

anales de psicología

Número Monográfico, 1989, 5, 1-12

©Secretariado de Publicaciones e Intercambio Científico.
Universidad de Murcia.

Evaluación de Programas: Aspectos Básicos

Manuel Ato García¹
Elena Quiñones Vidal
Agustín Romero Medina
Rafael Rabadán Anta

Universidad de Murcia

¹*Dirección de los autores:* Departamento de Metodología y Análisis del Comportamiento. Facultad de Filosofía, Psicología y Ciencias de la Educación. Universidad de Murcia. 30071 MURCIA.

anales de psicología

Número Monográfico, 1989, 5, 1-12

Abstract

We describe firstly the essential features of program evaluation and propose after a general methodological model for planning, monitoring and analysis of psychological and social interventions.

Resumen

En este trabajo se ofrece una introducción conceptual a la evaluación de programas, describiendo primero los rasgos esenciales que caracterizan la evaluación de programas para proponer a continuación un modelo metodológico general que permita valorar la planificación, administración y análisis de programas de intervención psicológicos y sociales.

Keywords

Program Evaluation, Estimating the Effects of Social Interventions, Experimental and Quasi-Experimental Methodology.

Palabras-clave

Evaluación de Programas, Estimación de los Efectos de Intervenciones Psicosociales, Metodología Experimental y Cuasi-Experimental.

1 Introducción

Una preocupación constante de las sociedades económicamente más avanzadas se centra en incrementar, hasta donde permiten sus propios recursos, el bienestar individual, familiar y social de sus ciudadanos.

En el caso de la sociedad española, con un volumen de ingresos y gastos de alrededor de doce billones de pesetas, donde el gasto público se sitúa en torno al 62.5%, los gastos derivados de la política económica alcanzan el 20% y las transferencias a Comunidades Autónomas para sufragar los cada vez más costosos servicios regionales y locales suponen casi el 9%, es notorio y evidente el interés del Estado por mejorar la calidad de vida individual y social del individuo.

Y aunque un análisis superficial de las diversas partidas que integran este enorme gasto público revela que, en su mayor parte, constituyen "gastos básicos" destinados a financiar un servicio público económicamente ruinoso pero políticamente indiscutible, un análisis en profundidad permitiría desvelar una gran cantidad de irracionalidades en el gasto, que se hacen tanto más patentes cuanto mayor es el volumen económico considerado.

Globalmente considerados, los recursos públicos que el Estado destina a mejorar la calidad de la vida individual y social de los ciudadanos se destinan a financiar lo que, desde la perspectiva de las Ciencias Sociales, se ha dado en llamar **programas de acción**, que podrían definirse como "... *una actividad o serie de actividades encaminadas a cambiar el estatus psicológico, social, económico o educacional de un individuo o grupo de individuos*" (Craig y Metze, 1979, p. 210). El impacto global de un programa de acción depende del interés que suscite en el ciudadano y de las repercusiones de índole económica, social y psicológica que de él puedan derivarse. En este sentido, constituyen programas de acción todas aquellas intervenciones sociales y comunitarias (véase a este respecto la reciente revisión de Gesten y Jason, 1987) que el Estado realiza en la sociedad y presentan una amplísima variedad que abarca desde programas de índole nacional (por ejemplo, los grandes programas educativos, sanitarios, de defensa, de desempleo, etc.) a programas de índole regional e incluso local, donde el alcance de la intervención no supera los límites de una Comunidad Autónoma o un determinado Ayuntamiento.

Ante el considerable volumen económico generado por la administración de un programa de acción, la **evaluación del programa** ha terminado convirtiéndose en un componente universalmente aceptado de aquél.

En este trabajo ofrecemos una introducción conceptual a la práctica de la evaluación de programas, con el propósito de pergeñar tanto un marco general comprensivo del proceso de evaluación en todos sus elementos cuanto un modelo básico del procedimiento metodológico que se encuentra a la base de una intervención psicológica, social o económica. Para ello, se trazan en primer lugar los rasgos definitivos esenciales de la evaluación de programas para proponer a continuación un modelo metodológico general del proceso.

2 ¿Qué es Evaluación de Programas?

La evaluación de programas es una actividad aplicada, en gran medida a-teórica, y multidisciplinaria que abarca prácticamente el ámbito de las denominadas Ciencias Sociales. Algunas definiciones seleccionadas permitirán entrever en qué consiste tal actividad:

"La evaluación de programas es la medición de las consecuencias deseables o indeseables de una acción psicosocial ejecutada para facilitar la obtención de metas perseguidas por el sujeto." (Riecken, 1972, p. 86).

