y Wholey (1981) son concluyentes al condicionarla a su utilidad práctica. Precisamente debido a estos elevados costes, Wholey (1983) plantea diferentes estrategias para determinar el grado de profundidad de la evaluación a desarrollar:

a) Nivel 1:

- Análisis de evaluabilidad, para averiguar si la evaluación va a ser útil y puede contribuir a la mejora de la intervención.
- Evaluación rápida retroalimentadora, que consiste en un control rápido del cumplimiento de los objetivos y que proporciona información para decidir si se realiza una evaluación más profunda y completa. Smith, y Brandon (2008) y Hendricks (1981) recopilan y desarrollan estrategias, mediante comparaciones con otras intervenciones, para extraer este *feedback*.

b) Nivel 2:

- Control de resultados, comparación con la situación previa.
- Evaluación intensiva. Diseños experimentales y cuasiexperimentales para establecer relaciones causales.

De acuerdo con este enfoque, Prottas (1979) entiende que la valoración de una política no puede llevarse a cabo sin la participación de las personas que la aplican, por su capacidad de tomar decisiones sobre su implementación, aunque Majone (1989) advierte de que un evaluador de intervenciones sociales y políticas desempeña un papel más parecido al de un abogado que al de un científico, debido a la clase de argumentación en la que se basa y a la multitud de fuentes de información. Browne y Widavsky (1987) proponen que la evaluación de cualquier intervención debe servir para mejorar el proceso de implementación e incluso pronosticar su funcionamiento futuro.

2.5. LO PRINCIPAL SON LOS BENEFICIARIOS

Stake (2004), psicólogo educativo, desde la perspectiva de la evaluación cualitativa, pionero en este enfoque, desarrolla un modelo que denomina evaluación comprensiva o “respondiente” (*responsive evaluation*). Para este autor, la finalidad de la evaluación es el descubrimiento del valor o mérito y de las deficiencias de una intervención, de manera que el evaluador decide y elige los aspectos a priorizar en el proceso de la evaluación. Se basa en la metodología del estudio de caso, junto con otros autores: Stake y Gjerde (1975) o Stake y Easley (1978). Para Stake (2004), las ciencias sociales no pueden establecer relaciones de causalidad y entiende que la evaluación experimental encuentra serias limitaciones. Otra de las claves del modelo de Stake (2004) consiste en situar el centro de atención en los intereses de los beneficiarios de la intervención, respondiendo a sus necesidades (*responsive evaluation*). La evaluación no debe servir al profesional ni tiene que centrarse en la difusión del conocimiento, debe ser útil a las personas a las que se dirige la intervención. Igual que Wholey (1983), defiende la implicación del evaluador en el desarrollo del programa. Johnson y Joslyn (1995) afirman que cuando se trata de un fenómeno político, rara vez es posible un diseño experimental, por tanto, el método adecuado es el cualitativo, de estudio de caso. Ahondando en esta idea, por su parte, Weiss y Rein (1970) señalan que centrarse en los resultados finales dificulta la obtención de la valiosa información procedente de los inevitables acontecimientos no previstos, que en absoluto son infrecuentes. Además, estos autores afirman que los programas son cambiantes en función del contexto en el que se desarrollan.

Para abordar los diferentes tipos de problemas en función de los mismos, Alkin y Christie (2005) y Alkin, Patton y Weiss (1990), desde este enfoque cualitativo, diferencian dos grandes grupos de modelos de evaluación en la base del árbol teórico:

- a) Un modelo prescriptivo, cerrado, regulado, con prescripciones y prohibiciones, dejando claro cómo debe realizarse la evaluación.
- b) Un modelo descriptivo, general, abierto que predice o en su caso explica las actividades y las consecuencias.

Provus (1969) propone un modelo de discrepancia para la evaluación, referido a programas educativos, pero aplicable a cualquier tipo de intervención, basado en las diferencias entre los objetivos del programa y su funcionamiento en un contexto aplicado. Hogwood y Gunn (1984) afirman que la evaluación de la intervención pública y de las políticas precisan analizar todas las fases del proceso, puesto que una valoración global está más alejada de la realidad, sin embargo, entienden que la investigación experimental, a pesar de su presupuesta objetividad, es la menos indicada, al menos para descubrir nuevas necesidades y problemas complejos.

Eisner (1985) habla de la evaluación como “crítica artística” y Patton (1990), incluso, llega a proponer un pensamiento creativo: "La investigación de evaluación cualitativa se basa tanto en el pensamiento crítico como en el creativo; tanto en la ciencia como en el arte del análisis" (Patton, 1990, p. 434). En lo que se refiere a la actitud del evaluador, Patton (1990) recomienda:

- a) Estar abierto a múltiples posibilidades.
- b) Generar una lista de opciones.
- c) Explorar varias posibilidades antes de escoger una.
- d) Hacer uso de múltiples formas de expresión tales como el arte, la música y las metáforas para estimular el pensamiento.
- e) Usar formas no lineales de pensamiento tales como ir hacia atrás y hacia delante y darle vueltas a un tema para lograr una nueva perspectiva.

- f) Divergir de las formas normales de pensamiento y trabajo, también para conseguir una nueva perspectiva.
- g) Confiar en el proceso y no amedrentarse.
- h) No tomar atajos, sino ponerle energía y esfuerzo al trabajo.
- i) Disfrutar mientras se ejecuta. (p. 434-435).

House (1980) señala que las conclusiones sobre la valoración de una intervención social no pueden ser extraídas exclusivamente por un evaluador, sino en un foro democrático conjunto con la participación de los grupos de interés, en el que se llegue a un consenso, teniendo muy presente la dimensión ética. House (1980) habla de una evaluación metafórica, en la que los programas son objetos (metáfora) que pueden ser observados, medidos y valorados. Bryk (1983), en esta línea de planteamiento, propone una evaluación basada en los diferentes niveles de grupos de interés. Fetterman (1998) afirma que es necesaria la riqueza que proporciona a la evaluación una pluralidad de visiones, debido a la complejidad de la realidad. Fullan (2005), desde una perspectiva sistémica, afirma que la evaluación debe contemplar todos los ámbitos y contar con todos los agentes interesados e implicados. Para Guba y Lincoln (1981), la recogida de información en la evaluación, no se puede llevar a cabo separada de la valoración cualitativa y la interpretación, es decir, evaluación y valoración van necesariamente unidas. Estos autores rechazan la utilidad de los métodos experimentales o cuasi-experimentales para la evaluación Lincoln y Guba (1985). Para ellos, como para Bryk (1983), es necesario tener muy presentes a los distintos grupos de interés. Por su parte, King, Morris y Fitz-Gibbon (1987) ponen de manifiesto la importancia de la evaluación de la implementación y el funcionamiento del programa y la participación en la misma de las personas cercanas al mismo. Por último, Kelly (1998) se muestra igualmente partidaria de procesos inductivos y de técnicas cualitativas, como entrevistas a personas representativas y seleccionadas, rechazando las técnicas puramente experimentales.

