

La Metodología Observacional en Evaluación de Programas

Observational Methodology in program Evaluation

M. Teresa Anguera-Argilaga¹
Universidad de Barcelona

RESUMEN

La abundante literatura existente sobre evaluación de programas pone de manifiesto el creciente interés que ha despertado en los últimos años. Entre los múltiples aspectos que se están introduciendo y desarrollando destacan las cuestiones metodológicas, que son probablemente las que han originado una importante laguna en este ámbito. La decisión acerca de cuál o cuáles sean las metodologías a utilizar nos obliga a delimitar las características, requisitos, recursos y aplicaciones de cada una de ellas, y, en este trabajo, específicamente de la metodología observacional. La metodología observacional, precisamente por ser inherente a situaciones en que entran en juego un elevado número de variables, y por deber garantizar la naturalidad de la situación adquiere una complejidad que implica un desarrollo adecuado, partiendo de unos criterios taxonómicos iniciales, y a lo largo de sus fases, culminando en la fase de análisis de datos.

Palabras clave: Metodología observacional, evaluación de programas.

ABSTRACT

The great amount of specific literature about program evaluation shows us a growing interest during later years. There are many aspects that being included and developed, and mainly some methodological questions; its traditional neglectance has originated an important problem to solve. The decision about which is the best methodology (or some joint methodologies) force us to delimitate the characteristics, conditions, resources and applications of each of them, and in this article, stressing observational methodology. Observational methodology is inherent to situations with an important number of natural variables. For this reason adequate develop-

¹ Quisiera expresar mi más profundo agradecimiento a Carlos Santoyo, por sus sugerencias y comentarios críticos, realizados durante su estancia en la Universidad de Barcelona. Para toda correspondencia, escribir a la autora: Departamento de Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Zona Universitaria de Pedralbes, 08028, Barcelona, España.

ment, from some initial taxonomic criteria until the data analysis, it requires of an through all phases.

Key words: Observational methodology, program evaluation.

1. PLANTEAMIENTO BÁSICO EN LA EVALUACIÓN DE PROGRAMAS

En los últimos años la evaluación de programas se halla en pleno proceso de expansión, como lo demuestra por una parte su rápido crecimiento y el interés despertado, tanto por las implicaciones metodológicas que ello va a comportar, como por parte de los profesionales que en ámbitos sumamente heterogéneos se hallan con la necesidad de analizar la medida en que un determinado plan de intervención genera unos resultados.

La abundante aparición de trabajos especializados en la literatura científica sobre evaluación de programas (manuales, artículos, ponencias en Simposiums, etc.) y el indudable esfuerzo que se ha desplegado en este mismo sentido a nivel institucional (organización de Congresos nacionales e internacionales, constitución de sociedades, reuniones de representantes de diversos países, etc.), constituyen un buen exponente del "estado de la cuestión" y de la voluntad existente por impulsarlo de forma creciente. No en vano se ha pronosticado por parte de autores de gran prestigio (Wertheimer, Barclay, Cook, Kiesler, Koch, Riegel, Rorer, Senders, Smith & Sperling, 1978) que la evaluación de programas constituye una de las áreas de trabajo en Psicología que más se van a expandir en vísperas del segundo milenio de nuestra era, y que el desarrollo de nuevas técnicas precisas y caracterizadas por su flexibilidad van a constituir uno de los desarrollos más prometedores en el futuro.

La principal cuestión que hasta ahora ha constituido un tema de debate es el propio concepto de evaluación de programas, que no es monolítico probablemente por la influencia diversificada de multitud de académicos de diferente formación y de profesionales que han pretendido ante todo resolver de forma empírica las situaciones con las que se enfrentaban. Existen aún diferentes concepciones perfectamente caracterizadas en su origen, pero que en la actualidad cada vez tienden más a confluir e incluso a solaparse. Entre ellas destacamos algunas:

a) Valoración de actividades como un "servicio dedicado a ..." (Cordray, Bloom & Light, 1987). Por ejemplo, en niños procedentes de hogares desestructurados, mujeres maltratadas, ancianos desvalidos que viven solos, alfabetización de internos de instituciones penitenciarias, etc. Prácticamente en todas y cada una de estas situaciones destaca la enorme complejidad que comportan, y en consecuencia la necesaria sistematización que se impone—aunque frecuentemente es muy deficitaria— para abordarlo con un mínimo de precisión y objetividad. Tradicionalmente, ha interesado mucho más su carácter asistencial que un análisis reflexivo sobre cómo se han programado los planes de acción.

b) Son bastantes los autores (Rossi & Freeman, 1985) que consideran la evaluación de programas como resultante de la mera aplicación "desde fuera" de diversas técnicas de investigación social que permitirán conocer el resultado de la implementación y uso de programas dirigidos a determinados colectivos de la población. Por lo general se persigue una objetividad en la obtención y tratamiento de la información, aunque los objetivos se hallan un tanto desvirtuados, debido en muchos casos a la indefinición, excesiva amplitud, o falta de concreción de sus objetivos. Por ejemplo respecto a los "slogans" de "Salud para todos", o "Ningún niño sin parque para jugar".

c) Aunque no tan minoritaria como pudiera preverse, existe igualmente una conceptualización de la evaluación de programas desde una perspectiva de prioridades políticas (Palumbo, 1987), atendiéndose en muchas ocasiones a razones de oportunidad. Así, ¿será más conveniente implementar un programa tendiente a eliminar la mendicidad existente en las calles de una ciudad, o bien se utilizarán estos recursos para aumentar las escuelas?

d) Existen también autores que han trabajado sobre un análisis y contraste de modelos teóricos en evaluación. Propiamente quedarían fuera del ámbito que nos ocupa, ya que su labor no implica una propuesta de nuevas conceptualizaciones, sino más bien de sistematización de las existentes; no obstante, con ello se favorece un análisis meta-evaluativo que sí juzgamos interesante porque revierte un espíritu de autocritica sumamente saludable (Stufflebeam, 1978).