"La evaluación de programas es la determinación de la medida en que un programa logró uno o más de sus objetivos, las razones por las cuales no los alcanzó y la relación existente entre los efectos del programa y una amplia diversidad de variables y características del programa." (Perloff, Perloff y Sussna, 1976, p. 570).

"La evaluación de programas es un proceso sistemático y continuo que proporciona información acerca del efecto o impacto de un programa con el objeto de contribuir a adoptar una decisión política de trascendencia." (Spaniol, 1975, p. 2).

"La evaluación de programas es un modelo racional de solución de problemas de carácter psicológico y/o social que requiere que: (1) los problemas se definan primero con claridad y después ... (2) se genere una amplia variedad de soluciones potenciales ... (3) algunas de las cuales se implementen y ... (4) finalmente se evalúen. (5) El conocimiento de las soluciones positivas se presentan a los responsables políticos quienes ... (6) lo utilizarán para tomar las decisiones políticas que correspondan." (Cook y Shadish, 1986, p. 197).

La conclusión general es la misma cualquiera que sea la definición de partida: la evaluación de programas es un proceso que consiste en valorar si un programa consiguió o no alcanzar los objetivos para cuya consecución fue ideado.

Tal proceso consta de al menos cuatro actividades básicas, a saber (véase Ato, 1988):

- identificar los objetivos del programa;
- transformar los objetivos en variables observables;
- recoger datos empíricos relativos a las variables identificadas;
- comprobar los objetivos logrados con los datos empíricos obtenidos, valorando estadísticamente el éxito o fracaso en su consecución y el balance económico del binomio coste/eficacia.

Aunque las cuatro actividades citadas poseen sus propias peculiaridades y plantean problemas técnicos que el evaluador debe resolver (véase Hanrahan y Reid, 1983), el énfasis metodológico del proceso se centra en las dos últimas actividades, que conciernen respectivamente a la medida de las variables implicadas (Lipsey, 1983) y al tipo de análisis estadístico que permita concluir si el efecto de la intervención fue o no positivo, una decisión que dependerá fundamentalmente del tipo de metodología empleada (experimental versus no experimental) y de la forma de resolver el denominado *sesgo de selección* (véase Saxe y Fine, 1983; Leamer, 1978; Mark y Cook, 1984; Heckman y Robb, 1985; Heckman y Hotz, 1989).

Sin embargo, considerado el proceso desde una perspectiva global, la evaluación de programas significa algo más que una mera recopilación y tratamiento de datos empíricos. Aunque estos aspectos deben adecuadamente ponderarse en el informe técnico que acompaña a la evaluación, ya que de ellos dependerá fundamentalmente la credibilidad y alcance del resultado obtenido, en la prioridad debe centrarse en una interpretación de los resultados que contemple la globalidad de los aspectos implicados en el programa, y fundamentalmente en su valoración política, social y económica.

Scriven (1967, p. 39-83) trazó una interesante distinción que se ha hecho clásica entre dos diferentes filosofías de la evaluación de programas en base a los objetivos perseguidos en los curricula educativos: la *evaluación formativa* produce información que se retroalimenta durante las diferentes fases del desarrollo del curriculum con el fin de contribuir a su mejora, y por tanto permite ir adaptando determinados aspectos del programa en función de los resultados obtenidos al final de cada fase; la *evaluación sumativa*, por el contrario, tiene lugar cuando el curriculum se ha completado y por ende produce información acerca de la eficacia para alcanzar los objetivos inicialmente formulados. Ambas filosofías no sólo no se contraponen sino que actualmente se entienden como dos aspectos complementarios del proceso de investigación evaluativa.

La evaluación de programas surgió como proceso formal en las sociedades democráticas más avanzadas y es una actividad técnica que requiere un equipo multidisciplinario de psicólogos, sociólogos y economistas; los dos primeros como expertos en métodos de medición, diseño y análisis de variables psico-sociales y los últimos como expertos en la valoración económica de la implementación del programa mediante el correspondiente análisis del binomio coste/eficacia.

3 Un modelo de diseño, administración y análisis de programas

Aunque existen muchos diferentes modelos de evaluación de programas (véase por ejemplo, Scriven, 1967; Weiss, 1972; Wortman, 1975; Cronbach, 1979, 1982; Rafter, 1984) el modelo presentado y desarrollado por Wortman (1975, 1983) ofrece, a nuestro entender, una visión global (pero suficientemente detallada) del proceso con la

suficiente flexibilidad como para permitir ser adaptado a las más diversas situaciones. Wortman concibe la evaluación de programas como una serie de actividades de retroalimentación interconexas que ocurren en función de la interacción de tres componentes básicos: las unidades de organización, el contexto conceptual y los procesos propiamente evaluativos.