2.6. LA PRIORIDAD ES LA POLÍTICA

Desde el ámbito de la psicología, Cronbach (Cronbach & Snow, 1977), adoptando un punto de vista ecléctico entre la postura clásica y el enfoque naturalista, entienden la evaluación como un conglomerado diverso de actividades y de metodologías cuantitativas y cualitativas al servicio de la toma de decisiones en la intervención, aunque este proceso no es directo, sino que requiere una contextualización y una interpretación (los autores utilizan el término “*social inquiry*” por ciencia social). Lo nuclear del proceso es la generalización, proporcionada por la validez externa, que no es exclusiva de los métodos cuantitativos. Cronbach y Shapiro (1982), afirman que los métodos experimentales y rigurosos encorsetan la evaluación y limitan sus posibilidades. A la hora de evaluar un programa, identifican cuatro elementos (UTOS): unidades, tratamiento, resultados y contexto. Cronbach desarrolló y extendió la diferenciación de Scriven entre evaluación formativa y sumativa. Lo primordial, por lo tanto, para estos autores es la toma de decisiones políticas (no de gestión del programa) y a ello se debe dirigir la evaluación. De alguna manera, existe una notable conexión entre el planeamiento de Cronbach y el de Weiss, al conceder gran importancia al contexto político, por un lado, y al avance de la ciencia de la evaluación, por otro. Así, este conocimiento, podrá ser transmitido y producirá generalizaciones, esto es validez externa, un concepto al que este autor concede gran importancia y que desarrolla en profundidad, junto con Snow (Cronbach & Snow, 1977).

2.7. LA EVALUACIÓN ADAPTADA

Rossi (Rossi & Freeman, 1985), sociólogo, cuyo origen se sitúa en el enfoque clásico-experimental, va evolucionando hasta representar como ningún otro autor la articulación y la integración del resto de las teorías desarrolladas sobre evaluación. Para llevar a cabo dicha postura integradora, desarrolla el enfoque de la evaluación comprensiva, adaptada y conducida por la teoría. De esta manera Rossi y Freeman (1985) se apoyan en tres fundamentos:

- a) *Compressive evaluation*: evaluación global o total.
- b) *Tailored evaluation*: evaluación a la medida o adaptada.
- c) *Theory-driven evaluation*: evaluación sobre el modelo teórico, desarrollada por Chen y Rossi (1980) y Chen (1990). Se trata de detectar,

en el caso de que no se alcancen los efectos deseados, si se trata de un defecto de planteamiento teórico en las bases del programa, si la causa es que se ha llevado mal a la práctica o si el programa no es adecuado o suficiente.

Para Rossi y para Lipsey y Pollard (1989), igual que la mayoría de los teóricos, la evaluación engloba tres componentes, cuyas acciones se realizan en 3 momentos diferentes:

- a) Análisis relacionados con la conceptualización y diseño de las intervenciones.
- b) Control de la realización y del desarrollo del programa.
- c) Valoración de la utilidad del programa, sus efectos y la relación entre costos-eficacia.

Además, introducen el análisis económico, sin embargo, su aportación diferencial, es que Rossi y Lyall (1978) no plantean la realización de los tres en todas las evaluaciones, sino que propone diferentes modelos de evaluación, adaptados a cada programa, a medida. Así lo aplican en su práctica evaluadora Berk, Lenihan y Rossi (1981). Y este planteamiento basado en modelos teóricos adaptados a cada intervención no excluye ninguna de las aportaciones de otros enfoques teóricos, desde los más experimentalistas, cuyas técnicas propone prioritarias, hasta los más cualitativos y creativos. De acuerdo con este principio, Rossi y Freeman (1985) defienden que el modelo de evaluación será diferente para cada tipo de programa, identificando tres categorías:

- a) Programas innovadores, en implantación.
- b) Programas en proceso de ajuste.
- c) Programas estables.

Rossi y Wright (1977) afirman que la evaluación puede realizarse a distintos niveles:

- a) La evaluación a nivel de los medios previstos y utilizados.
- b) La evaluación de realizaciones, es decir de los resultados.
- c) La evaluación del grado de eficiencia.

De acuerdo con este planteamiento integrador, otros autores como Pollit (1995), afirman que la evaluación de la intervención social es un conglomerado de procedimientos basados en diferentes teorías y en distintos procedimientos metodológicos, de manera que el evaluador debe seleccionar aquellos que mejor se adapten al programa objeto de evaluación, siempre basándose en criterios económicos, de eficacia y eficiencia. De la Orden (1990) establece los tres criterios de calidad: la eficacia, la eficiencia y la funcionalidad; en función de las relaciones entre los elementos de un programa: objetivos y metas, procesos, entradas y productos. De esta forma, establece que:

- a) La eficiencia se refiere a la relación de los productos con los procesos y las entradas.
- b) La eficacia es la relación entre productos y entradas.
- c) La funcionalidad se refiere a la relación entre los productos y los objetivos metas.

Owen y Rogers (1999) entienden la evaluación de programas como un proceso de aplicación de la investigación, que está determinada por la intervención, su objeto de estudio, su finalidad y su contexto. Se trata de un proceso adaptado, que debe ser riguroso y sistemático. Preskill y Catsambas (2006), desde el enfoque de la investigación apreciativa, apuntan que es preferible atender a las fortalezas de los programas y realizar evaluaciones ambientales, intentando potenciar lo que mejor funciona, que establecer límites a la propia evaluación, tratando de detectar fallos o de recortar actuaciones.

3. LA EVALUACIÓN EN LA ACTUALIDAD. LA INVESTIGACIÓN EVALUATIVA

A continuación se analizarán los distintos enfoques teóricos.