En esta diversidad de enfoques juega también un importante papel el carácter interdisciplinar de la evaluación de programas, en el que confluyen psicólogos, trabajadores sociales, sociólogos, economistas, etc., lo que supone la adopción de posiciones muchas veces distintas que no siempre son fáciles de conjuntar en la elaboración, implementación, monitorización y análisis de los efectos de un programa.

De ahí que exista una indudable necesidad de estructuración en este ámbito, en el que deben fijarse las coordenadas básicas para poder incidir en aquellas cuestiones de carácter metodológico que puedan constituir una aportación útil y tendente a una mejora en la objetivación de los efectos (Anguera, 1989b; Anguera y Blanco, 1988-91). Partimos para ello de una conceptualización según la cual un programa se halla formado por un conjunto heterogéneo de acciones incluso a veces con una estructura difusa, cuyo objetivo es resolver algún problema que incumbe a los sujetos (usuarios o beneficiarios), y una vez que se haya previamente fijado mediante un análisis de necesidades. En este mismo sentido, para Fernández Ballesteros (1987, p. 2) la evaluación de programas consiste en una "aplicación de procedimientos de investigación social para el enjuiciamiento o análisis sobre el mérito, valor o eficacia de acciones humanas concretas implantadas en la realidad social". La evaluación de programas implica una cadena de tomas de decisión que se inicia en el interés de los objetivos con base en las necesidades destacadas y evaluadas de acuerdo con alguno de los modelos existentes al efecto, y se concreta en una supuesta relevancia del impacto esperado de las

actividades programadas, de forma que se produzca un cambio efectivo en la dirección deseada con el mínimo costo posible.

Nosotros optamos claramente porque la evaluación de programas se constituya en un cuerpo sistemático de conocimientos que responda a la especificidad y complejidad de las múltiples situaciones —muy diversas temáticamente— que constituyen un campo de aplicación y desde la cual se deben impulsar los avances metodológicos necesarios que tiendan a una optimización constante en la evaluación de los efectos de intervención.

2. ¿POR QUÉ SE REQUIERE LA METODOLOGÍA OBSERVACIONAL EN EVALUACIÓN DE PROGRAMAS?

Un amplio consenso entre evaluadores, teóricos de la evaluación, instituciones que la reclaman y, en general, personas interesadas en el tema, lleva al convencimiento de que es necesario llevar a cabo una integración de distintas metodologías de forma que en su complementación se logren resolver cuestiones que, de lo contrario, tendrían difícil solución. El evaluador debe adoptar la metodología o combinación de ellas más adecuadas a la realidad específica en la que debe trabajar; la clave se halla en la pertinencia de las decisiones respecto al procedimiento en cada uno de los momentos del proceso —generalmente prolongado— que implica la evaluación de programas.

2.1. Cumplimiento de requisitos específicos

Durante la aún corta historia del moderno concepto de evaluación de programas se han esgrimido algunos argumentos que parecían encaminarla hacia las metodologías experimental (o bien cuasiexperimental) y selectiva³, dado que la mera existencia de un programa, desde su acepción convencional, se sitúa a contrapelo de la naturalidad de un determinado contexto por lo que implica de programación y, en consecuencia, de intervención. Pero sólo podemos afirmarlo parcialmente (interacción familiar, tipo de actividad profesional, relaciones sociales, etc.) y, en consecuencia, la metodología observacional puede desempeñar un importante papel en la evaluación de programas.

En sentido estricto, el criterio de intervención o manipulabilidad, que tradicionalmente ha servido de punto de referencia para la estructuración de las tres

³ La metodología selectiva se caracteriza esencialmente por un grado medio de intervención o control interno que se ejerce sobre ella, e implica una reconceptualización del antiguo término de "método correlacional", al cual consideramos carente de sentido por el hecho de que el análisis estadístico al que se someten algunos datos nunca puede ser criterio o punto de referencia para distinguir una metodología de otras. En la metodología selectiva tiene lugar la elicitación de una respuesta y el control que se ejerce es fundamentalmente de tipo indirecto, mediante la estandarización de las condiciones de la investigación, la formación de subclases o el control estadístico (Arnau, Anguera y Gómez, 1990).

grandes metodologías (observacional, selectiva y experimental) aquí no resulta más que parcialmente adecuado, por el simple motivo de que por tratarse de programas de intervención siempre existe manipulabilidad o control interno, y en consecuencia nunca se trata de acciones espontáneas. Por otra parte, y a pesar de que inicialmente la evaluación de programas implicaba la elección de un diseño experimental o cuasi-experimental (Pancer & Westhues, 1989), con la restricción que éstos comportan para la búsqueda de relaciones de causalidad debido a los controles que deben establecerse (Lipsey, Dunkle, Pollar & Stobart, 1985), ya no es admisible la aplicación de las pautas propias de la experimentación al ser conscientes del gran número de variables extrañas y artefactos experimentales que corren el riesgo de hallarse presentes (organísmicas, del sujeto, de funcionamiento institucional, de implementación de los programas, de funcionamiento comunitario, etc.), y las dificultades —muchas veces insalvables en su totalidad— que plantea su control mediante las técnicas habituales (mantenimiento de valores constantes, aleatorización, balanceo, contrabalanceo, etc.). Pese a que el seguimiento de este tipo de diseño ha sido una pretensión de gran parte de los evaluadores, en la práctica muy pocos trabajos en este campo han logrado con éxito el empleo de un diseño experimental de asignación de grupos al azar (Rutter y Giller, 1988).