Las UNIDADES DE ORGANIZACIÓN son todos aquellos elementos (individuos o grupos) que pueden interactuar durante la administración y evaluación de un programa e incluye, en un sentido muy amplio, la Sociedad donde se supone que radica el interés por remediar o mejorar algún problema psicológico, social o económico preocupante y, en un sentido más restringido, el organismo financiador, el equipo profesional, el coordinador del programa, el equipo de evaluación y uno (o más) grupo(s) seleccionado(s) de individuos que de algún modo u otro participan del mismo en algún momento de su desarrollo. En el esquema general del modelo (presentado en la figura 1, adaptada de Wortman, 1975, con algunas modificaciones), todas las unidades de organización ocupan la primera columna del mismo. Obsérvese el papel crucial que juegan el organismo financiador, como promotor o impulsor de un programa, a cuya responsabilidad queda supeditada tanto la selección de un problema y su traducción en un programa de actuación como la elección de un coordinador, la de un equipo profesional competente y la de un equipo de evaluación.

Por su parte, la segunda columna del esquema está constituida por el CONTEXTO CONCEPTUAL que se encuentra a la base del programa. Varios aspectos se engloban dentro de este contexto.

- En primer lugar, los **valores sociales** desempeñan un papel esencial en el proceso porque imponen la adopción de determinadas estrategias políticas y son hasta cierto punto inmunes al proceso propio y a los resultados (al menos a corto plazo) de la propia evaluación. Por ejemplo, un "programa de acción" puede adoptarse en base a una de dos estrategias políticas: la creación de un servicio o la administración directa de recursos. En el primer caso, se crea una organización que ofrece a un grupo de individuos una serie de servicios que le ayudarán a superarse a sí mismos (programas de adiestramiento o capacitación, por ejemplo); en el segundo, se proporciona directamente a los participantes los recursos necesarios para utilizarlos a su libre albedrío. En cualquier caso, los valores sociales establecen las condiciones ideológicas dentro de las cuales se implementará y evaluará el programa (véase Edwards, Guttentag y Snapper, 1975, para una detallada explicación de la importancia de los valores sociales en el diseño, implementación y evaluación de programas).
- En segundo lugar, una vez determinadas las condiciones del problema y la estrategia política correspondiente dentro de la cual se aplicará, en función de los valores sociales imperantes, el organismo financiador los convertirá en **objetivos concretos** mediante la aplicación de alguna filosofía de la acción, o conjunto de principios que representan una actitud o un modo de pensar acerca de los hechos sobre los que se pretende intervenir (véase a este respecto Washington, 1980).