3.1. ANÁLISIS DE LOS ENFOQUES TEÓRICOS

Simplificando, podemos agrupar los anteriores enfoques teóricos en tres categorías:

- El enfoque cuantitativo. Utilizaremos este término para denominar al que muchos autores se refieren con términos como clásico o científico-experimental.
- El enfoque cualitativo. Seleccionamos esta denominación, aunque encontramos otras como evaluación naturalista, pluralista, ambientalista, libre de objetivos...
- Enfoque integrador.

3.1.1. Limitaciones de la evaluación clásica - cuantitativa

Como hemos señalado, esta evaluación se basa en el establecimiento de relaciones causales entre la intervención y los efectos producidos. Para ello, se procede a:

- Formulación de objetivos y metas.
- Diseño e implantación de un programa de intervención social encaminado a conseguir los objetivos.
- Comprobación de que se han logrado dichos objetivos.

Según Ballart (1992), por un lado, la definición de los objetivos ya conlleva serias dificultades, puesto que deben ser operacionales, medibles y que faciliten la formulación de indicadores cuantitativos. Por otro lado, las estrategias no suelen ser sencillas, lo que dificulta el establecimiento de relaciones causales directas. Y, por último, la aplicación de diseños experimentales, con grupos de control aleatorios, fuera del laboratorio, resulta prácticamente imposible en infinidad de ocasiones, además de ser susceptible de cuestionamiento ético. En las situaciones reales, realizar medidas antes y después de la intervención y aislar otras variables que puedan influir durante el proceso es muy difícil.

En opinión de Weiss (1972), además, encontramos otras limitaciones. Por una parte, de tipo temporal, al tratarse de evaluaciones, a menudo largas en el tiempo, lo que supone que pierden utilidad porque no llegan a tiempo para tomar las decisiones adecuadas y, por otra parte, de tipo económico, al tratarse de estudios de elevado coste. Por último, los resultados obtenidos a través de estos métodos son resultados técnicos, de interpretación ambigua, que, además, suelen concluir que la eficacia de los programas es muy similar,

de bajos resultados, sin ofrecer diferencias entre los mismos, lo que disminuye la utilidad práctica de la propia evaluación.

3.1.2. Debilidades de la evaluación cualitativa

Agrupamos aquí buena parte de los enfoques teóricos analizados, de manera que vamos a extraer los aspectos comunes a todos ellos. Siguiendo a Ballart (1992), se trata de cuatro puntos o características:

1. Variedad del proceso de evaluación, de forma participativa, atendiendo a un amplio abanico de dimensiones, de participantes... Se propone una evaluación pluralista.
2. Utilidad de las conclusiones de la evaluación. Todos estos enfoques conceden gran importancia a que los resultados de la evaluación puedan ser utilizados y, de hecho, se utilicen, ya sea en el ámbito de la política, de la gestión o en favor de los intereses de los beneficiarios.
3. Metodología cualitativa, como la única vía para profundizar en el contenido del programa, interpretarlo y entenderlo, así como el contexto en el que se desarrolla.
4. Metodología dinámica, sujeta a cambios durante el propio proceso de evaluación. Se rechazan los métodos experimentales que requieren ser aplicados en una situación fija, irreal.

Mientras la evaluación clásica centra el foco de atención en observar los efectos de la intervención y compararlos con la situación inicial, los enfoques de carácter más cualitativo lo hacen en el proceso y en la implementación, para identificar los aspectos que puedan explicar dichos efectos.

Es evidente que una de las limitaciones de los enfoques naturalistas es la debilidad de sus conclusiones al no poder controlar algunos sesgos como los intereses de los participantes en el proceso evaluador, por ejemplo, responsables técnicos, políticos, beneficiarios o el propio evaluador. Este sistema de participación conlleva un proceso negociador, que, por una parte, puede terminar siendo no representativo y que, por otra, también requiere tiempo, con lo que no se soluciona una de las limitaciones importantes de la evaluación experimental, como establece Ballart (1992).

3.2. ENFOQUE INTEGRADOR

Resumiendo, la evaluación clásica encuentra dos grandes limitaciones, por un lado, la dificultad de aplicar diseños experimentales fuera del laboratorio, y por otro, la escasa utilidad de los resultados que ofrece. Por su parte, la evaluación pluralista se topa con problemas de validez. Una postura integradora puede minimizar ambas limitaciones, sin embargo, incrementa la necesidad de recursos y alarga aún más el proceso temporal. A la luz de este análisis, la opción teórico-metodológica que parece más acertada es la evaluación adaptada, propuesta y desarrollada por Rossi y Freeman (1985), basada en el estudio de caso y planteando una investigación evaluativa a medida del programa a evaluar y de las posibilidades de llevarla a cabo. Esta es la perspectiva teórica, de carácter holístico, que se adopta en este trabajo. La evaluación que aquí se lleva a cabo viene definida por la naturaleza, las peculiaridades y la propia esencia de la intervención objeto de valoración, la cual es el primer gran objetivo del estudio. Se trata pues, de una metodología de estudio de caso, que, como hemos visto, en absoluto es incompatible con una posterior generalización, tal y como establecen Cronbach y Shapiro (1982), lo que posibilita, a priori, el segundo de los objetivos de este trabajo: el diseño y la validación de un modelo de evaluación de otros programas de similar naturaleza. Sin embargo, esta perspectiva de evaluación sí que requiere que dicho modelo, así como sus posibles instrumentos y herramientas sean adaptables a las características de otras intervenciones. Más adelante, se justifica y se desarrolla esta proposición.

4. LA PRÁCTICA DE LA EVALUACIÓN DE PROGRAMAS. METODOLOGÍA

En este apartado, se van a estudiar algunas aportaciones, que vamos a denominar perspectivas metodológicas, para diferenciarlas de los anteriores enfoques teóricos. Si bien, no podemos hablar de teoría y de metodología como si se tratara de dos áreas separadas, pues tanto los primeros como los segundos contienen elementos teóricos y metodológicos, es cierto que las propuestas que revisaremos a continuación tienen un carácter claramente aplicado. Además, todas ellas pueden ser conectadas con alguno de los enfoques anteriores.