Además, y al margen de un mayor desarrollo en el grupo siguiente, existen otros motivos que apoyan la *pertinencia e, incluso, la necesidad* —diríamos que básica o primaria— de la *metodología observacional en evaluación de programas*:

a) Por sus especiales características que la hacen flexible y adaptable a la especificidad de cada situación, pero que con la máxima objetividad y rigor permite la obtención de registros de elevada calidad informativa que pueden someterse a las técnicas de análisis de datos que mejor se adecúan a la obtención de aquellos resultados que no pretenden ser otra cosa que la respuesta a los objetivos planteados en cada caso, y en función del objetivo del programa, que nos deberá guiar en el estudio de la relación existente entre las variables implicadas en cada situación concreta estudiada.

b) Por el interés creciente que despierta el estudio de casos particulares. Es cierto que la evaluación de programas, considerada genéricamente, está estructurada para atender a colectivos en los que incide una determinada carencia, y que en principio podría considerarse nomotética por naturaleza. Pero no es menos cierto, y la experiencia diaria nos lo demuestra, que se requiere cada vez con mayor insistencia un tratamiento a partir de unidades (sujeto, díada que configura una pareja, unidad familiar, etc.) que nos sitúa en una clara perspectiva idiográfica, en la cual resultan forzadas determinadas técnicas de recogida de datos, como la encuesta, que son propias de la metodología selectiva.

c) Formalmente se cumplen los requisitos de observabilidad, dado que es posible captar el significado de una conducta manifestada que haya sido adecuadamente seleccionada en función de las conductas incluidas en o relacionadas con el objetivo (Anguera, 1986a, 1986b, 1989a). Se trata, pues de articular una percepción de la realidad con su adecuada interpretación, siempre que ésta asu-

ma la objetividad resultante de las distintas cribas y optimizaciones a las que se somete el dato registrado. La "captación" del significado de una situación en la que se involucran determinados comportamientos y/o la ocurrencia de ciertos hechos implica la síntesis adecuada entre percepción de los mismos y el hecho de que se les confiera un contenido, evitando toda posible distorsión en la representación de la realidad (Fassnacht, 1982) en sus diferentes niveles, a la vez que la evalúa a partir de unas coordenadas que responden a la ecuación personal del observador y a un marco de referencia teórico determinado; marco teórico, por otra parte, que, respecto a teorías de la evaluación, ha pasado de la diversidad de un próximo pasado a la integración actual (Shadish, Cook, & Leviton, 1991).

2.2. Naturaleza de la respuesta

Además de los requisitos específicos que acabamos de mencionar, y en aras de un mejor aprovechamiento de las posibilidades de la metodología observacional, se sugiere la adopción de otro criterio distinto al de la manipulabilidad o control interno, aquí mucho más adecuado, que es el de la *naturaleza de la respuesta educada al(a los) sujeto(s)* en función de la acción concreta a la que se le(s) somete, y que deberá combinarse, a efectos de la materialización del diseño correspondiente, con la *perspectiva diacrónica/sincrónica/mixta* adoptada.

A su vez, la respuesta del(de los) sujeto(s) puede o no implicar la cumplimentación de una prueba o sistema standard de recogida de información, y en cada negativo, puede ser o no de carácter perceptible. De forma esquemática, se expresa en la Figura 1.

Cada una de las posibilidades expresadas puede ser considerada desde una recogida puntual de datos (perspectiva sincrónica) o en su seguimiento continuado (perspectiva diacrónica), siendo factible una combinación de ambas (perspectiva mixta) en aquellas ocasiones en que interesa una proximidad (que puede llegar a ser un caso extremo una sucesión de puntos en el tiempo) de los momentos de recogida puntual de información.

La indudable cientificidad del procedimiento observacional, siempre que se cumplan los requisitos específicos pertinentes, ha facilitado en la actualidad el avance del propio "corpus" de conocimientos en que —como hemos antes indicado— debe cristalizar la evaluación de programas, la cual, por el número elevado de elementos implicados, la ausencia de límites estrictos o la dificultad de trazarlos, el peso relativo y variable de la conducta individual, y la necesidad de efectuar un seguimiento prolongado en el tiempo, acusó en las dos décadas anteriores la falta de una fundamentación metodológica —si no un deterioro progresivo en muchos casos— que permitiese impulsarlo adecuadamente, trascendiendo por supuesto del mero estudio descriptivo tradicional apenas sistematizado a un rango considerablemente superior en que no están ausentes ni la sofisticación técnica ni la complejidad de análisis, racionalizando en grado máximo toda la gama de recursos posibles (Ato, Quiñones, Romero y Rabadán, 1989).