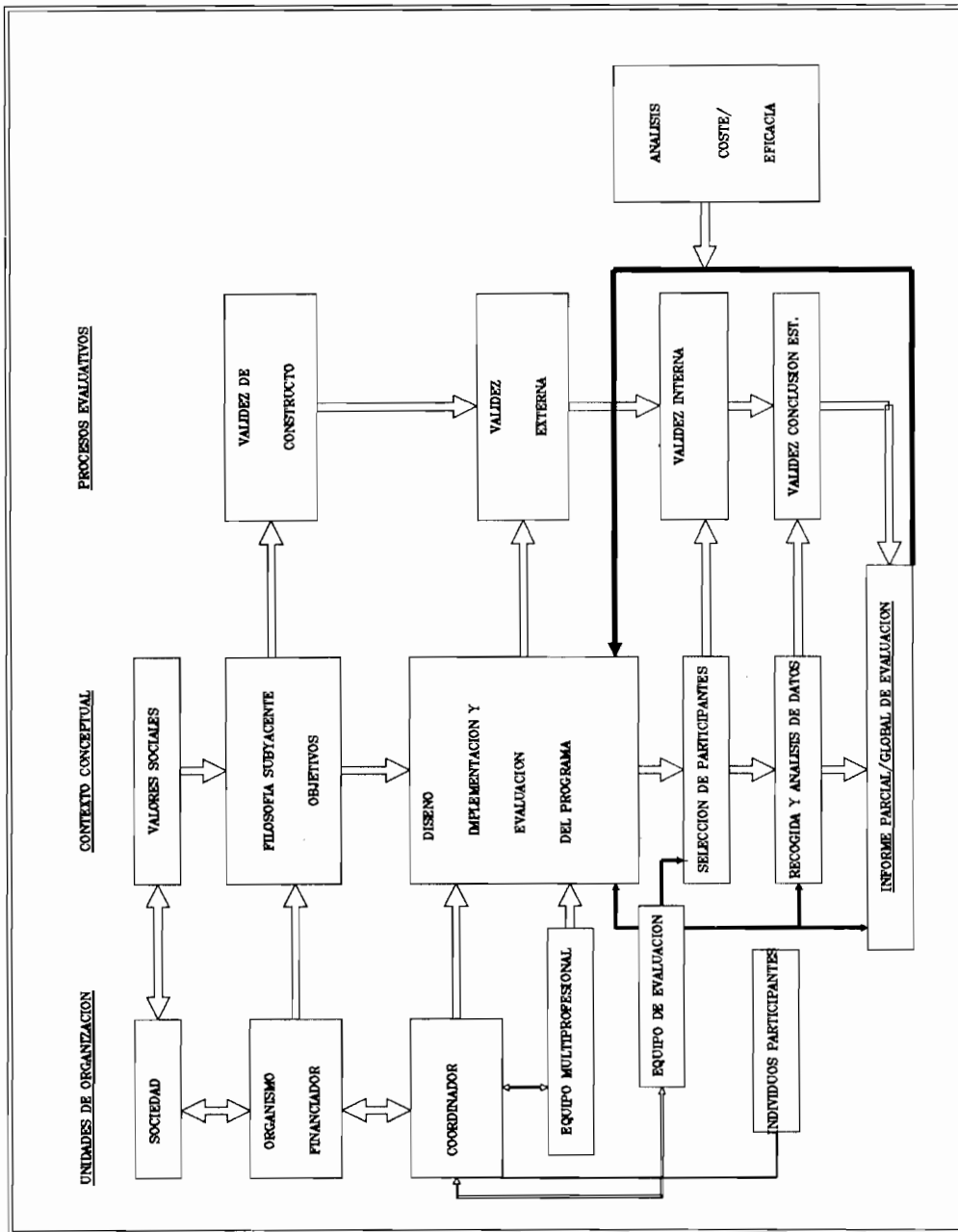


Figura 1: Un modelo de diseño y análisis de programas.

En la mayoría de las ocasiones, el organismo financiador carece de información relevante para traducir en objetivos un problema determinado. Se hace entonces necesario realizar una evaluación inicial del problema, cuyo conocimiento pueda permitir a los responsables políticos la especificación de los objetivos generales que se pretenden conseguir para transformar la problemática psicosocial objeto de la intervención. No obstante, la transformación de un programa en una serie de objetivos implica necesariamente un proceso de "operacionalización" de conceptos y variables que debe incluir qué comportamientos deben ser observados y/o medidos y especificar bajo qué condiciones se alcanzarán los objetivos propuestos.

- En tercer lugar, una vez identificados y operativamente definidos los objetivos del programa, el paso siguiente consiste en **diseñarlo** primero, **administrarlo** después y finalmente **evaluarlo**. Por regla general (véase Rossi y Wright, 1984) el proceso de diseño supone una puesta en común del coordinador del programa, en representación de los intereses del organismo financiador, el equipo profesional que administrará el programa, y el equipo de evaluación, cada uno de los cuales deberá aportar sus propios puntos de vista para fijar una serie de actividades a realizar en orden a lograr los objetivos propuestos y a especificar el modo como se realizará la correspondiente evaluación.
- Y finalmente, un conjunto de actividades que dependen directamente del equipo de evaluación, aunque se pergeñan durante el diseño del programa, son la **selección del contexto** situacional y de los **individuos** participantes y la **elección de los métodos** que se emplearán en la **recopilación** y el **análisis** de los datos. El primer aspecto tiene interés ante el deseo de buscar la mayor representatividad posible de la población a la cual se destina el programa y de generalizar el posible impacto de aquél al mayor número de individuos posible. El segundo aspecto depende en general del primero y su espectro de posibilidades es amplísimo, abarcando desde la utilización de diseños de observación natural (véase Anguera, 1978) al diseño experimental auténtico tanto univariable (véase Riecken y Boruch, 1974; Saxe y Fine, 1983; Arnau, 1986; Kish, 1987) como multivariable (Arnau, 1989 -en prensa-) y al diseño cuasiexperimental (Cook y Campbell, 1979; Trochim, 1984; Achen, 1986) e incluso los diseños de investigación cualitativa (véase Rafter, 1984; Anguera, 1986).