4.1. LA EVALUACIÓN DE SERVICIOS Y PROGRAMAS SOCIALES SEGÚN LA PROPUESTA DE AGUILAR Y ANDER-EGG (1994)

Aguilar y Ander-Egg (1994) entienden la evaluación como un proceso de investigación, en el que sólo es aplicable el método científico, de manera que se sitúan muy cerca del enfoque teórico clásico. La evaluación es, según estos autores:

...una forma de investigación social aplicada, sistemática, planificada y dirigida; encaminada a identificar, obtener y proporcionar de manera válida y fiable, datos e información suficiente y relevante, en qué apoyar un juicio acerca del mérito y el valor de los diferentes componentes de un programa (tanto en la fase de diagnóstico, programación o ejecución), o de un conjunto de actividades específicas que se realizan, han realizado o realizarán, con el propósito de producir efectos y resultados concretos; comprobando la extensión y el grado en que dichos logros se han dado, de forma tal, que sirva de base o guía para una toma de decisiones racional e inteligente entre cursos de acción, o para solucionar problemas y promover el conocimiento y la comprensión de los factores asociados al éxito o al fracaso de sus resultados. (p. 18).

La propuesta de estos autores para la clasificación de los tipos de evaluación, en función del criterio de los aspectos a evaluar es la siguiente:

a) Evaluación del diseño y conceptualización del programa, que, Según Aguilar y Ander-Egg (1994):

...juzga la pertinencia formal y potencial de un programa, teniendo en cuenta básicamente la coherencia interna entre sus distintos componentes. Comporta tres aspectos fundamentales:

- Estudio-investigación.
- Diagnóstico.
- Diseño y concepción del programa.” (p. 38).

b) Evaluación de la instrumentación y seguimiento del programa, que consiste en “...evaluar cómo está funcionando” (Aguilar & Ander-Egg, 1994, p. 18). Siguiendo la propuesta de estos autores, comprende varios aspectos:

- Cobertura.
- Implementación.
- Ambiente organizacional.
- Rendimiento personal.

c) Evaluación de la eficacia (resultados) y eficiencia (rentabilidad).

- De la eficacia (resultados). Aguilar y Ander-Egg (1994), entienden que es la parte más importante de todo el proceso evaluativo, quedando el resto de evaluación supeditada a este aspecto y distinguen entre:
 - Evaluación de resultados, cuando se focaliza en los destinatarios.
 - Evaluación de impacto, cuando se tiene en cuenta el conjunto de la población de la que los beneficiarios forman parte forman parte.
- De la eficiencia (rentabilidad), distinguiendo entre rentabilidad económica y rentabilidad política y social. Identifican tres tipos de análisis:
 - Coste-efectividad o coste-eficacia.
 - Coste-utilidad.
 - Coste-beneficio.

Por otra parte, Aguilar y Ander-Egg (1994), establecen unos principios o requisitos para la investigación evaluativa:

- Principios que resultan de las exigencias del método científico:
 - Validez.
 - Fiabilidad.
- Principios de carácter práctico:
 - Practicidad y utilidad.
 - Oportunidad.

4.2. LA EVALUACIÓN DEL IMPACTO SOCIAL SEGÚN EL ENFOQUE PROPUESTO POR PICHARDO (1993)

Pichardo (1993), que se sitúa igualmente cercana al enfoque clásico-experimental, define impacto social como: “...los cambios o variaciones deseados en los destinatarios de las políticas (sociales o no), programas y/o proyectos...” (Pichardo, 1993, p. 73) “De tal manera, que... puede observarse en tres niveles básicos:

- a) Los destinatarios de las acciones evaluadas.
- b) El medio institucional en el cual se desarrollan...
- c) El contexto en el cual se inscriben...” (Pichardo, 1993, p. 74).

Esta autora concibe el impacto social como un enfoque de evaluación, una perspectiva metodológica, no únicamente como un criterio de evaluación, percepción parcial e incompleta. A su vez, se trata de una síntesis de cuatro enfoques generales de evaluación:

- a) La evaluación privada o análisis de rentabilidad, buscando la maximización del beneficio obtenido. En esta evaluación, Pichardo incluye dos indicadores: la tasa interna de retorno (TIR) o ganancia obtenida de la inversión y el valor neto actualizado (VAN), que asume que la rentabilidad obtenida disminuye con el tiempo.
- b) La evaluación social o Análisis de Costo / Beneficio (ACB) desde el punto de vista público, como apuntan Cohen y Franco (1992), no privado, es decir, teniendo en cuenta los intereses del estado.
- c) El Análisis de Costo / Eficiencia (ACE), similar al anterior, pero la eficiencia no se mide exclusivamente en términos económicos, según Cohen y Franco (1992).
- d) Evaluación experimental o comparación cuantitativa entre lo programado y lo ejecutado, respecto de los resultados obtenidos, mediante métodos experimentales o cuasiexperimentales. Este enfoque se basa en dos criterios fundamentales:
 - La eficacia: consecución de los objetivos.
 - La eficiencia: conseguir dichos objetivos con los menores recursos posibles.

En cuanto a los requisitos que debe cumplir la evaluación del impacto social, Picardo (1993) destaca 5:

- a) Validez.
- b) Confiabilidad.
- c) Practicabilidad.
- d) Objetividad.
- e) Oportunidad.

Por otra parte, considera que pueden darse varias formas de evaluación de impacto social:

- a) Externa a la organización.
- b) Interna, dentro de la organización.
- c) Mixta.
- d) Autoevaluación, realizada directamente por los responsables de la intervención.
- e) Participante, incluyendo a los destinatarios.

Según la naturaleza y los fines de la evaluación, la autora identifica dos tipos:

- a) Exploratoria – Descriptiva, cuando la finalidad es recopilar información, especialmente cuantitativa.
- b) Analítica, cuya finalidad es la interpretación de la información.

Por último, asume la evaluación “...como proceso integral y continuo...” (Picardo, 1993, p. 178). En este proceso se suceden tres momentos o fases de la intervención, y, por tanto, de la propia evaluación:

- a) La evaluación ex-ante, que centra su interés en “determinar la viabilidad de generar el impacto social deseado. Los resultados de esta evaluación constituyen un insumo fundamental para la etapa de la formulación de las decisiones” (Picardo, 1993, p. 179)

Los parámetros en la evaluación en esta etapa inicial son dos: el análisis de viabilidad y el de trascendencia social.

- b) La evaluación concurrente, enfocada en “examinar la organización de los medios disponibles que garanticen el logro del impacto social deseado.