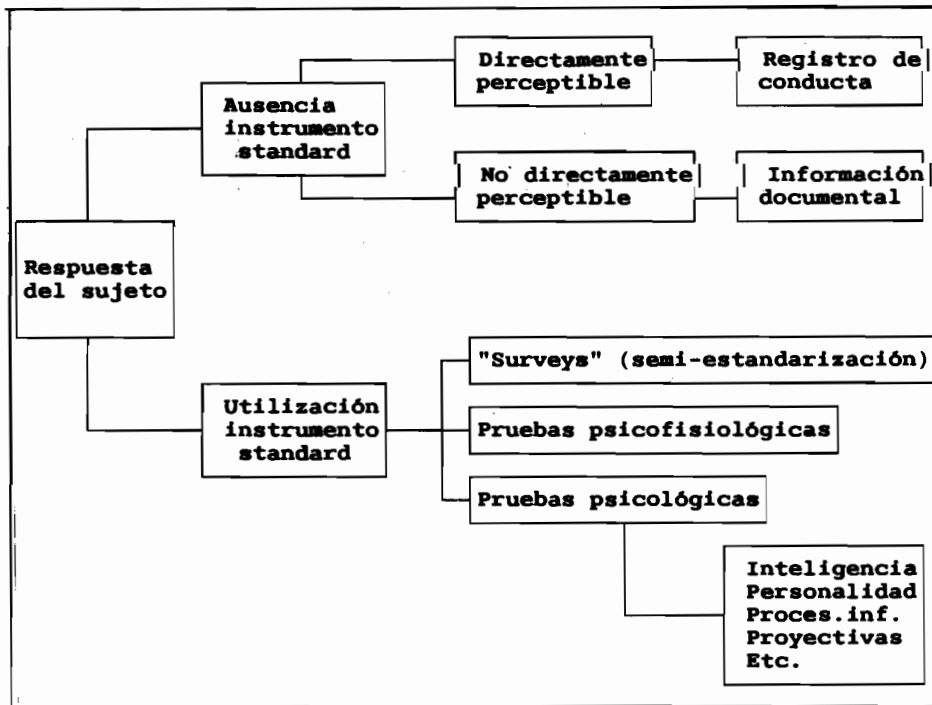


Figura 1. Criterio alternativo al de control interno en la consideración de las metodologías a seguir en evaluación de programas (Anguera y Redondo, en prensa).

3. RAZONES DE PERTINENCIA EN FUNCIÓN DE CRITERIOS TAXONÓMICOS

En distintas ocasiones nos hemos pronunciado sobre diversos criterios taxonómicos —y no sólo clasificatorios, dado que entre ellos también existe un rango jerárquico— que resultan útiles en metodología observacional (Anguera, 1985, 1988, 1990, 1991), y los hemos revisado paulatinamente dado el actual proceso de asentamiento de algunos de ellos y la polémica que generan. Con respecto, por supuesto, a planteamientos alternativos, desde nuestro punto de vista, se pueden establecer, entre otros (Anguera, 1979), cinco criterios taxonómicos que, una vez desarrollados, permiten encajar las diversas modalidades de observación resultantes en momentos y situaciones diversas que plantea la evaluación de programas.

Estos posibles criterios taxonómicos se hallan representados en la Tabla 1, y son descritos a continuación:

Tabla 1. Principales criterios taxonómicos útiles en metodología observacional

Rango	Estrategia	Observabilidad	Participación	Niveles de respuesta
Método	Via inductiva	Directa	No participante	No verbal
			Participante	Proxémica
Técnica	Via deductiva	Indirecta	Participac.-Observación	Vocal
			Autoobservación	Verbal

3.1. Rango procedimental

La metodología observacional, como su propio nombre indica, implica el seguimiento del método científico adaptado a las características propias que la definen (ausencia de control interno, naturalidad, del contexto, observabilidad, preferencia por lo idiográfico, perspectiva temporal mediante el seguimiento), y, como hemos mencionado en el punto anterior, es un hecho que en el ámbito de evaluación de programas es más frecuente el uso de la observación como técnica que como método, aunque éste tenga su razón de ser en tanto en cuanto *pueden implementarse programas diversos sin que se menoscabe o altere la integridad del contexto natural en que nos desenvolvemos*. Así, en el ámbito penitenciario, un programa lúcido deportivo (Redondo, Pérez, Agudo, Roca i Azpiazu, 1990) daría lugar a una evaluación perfectamente viable —y enfáticamente recomendable— mediante el método de observación.

Hay que tener en cuenta, sin embargo, que como *técnica* no significa más que una estrategia de recogida de determinado tipo de datos subordinados a las directrices de una metodología selectiva o experimental. Dado que aquí la observación tiene por misión únicamente el suministro de información de forma complementaria a las otras formas de recogida de datos, incluye el estudio de conductas generadas en situación artificial (Anguera, 1981; Anguera y Redondo, en prensa), lo cual permite una reducción considerable de los requisitos de partida (Riba, 1991) para su uso, sacrificándose, en contrapartida, las presentaciones de esta metodología (Bakeman & Gottman, 1987), que Bakeman y Gottman (1989, p. 23) definen magistralmente como “vía específica de acceso a la cuantificación de la conducta. Esta aproximación se refiere típicamente a la ocurrencia espontánea de la conducta que se observa en contextos naturales”. Son muchos los programas que utilizan la observación como técnica; por ejemplo, en progra-

mas de mejora de bienestar subjetivo en ancianos institucionalizados (Fernández-Ballesteros, Izal, Díaz, González y Souto, 1988).

3.2. Estrategia

Aunque esta cuestión, como la relativa al rango procedimental, se halla inserta en cuestiones de metodología básica, aquí revierte de forma importante.

Mediante, la *vía inductiva*, de claro carácter ascendente, partiríamos poco menos que de una "tábula rasa", y, sin el marcaje de un "corpus" teórico, se irían sucediendo los registros correspondientes a observaciones de casos particulares, de forma que el objetivo se hallaría en conocer el transcurso de las conductas estudiadas si el estudio se llevara a cabo durante un período de tiempo relativamente prolongado, y se manifestara algún tipo de regularidad en los registros de las diversas unidades homogéneas consideradas. Buena parte de los estudios de evaluación ambiental reúnen estas características, dada la individualización que supone cada contexto específico (Fernández Ballesteros, Díaz, Izal y González, 1987).

Por el contrario, en la *vía deductiva* se toma como referencia inmediata una construcción teórica con mayor o menor zona de cobertura y diferente grado de consolidación (Chen & Rossi, 1983), pero que en cualquier caso permite una contrastación de la hipótesis relativa a la efectividad del tratamiento o la relación existente entre diferentes elementos implicados (que por otra parte, deben ser observables o haberse operativizado mediante indicadores). Una ilustración interesante es la de aquellos programas de prevención de enfermedades (cardio-vasculares, cáncer, obesidad, diabetes, etc.) a partir del conocimiento de los efectos que producen las distintas sustancias contenidas en la dieta y su cantidad (Smith & Bobrow, 1991).