La tercera columna del diagrama del modelo de Wortman que comentamos, bajo la denominación general de **procesos evaluativos**, contiene todos aquellos procedimientos conectados con la valoración de las diversas etapas del programa, y se ha tomado del esquema de evaluación sugerido por Cook y Campbell (1976, 1979; véase también Brinberg y McGrath, 1985; Krathwohl, 1985).

- La **validez de la conclusión estadística** designa un conjunto de preocupaciones asociadas con la elección de los procedimientos de recogida y análisis de datos más adecuados a las circunstancias del programa. los "programas de

acción” por lo general suelen presentar un impacto que en la mayoría de los casos puede parecer irrelevante (la magnitud del efecto es baja) o demorarse excesivamente en el tiempo. Es preciso cerciorarse de que se han empleado instrumentos lo suficientemente potentes para detectar la presencia de alguna mejora, por pequeña que sea, y para diferenciar el componente de variación estrictamente sistemático, presentado o no de forma acumulada en el tiempo, supuestamente debido al programa, del componente de variación estrictamente aleatorio.

- La **validez interna** refleja asimismo un conjunto de preocupaciones que permiten eliminar la confusión de las actividades propias del programa, cuyo impacto pretendemos evaluar, con otras actividades extrañas a aquél, que de un modo u otro pudieran producir una influencia relevante sobre los participantes. Más concretamente, supone un intento de cerciorarse de que el efecto puede atribuirse con absoluta seguridad al programa en sí, y no a actividades paralelas o incluso asociadas al programa. De ahí que sea crucial para la evaluación tanto la selección de los participantes como la del contexto donde se implementará el programa.
- La **validez externa** representa la capacidad para generalizar los resultados, tanto a otros participantes diferentes de los incluidos como a otros contextos diferentes al utilizado, como a otros momentos en el tiempo. De tal interés puede derivarse la adopción de todo un conjunto de cruciales decisiones políticas.
- La **validez de constructo** indica, finalmente, la valoración de la relación existente entre el comportamiento tal y como fue observado y medido y uno (o más) conceptos teóricos o constructos que se suponen que son el fundamento o la causa del comportamiento evaluado. Y puesto que tales constructos son en general los mismos que han marcado la traducción empírica de los objetivos del programa, el equipo de evaluación debe valorar el grado de vinculación existente entre el contexto teórico-filosófico y la observación y/o medida de los comportamientos objeto de evaluación.

Como se contempla en la Figura 1, un INFORME DE EVALUACIÓN deberá ser redactado al final de cada una de las fases de implementación del programa. Tal informe debe contener todo un conjunto de interpretaciones acerca de los datos empíricos recabados a lo largo de cada fase, junto con el oportuno diseño y análisis de datos, los procesos de medición y operacionalización de variables y una detallada valoración económica del binomio coste/eficacia (Levin, 1983). No obstante, y puesto que el efecto o impacto de un programa de acción no suele presentarse de forma inmediata ni alcanzar una magnitud suficiente, los informes de evaluación tienen un carácter transitorio y se utilizarán como guía de la implementación del programa hasta emitir un informe final que recapitulará todo el proceso de diseño, implementación, análisis y evaluación del programa correspondiente.