De ahí que la supervisión y el seguimiento de las acciones ocupen un papel central en este momento evaluativo” (Pichardo, 1993, p. 179).

Los parámetros para la evaluación que se realiza durante la ejecución del proyecto son: el análisis del rendimiento o del progreso y el de trascendencia social.

- c) La evaluación ex-post, centrada en “comprobar el grado de cumplimiento efectivo y eficiente de los objetivos del impacto social programado”.

La evaluación que se realiza una vez finalizada la intervención es el “análisis del impacto propiamente dicho” (Pichardo, 1993, p. 348).

4.3. LA EVALUACIÓN DE PROGRAMAS Y SERVICIOS PÚBLICOS PROPUESTA POR BALLART (1992)

Ballart (1992) plantea la siguiente definición:

...la evaluación de programas tiene que ver tanto con los efectos de un programa sobre sus beneficiarios, como con su implementación, como con el proceso de formulación de propósitos generales y objetivos más específicos, lo que tendría que permitir relacionar, de acuerdo con el modelo teórico, resultados finales, con la forma en que es administrado el programa y planificada la intervención (p. 19).

Ballart (1992) se sitúa en un enfoque claramente integrador y basa su definición en el planteamiento de Rossi y Freeman (1985) para quienes la evaluación es “...la aplicación sistemática de procedimientos de investigación social para valorar la conceptualización y diseño, implementación y utilidad, de los programas de intervención social” (Rossi & Freeman, 1985, p. 20).

De las líneas resumidas en esta definición, Ballart (1992) extrae las categorías para su clasificación:

- Según su propósito, puede ser evaluación formativa o de proceso y evaluación conclusiva o de resultados.
- Teniendo en cuenta la naturaleza del estudio, distingue entre evaluación descriptiva y evaluación analítica.

- Según la relación de los evaluadores con la organización que desarrolla el proyecto, puede distinguirse entre evaluación interna y externa.
- Por último, incluye la meta-evaluación o síntesis de los resultados de varias evaluaciones.

A su vez, en lo que respecta al proceso evaluación, Ballart (1992) establece cuatro fases:

- a) Evaluación de la Conceptualización y del Diseño de los Programas. Se trata de un diagnóstico de las necesidades, de una valoración del diseño de la intervención y de su coherencia y racionalidad.
- b) Evaluación de la Implementación de los Programas. Consiste en "...una supervisión o control del funcionamiento del programa que indudablemente puede interesar a los gestores del programa que pueden tomar acciones de tipo correctivo en relación con problemas de gestión..." (Ballart, 1992, p. 76).
- c) Evaluación de la Eficacia o Impacto de los Programas. "Una evaluación de eficacia o impacto examina la medida en que éste produce un cambio en la dirección deseada..." "La cuestión esencial en la evaluación de impacto es, pues, separar los efectos del programa de otros efectos externos para determinar hasta qué punto el programa es responsable de los cambios observados durante y después de su puesta en práctica." (Ballart, 1992, p. 76). Para este autor, no hay diferencias entre la evaluación de impacto y evaluación de eficacia.
- d) Evaluación de la Eficiencia de los Programas. Implica una evaluación de eficacia previa. Según este autor, "...el impacto de un programa deberá ser juzgado en función de su coste ya que puede ocurrir perfectamente que un determinado programa está produciendo unos resultados escasos por los costes en que ha incurrido." (Ballart, 1992, p. 77).

4.4. LA EVALUACIÓN SEGÚN EL ENFOQUE DE LA UNESCO. SIEMPRE (1999)

El Sistema de Información, Monitoreo y Evaluación de Programas Sociales de la UNESCO, SIEMPRO (1999), desde un enfoque integrador igualmente:

...concibe la evaluación como un proceso permanente y continuo de indagación y valoración de la planificación, la ejecución y la finalización del programa social. Su finalidad es generar información, conocimiento y aprendizaje dirigidos a alimentar la toma de decisiones oportunas y pertinentes para garantizar la eficiencia, la eficacia y la calidad de los procesos, los resultados y los impactos de los programas... (p. 55).

Estos autores, SIEMPRO (1999). distinguen tres fases, con sus correspondientes instrumentos, que configuran lo que denominan la gestión integral de un programa social:

a) Planificación

- La evaluación ex-ante. Evaluación que se lleva a cabo antes de empezar la intervención y consiste en "...un análisis de la formulación del programa para conocer si los elementos operativos del mismo ... resultan consistentes..." e "...indaga la viabilidad de los componentes y de las actividades planificadas ... De esta manera, se busca racionalizar la inversión social." (p. 74). En resumen, consistencia y viabilidad.

b) Ejecución

- El monitoreo. Es un proceso permanente que se desarrolla a lo largo de toda la fase de ejecución que "...consiste en el análisis periódico ... con el objeto de establecer el grado en que las actividades de cada componente, el cronograma de trabajo, las prestaciones y los resultados se cumplen de acuerdo con lo planificado. Gracias a este control, se pueden detectar deficiencias e incongruencias y corregir o replantear oportunamente." (p. 87). En síntesis, posibilitar los ajustes.
- La evaluación diagnóstica. Es una evaluación puntual y profunda, que se lleva a cabo en un momento concreto, durante la ejecución, para "...detectar y analizar problemas y obstáculos con miras a lograr un impacto más significativo del programa y una mayor eficacia y

eficiencia en la utilización de recursos.” (p. 93). En resumen, anticipación de resultados y acciones de mejora.