En evaluación de programa, la *disyuntiva entre vía inductiva y deductiva* (consideradas de forma "pura") nos retrotrae a la dicotomía relativa al modelo previo de evaluación de que se parte (Veney & Kaluzny, 1984), y esencialmente en la diferenciación entre el *modelo lineal* y el *no lineal*.

En el primero se plantea una secuencia lógica de fases que implica la planificación (construcción) del programa, su implementación, y posteriormente su evaluación, en una estructura cuasi-cerrada en la que si bien establece un proceso de "feed-back" para una re-planificación del programa, la modificación que se incluya formará parte de la siguiente campaña o ronda de aplicación.

En el modelo no lineal o interactivo, sin embargo, aunque probablemente se adapta menos a los calendarios de la Administración, mejoran sensiblemente sus posibilidades y aumentan las ventajas que se obtienen, pues es mucho más flexible al considerar la evaluación como parte integral de todo el ciclo. La evaluación de programas se hace presente desde la planificación del programa en función de los problemas específicos que se trata de resolver, contrastando con

otros posibles programas alternativos en cuanto a disponibilidad de recursos, estructuración, periodificación, complejidad de las acciones a emprender, etc.; igualmente se plantea la evaluación en la fase de implementación respecto a la posible necesidad de realización de encuestas, monitorización del personal necesario, selección de la población afectada, etc. Y una vez que el programa se ha implementado, nos podemos formular varias preguntas al efecto: ¿Es efectivo y eficiente? ¿son razonables los costos? ¿se habría solucionado el problema si no se hubiera aplicado el programa? ¿podría otro programa haber resuelto el problema de forma más óptima? El estudio de los efectos y del impacto del programa constituyen los últimos eslabones de la evaluación a que se ve sometido todo el proceso.

Además, existe una correspondencia entre estos modelos —lineal y no lineal— y los de *evaluación sumativa y formativa*, respectivamente (Anguera y Redondo, en prensa). La evaluación sumativa se refiere a las actividades asociadas con los efectos del programa y con unas connotaciones de control, mientras que la evaluación formativa, por la que abogamos, se refiere a las actividades asociadas con las operaciones continuas del programa a lo largo de todo el proceso. Por ejemplo, el estudio de predictores de la delincuencia a partir del estudio y análisis previo de carrera delictivas (historial), o la evaluación continuada de las personas en prisión en las fases inicial, durante el internamiento y a la salida para el posterior estudio evaluativo de la reintegración social o de la reincidencia. En este sentido, Fitz-Gibbon & Morris (1987) consideran que la evaluación es formativa si la tarea del evaluador es la de ayudar a planificar y desarrollar el programa.

Una vez definido el modelo, un programa incorpora un rango de actividades que puede ser cuán amplio se desee; esto nos muestra la útil complementariedad entre las vías inductiva y deductiva, dado que en función del marco teórico (explicado mediante el modelo) se programan las distintas acciones que configuran el programa, y, además, especialmente en los inicios de la implementación, el registro de las conductas ejecutadas en cada uno de los sujetos o unidades consideradas, funciona a modo de vía inductiva, con la plasticidad que permite introducir mejoras en las técnicas utilizadas, intensificar la formación del personal, ajustar mejor el previsible monto de costes, etc.

Así, en el ámbito sanitario, “el objeto de la evaluación es mejorar los programas y la infraestructura sanitaria para su ejecución y orientar la distribución de los recursos entre las actividades ya emprendidas o que pueden desarrollarse en el futuro. La evaluación es un medio sistemático de aprender empíricamente y de utilizar las lecciones aprendidas para el mejoramiento de las actividades en curso y para el fomento de una planificación más satisfactoria mediante una selección rigurosa entre las distintas posibilidades de acción futura” (OMS, 1981, p.5).

En programas penitenciarios la evaluación debe dirigirse, desde una perspectiva ambiental que implica el establecimiento de nuevos climas sociales (Moos, 1982), a un replanteamiento de todo el proceso de intervención (Clemen-

te y Gil, 1985), orientado éste hacia nuevas direcciones, como la del establecimiento de fases progresivas (Redondo, Roca, Pérez, Sánchez y Deumal, en prensa), priorizar la prevención, fomentar la adecuación selectiva del régimen abierto (Redondo, en prensa), etc.

En el ámbito de Servicios Sociales, a pesar de un indudable interés por el tema en los países en los que desde el propio Estado existe una directa asignación de recursos para resolver carencias propias de colectivos marginales, o que viven en un permanente u ocasional estado de necesidad (Kahn & Kamerman, 1980), y en donde España no es una excepción (Culyer, 1982; Del Valle, 1987; Chacón, 1989; Rueda 1989; Luque, 1989), se está trabajando en la concreción de estrategias que faciliten la operativización del denominado modelo asistencial o comunitario, predominando en la actualidad las opiniones favorables al uso de metodologías flexibles (López Blasco, 1984; Muñoz, 1984), referidas tanto a la metodología observacional desde las estrategias inductiva y deductiva, como a otras metodologías complementarias, especialmente la selectiva, ("surveys", determinadas modalidades de entrevistas, etc.).