4 Referencias bibliográficas.

- Achen, C. H. (1986). *The Statistical Analysis of Quasi-experiments*. Berkeley, CA: University of California Press.
- Anguera, M.T. (1978). *Metodología de la Observación en las Ciencias Humanas*. Madrid: Cátedra.
- Anguera, M.T. (1986). Posibilidades de la metodología cualitativa versus cuantitativa. *Revista de Investigación Educativa*, 3(6), 127-144.
- Arnau, J. (1986). *Diseños Experimentales en Psicología y Educación, Vol I*. Méjico: Trillas.
- Arnau, J. (en prensa). *Diseños Experimentales Multivariables: Alternativa Analítica a la Investigación Psicológica y Pedagógica*. Méjico: Trillas.
- Ato, M. (1988). *La Evaluación de Programas*. Informe técnico no publicado presentado a la Conserjería de Servicios Sociales. Murcia.
- Brinberg, D. y McGrath, J. E. (1985). *Validity and the Research Process*. Beverly Hills; CA: Sage Publications.
- Cherulnik, D. D. (1983). *Behavioral Research: Assessing the Validity of Research Findings in Psychology*. New York, NY: Harper and Row.
- Cook, T. D. y Campbell D. T. (1976). The design and conduct of quasi-experiments and true experiments in field settings. En M. Dunnette (ed.): *Handbook of Industrial and Organizational Psychology*, p. 223-326. Skokie: IL: Rand McNally.
- Cook, T. D. y Campbell, D. T. (1979). *Quasixperimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin.
- Cook, T. D. y Shadish, W. R. (1986). Program evaluation: the worldly science. *Annual Review of Psychology*, 37, 193-231.
- Craig, J. R. y Metze, L. P. (1979). *Methods of Psychological Research*. Boston, MA: W.B. Saunders (Trad. castellana: Interamericana, 1982).
- Cronbach, L. J. et al. (1979). *Toward Reform of Program Evaluation: Aims, Methods and Institutional Arrangements*. San Francisco, CA: Jossey-Bass.
- Cronbach, L. J. (1982). *Designing Evaluations of Educational and Social Programs*. San Francisco, CA: Jossey-Bass.
- Edwards, W., Guttentag, M. y Snapper, K. (1975). *Handbook of Program Evaluation* (2 vols.). Beverly Hills, CA: Sage.

- Gesten, E. L. y Jackson, L. A. (1987). Social and community interventions. *Annual Review of Psychology*, 38, 427-459.
- Hanrahan, P. y Reid, W. J. (1984). Choosing effective interventions. *Social Service Review*, 3(June), 244-58.
- Heckman, J. J. y Hotz, J. (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training (with comments). *Journal of American Statistical Association*, 84, 862-80.
- Heckman, J. J. y Robb, R. (1985). Alternative methods for evaluating the impact of interventions. En J.J. Heckman y B. Singer (eds.): *Longitudinal Analysis of Labor Market Data*, p. 156-245. New York, NY: Cambridge University Press.
- Judd, C. M. y Kenny, D. A. (1981). *Estimating the Effects of Social Interventions*. Cambridge, MA: Cambridge University Press.
- Kish, L. (1987). *Statistical Design for Research*. Nueva York, NY: Wiley.
- Leamer, E. (1978). *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York, NY: Wiley.
- Levin, H.M. (1983). *Cost-effectiveness: a primer*. Beverly Hills, CA: Sage.
- Lipsey, M. W. (1983). A scheme for assessing measurement sensitivity in program evaluation and other applied research. *Psychological Bulletin*, 94, 152-65.
- Mark, M. H. y Cook, R. D. (1984). Design of randomized experiments and quasiexperiments. En L. Rutman (ed.): *Evaluation Research Methods: a Basic Guide*, p. 65-119. 2nd Edition. Beverly Hills, CA: Sage Publications.
- Perloff, R., Perloff, E. y Sussna, E. (1976). Program evaluation. *Annual Review of Psychology*, 27, 569-594.
- Rafter, D.O. (1984). Three approaches to evaluation research. *Knowledge*, 6(2), 165-185.
- Riecken, H.W. (1972). *Social experimentation: a method for planning and evaluating social interventions*. Nueva York: Academic.
- Rossi, P.H. y Freeman, H.E. (1985). *Evaluation: a systematic approach*. 3rd Edition. Beverly Hills, CA: Sage.
- Rossi, P.H. y Wright, J.D. (1984). Evaluation research: an assessment. *Annual Review of Sociology*, 10, 331-352.
- Saxe, L. y Fine, M. (1983). *Social Experiments: Methods for Design and Evaluation*. 2nd Edition. Beverly Hills, CA: Sage Publications.

- Scriven, M. (1967). The methodology of evaluation. En R. Tyler et al. (Eds.): *Perspectives on curriculum education*. Chicago: Rand McNally.
- Trochim, W. M. K. (1984). *Research Design for Program Evaluation: The Regression-Discontinuity Approach*. Beverly Hills, CA: Sage Publications.
- Trochim, W. M. K. (1986). *Advances in Quasi-experimental Design and Analysis*. San Francisco, CA: Jossey-Bass.
- Washington, R.O. (1980). *Program evaluation in the human services*. Washington, DC: University Press of America.
- Weiss, C.H. (1972). *Evaluation Research*. Englewood Cliffs, NJ: Prentice-Hall. (Trad. castellana: Trillas, 1980).
- Wortman, P. (1975). *Evaluation research: a psychological perspective*. *American Psychologist*, 30, 560-569.
- Wortman, P. (1983). *Evaluation research: a methodological perspective*. *Annual Review of Psychology*, 34, 223-240.