- La evaluación desde la perspectiva de los beneficiarios. Es la llamada evaluación participativa y pretende incorporar “...la perspectiva de las poblaciones beneficiarias en el proceso de análisis de las fortalezas y las debilidades de los programas y de los impactos que generan en las condiciones de vida de esas poblaciones” (p. 97). Cousins y Whitmore (1998) diferencian dos tipos de evaluación participativa: la evaluación participativa práctica (*practical participatory evaluation*) y la evaluación participativa transformadora o emancipadora (*transformative participatory evaluation*).

c) Evaluación

- La evaluación de resultados finales. Aunque se planifica con anterioridad, se lleva a efecto en el momento de la finalización de la intervención y pretende “...indagar y analizar el cumplimiento de las metas previstas para cada uno de los objetivos específicos del programa (metas de resultado) en términos de su efectividad, su eficacia y su eficiencia, para obtener conclusiones...”. (p. 101). En síntesis, eficacia y eficiencia inmediata.
- La evaluación de impacto. Como se ha señalado más arriba, Aguilar y Ander-Egg (1994) basan la diferencia entre evaluación de resultados, cuando se focaliza en los destinatarios e impacto, cuando se tiene en cuenta el conjunto de la población. En contraste con este planteamiento, el SIEMPRO (1999), al igual que otros autores como Borus y Tash (1970), entiende que la evaluación de impacto es la que se realiza un tiempo después de finalizada la ejecución de la intervención. La evaluación de impacto es, según SIEMPRO (1999) “...el análisis que permite conocer si el programa social ha generado cambios relevantes en las condiciones de vida de la población beneficiaria.” (p. 106). Para la evaluación de impacto en concreto, el enfoque adoptado en este manual es el de la perspectiva clásica, tratando de establecer relaciones causales, por medio de la aplicación de diseños experimentales, cuasiexperimentales y no experimentales.

5. LA EVALUACIÓN. ELEMENTOS PARA UNA DEFINICIÓN

En el contexto teórico de integración de enfoques en el que nos encontramos en la actualidad, cuyo planteamiento se asume y se comparte en este trabajo, más que tratar de ofrecer una definición del concepto de evaluación de programas, que difícilmente va a aportar luz a la comprensión del término y que tampoco va a enriquecerlo, vamos a tratar de encontrar todos los componentes que integran el marco conceptual de la evaluación. Se trata de realizar una identificación de todos aquellos elementos que hemos ido viendo en las diferentes perspectivas teóricas y metodológicas anteriores y tratar de combinarlos y relacionarlos en un sistema.

La evaluación de la intervención es un *proceso*, lo que significa que se trata de una sucesión temporal de fases. Pero no hablamos de un proceso automático. Aunque puede y debe ser programado y en la medida de lo posible, sistemático, su desarrollo es complejo y de carácter variable y dinámico. Llevarlo a cabo requiere, sin duda, amplios conocimientos, pero también habilidades, es decir, se trata de una competencia profesional. ¿Estamos hablando de una disciplina científica, y, por tanto, de una categoría profesional emergente? Para contestar afirmativamente a esta pregunta, debemos identificar su objeto de estudio, pero fundamentalmente su método propio y, además, es preciso asegurarnos de que tiene una aplicación práctica.

Su campo de estudio, su objeto, algo que apuntaba Scriven y en lo que coinciden absolutamente todos los enfoques teóricos que acabamos de revisar, es indagar y e incluso calcular el valor, el mérito de una intervención social. Esto es *valoración*. No entraremos en el debate sobre valoración o evaluación, en este trabajo el concepto valoración se utiliza para referirse al mero hecho de valorar algo a la luz de la información hallada. Entenderemos evaluación como todo el proceso de investigación y valoración como una parte fundamental del mismo, que es, a su vez, el propio resultado final.

Siguiendo con la finalidad del proceso evaluador, otro de los factores comunes de todas las teorías es la búsqueda de una *relación causal* entre la

intervención y sus efectos. Los cambios conseguidos deben ser una consecuencia directa de la intervención y sólo de ella. Este es el punto nuclear para la búsqueda de un método. Lo que nos habilita para aproximarnos a esta relación causa-efecto es el grado de *validez*. Como en cualquier ciencia, renunciamos a la búsqueda de la verdad absoluta, por lo tanto, hablamos siempre en términos de probabilidad. La validez interna, es la que nos permite establecer que, en un programa concreto, los resultados son fruto de su implementación. La validez externa nos legitima para predecir que la misma intervención en otro lugar, en otro contexto, en distinto momento o con diferente población, produciría los mismos efectos, esto es generalización de las conclusiones. Si, establecida esta relación causal, los resultados obtenidos son los esperados, los deseados y se corresponden con las metas y los objetivos planteados, el programa tiene una característica, la *eficacia*. Si, además, se ha realizado con los mínimos recursos posibles, hablamos de *eficiencia*. Se trata de un concepto relativo, sólo se puede afirmar que una intervención es eficiente, comparándola con otras similares en objetivos.

Con todo, la evaluación es *información* organizada, que da lugar a *conocimiento científico*. La investigación evaluadora va más allá de la evaluación de un programa o una serie de ellos, persigue el avance de la ciencia y la construcción paulatina de un marco conceptual y relacional de la intervención social. Busca la formulación de teorías explicativas, de nuevas hipótesis y posibilita la generación de nuevas líneas de investigación.

En lo concerniente al método, sólo cabe uno: el *método científico*. Con las características y la pluralidad de formas que adopta la aplicación de este método en las ciencias sociales. Esto se traduce en la utilización de una metodología *cuantitativa*, en la que se establezcan mediciones, indicadores..., asegurándose de la *fiabilidad* de los instrumentos de medida, esto es consistencia en las mediciones, y una metodología *cualitativa*, mediante la que se interpretan las observaciones y la información extraída, numérica o no. Por ello, en el proceso de evaluación, se utilizan diseños experimentales en la medida de lo posible, cuasiexperimentales,

correlacionales y no experimentales y se recurre a técnicas estadísticas y otras como la observación directa, la entrevista, la participación de agentes implicados y grupos de interés, los grupos de discusión o la revisión documental.

Por último, en lo referente a la aplicación y a su utilidad práctica, característica imprescindible para cualquier disciplina y para todo campo de la ciencia, arriesgando, podemos sintetizar y agrupar las diferentes aportaciones en una afirmación: sirve para *tomar decisiones* y para *formular políticas e intervenciones*, se sitúe en el centro de atención la política, la gestión o los beneficiarios. Con todo ello, los resultados de los procesos de evaluación explican, con muchas limitaciones predicen y siempre arrojan luz para establecer e implantar medidas que contribuyan eficaz y eficientemente, a lo que en este trabajo entendemos que es el fin último de la intervención social: *solucionar problemas sociales*.