3.3. Observabilidad

Se trata de un tema complejo que en la actualidad se halla sometido a polémica que raya en lo epistemológico; no olvidemos al efecto que la introspección, desde la época de Wundt, ha basculado en su valoración, e igualmente podríamos decir, pasando al terreno de técnicas de recogida de datos, que los diarios constituyen un material rico en información para el estudio de casos particulares, o que tanto la desvirtuada calidad de dicha información como la distorsión que implicaría su interpretación por el elevado grado de inferencia la desproveerían totalmente de valor. Y ambos ejemplos sólo constituyen una pequeña muestra de posiciones contrapuestas. Dado que aquí nos interesa tan solo desde su incidencia y repercusión en la evaluación de programas, distinguiremos entre observación directa e indirecta (Anguera, 1990).

La *observación directa* —observación en el estricto sentido del término— es la que permite el registro "in vivo" de conductas directamente perceptibles mediante el canal sensorial pertinente (visual y/o auditivo generalmente), con muy escaso nivel de inferencia, y en donde predomina el componente perceptivo sobre el interpretativo, pero sin que ello signifique que el primero anule al segundo, ya que de ser así se correría el riesgo de pretender que un medio técnico de registro (magnetofón, video, etc.) pudiera sustituir al observador humano, con lo cual se trataría de una mera transcripción vacía de contenido.

Su aplicabilidad en evaluación de programas es altísima, y cabe pensar, como ejemplos, en programas de intervención educativa (Forns i Gómez, 1990), de rehabilitación de drogadictos (Pleite y Pardo, s/f), de mejora de la interacción en niños en régimen institucional (Fernández del Valle, 1991a), etc.

La *Observación indirecta*, muy cuestionada desde diversos ángulos, implica la existencia de conductas encubiertas que requerirán una inferencia y, por tanto, una carga interpretativa que puede redundar en detrimento y menoscabo de la objetividad requerida en toda metodología científica. De ahí los recelos que inspira y su carácter un tanto ambiguo a la hora de considerar su inclusión. La definición atribuida al término en el glosario elaborado por Anguera, Behar, Blanco, Carreras, Losada, Quera y Riba (en prensa) es "Tipo de observación correspondiente al estudio de conductas encubiertas, y, por tanto, no perceptibles en el momento del estudio. Por lo tanto, siempre requiere un considerable grado de inferencia".

No obstante, en los últimos años se ha avanzado en este sentido, y el estudio de los procesos cognitivos superiores ha impulsado a empujar los límites que circundan a la observación indirecta. Consideramos que forma parte de ella (Anguera, 1990; Anguera y Redondo, en prensa):

a) Indicadores externos de un estado emocional o de una situación que requiera inferencia. En un ejemplo correspondiente a un programa de alfabetización en un centro penitenciario (Redondo, Pérez, Agudo, Roca i Azpiazu, 1990), cuando a un interno que se halla en la escuela de la prisión se le pregunta y baja la cabeza, evita intercambios de mirada, se produce una sudoración no justificada por la temperatura ambiente, tartamudea o presenta "lapsus linguae", etc., da lugar a una configuración global cuya principal problemática sería la de su validación para inferir la correspondiente motivación.

b) Textos documentales obtenidos por la grabación de la conducta verbal de un sujeto, y que pueden ser sometidos a un análisis de contenidos, proceso que corre en paralelo con el de observación directa (Markoff, Shapiro & Weitman, 1975), pero con la diferencia fundamental —que constituye un indudable riesgo— de delimitación de las unidades lingüísticas y su codificación (Krippendorf, 1980). Por ejemplo, en un programa de mejora de calidad de atención sanitaria a pacientes que se hallan en fase terminal, y se analizan sus grabaciones (Farré, 1990), o en programas de recuperación de minusválidos físicos (Alvira, 1989), o en programas de atención a vagabundos sin hogar (Generalitat de Catalunya i Ajuntament de Barcelona, 1989) cuando se les pregunta por su situación, y se les insta a que accedan a pernoctar en residencias al efecto.

c) Autoinformes procedentes del registro propio de la autoobservación, que pueden o no corresponder a conductas heteroobservables. En el primer caso no existen problemas relativos a la garantía de validez; por ejemplo, en programas de deshabitación al tabaco (Richart y Cabrero, 1988), o en Psicología Deportiva (Cerro, 1980), o en el tratamiento de la obesidad (Mahoney, Moura y Wade, 1973) o de la enuresis (Avia, 1981).

Más cuestionables son los autoinformes resultantes de autoobservación de conductas no hetero-observables, ya que es cuando entra en juego la explicitación de vivencias por parte del sujeto, y por supuesto si se trata de autorregistros retrospectivos (Levinson, Gordon, & Skeff, 1990). Con el fin de facilitar al máxi-

mo de objetividad (Vizcarro y García Marcos, 1981), pero con importantes dudas acerca de su validez (Howard, Schmeck & Bray, 1979), existen numerosos ejemplos de inventarios al efecto, como el de conductas obsesivo-compulsivas (Hodgson & Racman, 1977), que consta de 30 ítems referidos a rituales obsesivos, que aunque son externamente observables, se presentan de forma "personalizada", como "evito usar teléfonos públicos porque temo contagiarme", "normalmente no tardo mucho tiempo en vestirme por la mañana", "generalmente pienso que no he hecho algo bien, aunque lo haya hecho con mucho cuidado", etc.

d) Datos idiográficos de naturaleza diversa obtenidos oralmente a partir del interesado o terceras personas en la resolución de pruebas / confección de fichas / protocolos (Ericsson & Simon, 1984), o análisis de tareas (Rodrigo, 1982). Por ejemplo, en programas de entrenamiento en habilidades de memoria (Fernández-Ballesteros, Montorio, Izal, Guerrero, Llorente & Hernández, 1990), o en programas de detección, registro e intervención de malos tratos en la infancia (Ferrero, 1990).

e) Datos de carácter nomotético que proceden del análisis de masas de población, generalmente de carácter estadístico, y a los que no puede negarse su carácter de observación indirecta. Por ejemplo en programas paliativos del ausentismo escolar (Fernández del Valle, 1991b).

f) Aunque en menor medida, y generalmente de carácter anecdótico, no podemos olvidar la información suministrada por la existencia de elementos materiales (denominados "agregativos" o de "productos permanentes" que denotan la previa existencia de ciertas conductas, las cuales son inferidas con riesgo variable (existencia de huellas, pisadas, colillas, etc. En un determinado lugar), o bien la información que se obtiene por el desgaste de uso de objetos, suelos, etc., y que en la actualidad se puede operativizar con elevada precisión. Por ejemplo, en un programa de adquisición de hábitos de aseo personal, el registro de elementos característicos de desaliño.