6. BIBLIOGRAFÍA

- Abramson, M., & Wholey, J. (1981). Organization and management of the evaluation function in a multilevel organization. *New Directions for Evaluation*, 1981(10), 31-48.
- Achúcarro, C., & Guillén, C. (1996). Aspectos históricos, teóricos y metodológicos de la evaluación de programas. . In *Intervención psicosocial: elementos de programación y evaluación socialmente eficaces*, 56-89.
- Aguilar, M., & Ander-Egg, E. (1994). *Evaluación de servicios y programas sociales*. Buenos Aires: Lemen.
- Alkin, M., & Christie, C. (2005). Unraveling theorists' evaluation reality. *Directions for Evaluation*, (106)., 111-128.
- Alkin, M., Patton, M., & Weiss, C. (1990). *Debates on evaluation*. Newbury Park: Sage.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Madrid: CIS Centro de Investigaciones Sociológicas.

- Anguera, M., & Chacón, S. (2008). Aproximación conceptual en evaluación de programas. En M. T. Anguera, S. Chacón, & A. Blanco-Villaseñor, *Evaluación de programas sanitarios y sociales. Abordaje metodológico* (p. 17-36). Madrid: Síntesis.
- Ballart, X. (1992). *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudios de caso*. Madrid: INAP.
- Ballart, X. (1996). Modelos teóricos para la práctica de evaluación de programas. En Brugué, Q. y Subirats, J., *Lecturas de gestión pública* (p. 321-351). Madrid: INAP. Obtenido de <http://scioteca.caf.com/bitstream/handle/123456789/1008/Version%20Web%20Evaluacion%20de%20politicas%20.pdf?sequence=1&isAllowed=y>
- Berk, R., Lenihan, K., & Rossi, P. (1981). Crime and Poverty: Some Experimental Evidence from Ex-Offenders. En H. E. Freeman, & M. A. Solomon, *Evaluation studies review annual, 6*. (p. 339-359). Beverly Hills, CA.: Sage.
- Bickman, L., & Rog, D. (2008). *The Sage handbook of applied social research methods*. Thousand Oaks, CA: Sage.
- Bingham, R., & Felbinger, C. (2002). *Evaluation in practice: A methodological approach*. New York: Chatham House Publishers.
- Borus, M., & Tash, W. (1970). *Measuring the Impact of Manpower Programs: A Primer*. Ann Arbor, Michigan: The University of Michigan-Wayne State University, Institute of Labor and Industrial Relations.
- Bouzas, R. (2005). Los caminos de la evaluación de políticas públicas: una revisión del enfoque. *Revista de investigaciones políticas y sociológicas, 2*, 69-86.
- Browne, A., & Wildavsky, A. (1987). What should evaluation mean to implementation. *The politics of program evaluation, 15*, 146.
- Bryk, A. S. (1983). *Stakeholder Based Evaluation*. San Francisco: Jossey-Bass.
- Bueno, C., & Osuna, J. (2012). La evaluación de políticas públicas en las ciencias sociales: entre el ser y el deber ser. *Prisma Social: revista de investigación social, 9*, 176-208.

- Bustelo, M. (1999). Diferencias entre evaluación e investigación: una distinción necesaria para la identidad de la evaluación de programas. *Revista Española de Desarrollo y Cooperación*, 4, 9-29.
- Campbell, D., & Boruch, R. (1975). Making the case for randomized assignment to treatments by considering the alternatives: Six ways in which quasi-experimental evaluations in compensatory education tend to underestimate effects. En C. A. Bennett, & A. A. Lumsdaine, *Evaluation and experiment: Some critical issues in assessing social programs* (p. 195-296). New York: Academic Press.
- Campbell, D., & Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Cohen, E., & Franco, R. (1992). *Evaluación de proyectos sociales*. México: Siglo XXI.
- Cousins, J., & Whitmore, E. (1998). Framing participatory evaluation. *New directions for evaluation*, 80, 5-23.
- Crane, J. A. (1988). Evaluation as Scientific Research. *Evaluation Review*, 12(5), 467-482.
- Cronbach, L., & Shapiro, K. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Chelimsky, E. (2006). The purposes of evaluation in a democratic society. *Handbook of evaluation: policies, programs and practices*. Obtenido de <https://search-proquest-com.ezproxy.uned.es/docview/36662947?accountid=14609om.ezproxy.uned.es/docview/36662947?accountid=14609>
- Chen, H., & Rossi, P. (1980). The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social Forces*, 59. Obtenido de <https://search-proquest-com.ezproxy.uned.es/docview/38276265?accountid=14609>
- Chen, H.-T. (1990). *Theory-driven evaluation*. Newbury Park, CA: Sage.
- De la Orden, A. (1990). Evaluación de los efectos de los programas de intervención. *Revista de investigación educativa*, 8(16), 61-76.

- De Miguel, M. (2000). La evaluación de programas sociales: fundamentos y enfoques teóricos. *Revista de Investigación Educativa*, 18(2), 289-217.
- Dunn, W. N. (1994). *Public Policy Analysis*. New Jersey: Prentice Hall.
- Eisner, E. W. (1985). *The art of educational evaluation: A personal view*. London: Taylor & Francis.
- Fernández-Ballesteros, R. (1992). El proceso como procedimiento científico y sus variantes. *Introducción a la evaluación psicológica*, 1.
- Fernández-Ballesteros, R. (1996). *Evaluación de programas: una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Fetterman, D. M. (1998). *Ethnography*. Thousand Oaks: Sage.
- Freeman, H., & Solomon, M. (1981). *Evaluation Studies Review Annual*. Beverly Hills: Sage.
- Fullan, M. (2005). *Leadership and sustainability: System thinkers in action*. Thousand Oaks: Corwin Press.
- Guba, E., & Lincoln, Y. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Hendricks, M. (1981). Service delivery assessment: Qualitative evaluations at the cabinet level. *New Directions for Evaluation*, 1981(12), 5-24.
- Hernández, M. (2002). Evaluación del desempeño de las organizaciones públicas a través de la calidad. *VII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública*. Lisboa.
- Hogwood, B., & Gunn, L. (1984). *Policy analysis for the real world*. Oxford: Oxford University Press.
- Horst, P., Nay, J. N., Scanlon, J. W., & Wholey, J. S. (1974). Program management and the federal evaluator. *Public Administration Review*, 34(4), 300-308.
- House, E. R. (1980). *Evaluating with Validity*. Beverly Hills, CA: Sage.
- Johnson, J., & Joslyn, R. (1995). *Political Science Research Methods*. Washington, DC: Congressional Quarterly.
- Kelly, R. M. (1998). An inclusive democratic polity, representative bureaucracies, and the new public management. *Public administration review*, 58(3), 201-208.