3.4. Participación

El nivel de participación se puede presentar en un "continuum" en razón de la mayor distancia a la identificación entre observador y observado, dando lugar a las cuatro modalidades siguientes, todas las cuales son de aplicación en evaluación de programas:

a) La observación no participante tiene lugar cuando no existe interacción entre observador y observado, lo cual confiere una ventaja respecto a independencia y neutralidad de la información, aunque en ocasiones genera inconvenientes de inviabilidad o dificultad de acceso a la información. Por ejemplo, en programas institucionales de prevención de accidentes y seguridad vial, de limpieza en calles y lugares públicos de una ciudad, campaña de higiene bucal en las escuelas, etc.

b) La observación participante supone una interacción entre observador y observado, pudiendo variar notablemente en función de las circunstancias y la situación. Por ejemplo en programas de prevención de adicción a la droga en centros de acogida, y puestos en práctica por parte de profesionales distintos de los educadores u otros que estén en permanente contacto con los muchachos.

c) La participación-observación resulta de una intensificación de la observación participante, cuando un miembro de un grupo adquiere la cualidad de observador de otro(s) perteneciente(s) al mismo grupo. Por ejemplo, en los programas de educación materno-infantil (Richart y Cabrero, 1988), cuando madre e hijo son destinatarios del programa, pero la madre lleva a cabo total o parcialmente registros de la conducta de su hijo.

d) La auto-observación implica el grado más elevado de participación en la observación, donde el observador es a la vez sujeto y objeto. Al referirnos a la observación indirecta, que permite contemplar la auto-observación desde otra perspectiva, ya hemos comentado la polémica existente, y que implica el doble proceso que consiste, por una parte, en atender deliberadamente a la propia conducta, y, por otra, en registrarla mediante alguna técnica especialmente indicada. Por ejemplo, en conductas que pertenecen a la esfera de la privacidad, como programas de resolución de conflictos a nivel familiar en donde el estudio de la interacción entre sus miembros se convierte en el eje principal (Gottman, 1980), o los programas citados de enuresis (Avia, 1981).

3.5. Niveles de respuesta

Las respuestas/acciones elicítadas al(a los) sujeto(s) o producidas espontáneamente por ellos se hallan formando parte de un plan de acción a través de una serie de actividades de carácter heterogéneo y orden diverso, que consecuentemente involucran los distintos niveles de respuesta que puedan establecerse, con la característica común de su observabilidad (Norris, 1981, 1984). La distinción clásica entre los niveles simples, propuesta por Weick (1968), que establece las conductas no verbal, espacial o proxémica, vocal o extralingüística y verbal o lingüística, consideramos que sigue teniendo vigencia por su cobertura respecto a la observabilidad y porque permite el análisis de la concurrencia entre varios de estos niveles simples, si bien es cierto que sigue teniendo pendiente el problema relativo al establecimiento de unidades de conducta (Anguera, 1988, 1990, 1991).

Dado que en este artículo estamos estudiando el enclave y posibilidades de la metodología observacional en evaluación de programas, se puede justificar con relativa sencillez que el hecho de que se pueda considerar con exhaustividad la naturaleza o modalidad de la respuesta del(de los) sujeto(s) hace posible que, tanto en la elaboración del plan de un programa, como en la recogida de datos simultánea a la implementación del mismo o incluso posterior codificación a partir de su grabación y correspondiente almacenamiento, no existían prácticamente

problemas técnicos pendientes de resolución, y tan sólo relativos a la toma de decisión sobre el alcance mínimo —con significado o atribución de contenido— de las unidades a considerar en el registro (Lehner, 1979; Zeiler, 1986).

4. ADECUACIÓN DE TÉCNICAS DE REGISTRO Y OPTIMIZACIÓN DE DATOS EN LA IMPLEMENTACIÓN DEL PROGRAMA

Una vez elaborado un programa, debe llevarse a la práctica de acuerdo con el plan previsto y la temporalización (“timing”) efectuada, constituyendo la operación que se denomina *implementación* de dicho programa. En la actualidad se ha dañado ya la expresión de “evaluación de la implementación” (“implementation assessment”) (Pancer & Westhues, 1989), que se refiere a qué acciones deben llevarse a cabo, cuándo y por quién, con el fin de que la alternativa seleccionada sea la que maximice las posibilidades de responder a los objetivos del programa. Su éxito dependerá no sólo del cierto o errores del programa en sí mismo, sino de la medida en que estuviera adecuadamente especificado —en la actualidad observamos cómo muchos programas adolecen de suficientes detalles en este sentido— y de la formación de las personas que colaboran en la puesta en marcha de las acciones previstas y la correspondencia recogida de información y optimización de esta información.

Tal y como señalan Rutter y Giller (1988), y de modo especial en el terreno criminológico y penitenciario, la mayoría de los informes empíricos adolecen de información acerca de cómo se llevó a cabo empíricamente la intervención, lo que impide o dificulta la valoración objetiva del proceso de implementación. Aquí nos referimos a literatura específica en el ámbito de la observación para hacer hincapié en la trascendencia de la construcción del sistema de categorías, su utilidad en el registro, y la revisión de la calidad de los datos en aras de su mejora.