- King, J., Morris, L., & Fitz-Gibbon, C. (1987). *How to assess program implementation*. Thousand Oaks: Sage.
- Langbein, L., & Felbinger, C. (2006). *Public program evaluation: a statistical guide*. New York: M. E. Sharpe.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage.
- Lipsey, M., & Pollard, J. (1989). Driving toward theory in program evaluation: More models to choose from. *Evaluation and Program Planning*, 12(4), 317-328.
- Madaus, G., Stufflebeam, D., & Scriven, M. (1983). Program evaluation. *Evaluation models*, 3-22.
- Majone, G. (1989). *Evidence, argument and persuasion in the policy process*. New Haven, CN: Yale University Press.
- Martínez-Mediano, C. (1998). La teoría de la evaluación de programas. *Educación XX1*, 1. doi:<http://dx.doi.org/10.5944/educxx1.1.1.398>
- Monnier, E. (1995). *Evaluación de la acción de los poderes públicos*. Madrid: Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda.
- Nioche, J. P. (1982). De l'évaluation à l'analyse des politiques publiques. *Revue française de science politique*, 32(1), 32-61.
- Owen, J., & Rogers, P. (1999). *Program Evaluation: Forms and Approaches*. Thousand Oaks: Sage.
- Palumbo, D. J. (1987). *The politics of program evaluation*. Newbury Park, CA: Sage.
- Parlett, M., & Hamilton, D. (1972). Evaluation as illumination: A new approach to the study of innovatory programs. *Occasional paper*, 9, 55-73.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Pichardo, A. (1993). *Evaluación del impacto social: el valor de lo humano ante la crisis y el ajuste*. Buenos Aires: Lumen-Hvmanitas.
- Pollit, C. (1995). Justification by works or by faith? Evaluating the new public management. *Evaluation*, 1(2), 133-154.
- Preskill, H., & Catsambas, T. (2006). *Reframing evaluation through appreciative inquiry*. London: Sage.
- Prottas, J. (1979). *People-Processing: The Street-Level Bureaucrat in Public Service Bureaucracies*. Lexington, MA: Lexington Press.

- Provus, M. M. (1969). *The Discrepancy Evaluation Model, An Approach to Local Program Improvement and Development*. Washington, DC: Bureau of Research, Office of Education.
- Riecken, H., & Boruch, R. (1974). *Social experimentation: a method for planning and evaluating social intervention*. New York: Academic Press.
- Rossi, P., & Freeman, H. (1985). *Evaluation: A Systematic Approach*. Beverly Hills: Sage.
- Rossi, P., & Lyall, K. (1978). An overview evaluation of the NIT experiment. En M. Cook, K. Del Rosario, K. Henningan, M. Mark, & W. Trochim, *Evaluation studies review annual*, 3 (p. 412-428). Beverly Hills, CA: Sage.
- Rossi, P., & Wright, S. (1977). Evaluation research: An assessment of theory, practice, and politics. *Evaluation Quarterly*, 1(1), 5-52.
- Scriven, M. (1967). The Methodology of Evaluation. En R. Tyler, R. Gagne, & M. Scriven, *Perspective of Curriculum Evaluation. Vol 1* (p. 39-83). Chicago: Rand McNally.
- Scriven, M. (1973). Goal-Free Evaluation. En E. R. House, *School Evaluation. The Politics and Process* (p. 319-328). Berkeley, CA: Mc Cutehan.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Shadish, W., & Reichardt, C. (1987). The intellectual foundations of social program evaluation: The development of evaluation theory. *Evaluation studies review annual*, 12, 13-29.
- Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Shortell, S., & Richardson, W. (1978). *Evaluation designs. Health program evaluation*. St Louis: Mosby.
- SIEMPRO. Gestión Integral de Programas Sociales Orientada a Resultados. (1999). *Manual Metodológico para la Planificación y Evaluación de Programas Sociales*. Brasil: UNESCO, Fondo de Cultura Económica.
- Smith, E., & Tyler, R. (1942). *Appraising and recording student progress*. Oxford: Harper.
- Smith, N., & Brandon, P. (2008). *Fundamental issues in evaluation*. New York NY: Guilford Press.
- Stake, R. (2004). *Standards-based and responsive evaluation*. Thousand Oaks, CA: Sage.

- Stake, R., & Easley, J. (1978). Case Studies in Science Education. *Evaluation Report Series, Vol I.*
- Stake, R., & Gjerde, C. (1975). An Evaluation of TCITY: The Twin City Institute for Talented Youth. *Evaluation Report Series, 7.*
- Stufflebeam, D., & Schikfield, A. (1987). *Evaluación Sistemática. Guía teórica y práctica.* Madrid: Paidós/MEC.
- Suchman, E. A. (1968). *Evaluation research: Principles and practice in public service and social action programs.* New York: Russell Sage Foundation.
- Tyler, R. W. (1986). Changing Concepts of Educational Evaluation. *International Journal of Educational Research, 10(1).* doi:[https://doi.org/10.1016/0883-0355\(86\)90008-X](https://doi.org/10.1016/0883-0355(86)90008-X)
- Weiss, C. (1972). *Evaluation research.* Englewood Cliffs, NJ: Prentice-Hall.
- Weiss, C. H. (1975). *Investigación evaluativa: Métodos para determinar la eficiencia de los programas en acción.* México: Trillas.
- Weiss, C. H. (1998). *Methods for studying programs and policies.* London: Prentice Hall.
- Weiss, C., & Bokuvalas, M. (1980). *Social science research and decision-making.* New York: Columbia University Press.
- Weiss, R., & Rein, M. (1970). The evaluation of broad-aim programs: Experimental design, its difficulties, and an alternative. *Administrative Science Quarterly, 15(1),* 97-109.
- Wholey, J. S. (1979). *Evaluation: Promise and Performance.* Washington, DC: Urban Institute.
- Wholey, J. S. (1983). *Evaluation and Effective Public Management.* Boston: Urban Institute.
- Wholey, J.S.; Abramson, M. A. y Bellavita, C. (1986). *Performance and credibility: developing excellence in public and nonprofit organizations.* Lexington, MA: Lexington.
- Wolf, R. M. (1979). *Evaluation in education: Foundations of competency assessment and program review.* New York: Pagers Publishers.