En virtud del *proceso de categorización* se elabora el “instrumento de medida” de la conducta, de carácter por supuesto no estándar (Figura 1), sino, que por el contrario, se sustenta en la “ad-hocidad” de la situación y respuestas a las que se dirige —o que median en— un determinado programa. De forma muy esquemática (Anguera 1991), es precisamente a partir de una lista de unidades de conducta (denominada “lista de rasgos”) obtenida con garantías de exhaustividad (número suficiente de sesiones previas de carácter exploratorio que permite elaborar un repertorio), que se inicia el proceso, guiado en mayor o menor medida por un “corpus” teórico (Bickman, 1987) relativo al ámbito en que se inscribe el programa, y que, en virtud de su distinto grado de solidez y formalización, por una parte, y de la decisión del evaluador, y por otra, se procede a la agrupación de unidades de conducta por semejanza o finalidad en función de unos criterios previamente establecidos, con lo cual se proponen unas primeras categorías provisionales, que se someten a prueba en sesiones de observación destinadas al efecto, se modifican adecuadamente (mediante desglose, agrupamiento, incorpo-

ración de nuevas propuestas, eliminación, incorporación de matrices distintos en la definición, reordenación, etc.), y así sucesivamente de forma interactiva hasta llegar al sistema de categorías considerado definitivo (sin olvidar la necesidad de revisión periódica para adaptarlo a la dinámica de surgimiento de nuevas posibles categorías, adaptación de las existentes a nuevos matices, o eliminación ante la extinción de las conductas que se le asignaban), que deberá cumplir las dos condiciones básicas de exhaustividad y mutua exclusividad en el sector del comportamiento en que se asienta el programa.

Una vez elaborado el sistema de categorías, la materialización del registro es sumamente simple (Blanco y Anguera, 1987, 1991), pudiendo adoptarse cualquiera de los sistemas existentes o crear otros nuevos para recoger y almacenar la información. Igualmente, la naturaleza de la codificación se establecerá en función de su pertinencia en las situaciones en que vaya a utilizarse. Así, por ejemplo, el sistema de categorías elaborado por Santoyo y Espinosa (1987) es perfectamente adecuado para el estudio de las propiedades bidireccionales propias de la interacción social en niños, y está adaptado a estas circunstancias; si interesa el estudio de la interacción social en una residencia geriátrica, el interés central es igualmente al de conductas de intercambio, la estructura global de preferencias sociales es igualmente pertinente, la representación de la dirección de las elecciones entre sujetos y las condiciones de interacción recíprocas será igualmente necesaria, pero, como imperativo de la "ad-hocidad" mencionada, deberán adaptarse algunas de las categorías (actividad académica, zona de juego, estructuras, etc.) a la actividad desarrollada por los ancianos en la residencia.

Y finalmente, se impone la optimización del registro a partir de una serie de precauciones que se materializan en la toma de decisiones acerca de los criterios de muestreo (Anguera, 1983; Quera, 1991), y la garantía acerca de la fiabilidad, precisión y validez (Blanco, 1983, 1989; Blanco y Anguera, 1984;), la cual alcanza su máximo exponente con la aplicación de la teoría de la generalizabilidad a datos observacionales (Blanco, 1986a, 1986b, 1989; Blanco, Losada y Anguera, 1991, en prensa), la cual, a su vez, hace posible el análisis económico de los programas sometidos a evaluación, y en consecuencia, la prospectiva en la evaluación de su impacto, en virtud del balance establecido entre los costes que implica el programa (instrumentos, sesiones, evaluadores, modalidades de registro necesarias, etc.), y las respectivas posibilidades en cada caso respecto al alcance de los resultados.

Complementariamente con la implementación de un programa destaca su *monitorización*, que a su vez, constituye un componente crítico de la evaluación (Vency & Kaluzny, 1984). La monitorización de un programa es la comparación entre el plan trazado y la realidad: Si no coinciden, o bien el plan puede ser modificado o las acciones llevadas a cabo desfiguran la realidad, y, en cualquier caso, son críticos los datos generados. Los aspectos vigilados mediante la monitorización pueden ser muy variados; Si los costos coinciden con el importe presupuestado, si el personal disponible responde a los requisitos fijados, si la

periodificación de las actividades se ajusta al calendario trazado, y si se obtienen los resultados proyectados. No obstante el hecho de que algunos autores, como Freeman & Rossi (1981), establezcan ciertas restricciones en este sentido, la monitorización "... es la menos conocida pero probablemente la más practicada de las formas de evaluación" (ERS Standards Committee, 1982, p. 10).

Una de sus modalidades es la realización de auditorías (Posavac & Carey, 1985), que consisten en un examen sistemático de los "inputs" y "outputs" del programa como prelude de un riguroso estudio de los resultados del programa. En la práctica muchas auditorías llegan a transformarse en rutinarias, pero el peligro que allí acecha deriva del hecho de que los programas responden a diferentes necesidades en momentos distintos. En cualquier caso, no sólo es deseable sino necesaria la realización de breves informes periódicos que den cuenta de la marcha en la implementación del programa y de los datos que se recogen, y que siempre podrán evitar el riesgo de una interpretación subjetiva de éstos.

5. ANÁLISIS DE DATOS EN FUNCIÓN DEL DISEÑO DEL PROGRAMA

Precisamente un análisis del balance contenido-procedimiento en evaluación de programas nos orientan a la evaluación entendida como proceso o como fuente de resultados. Mientras que en la segunda la actividad del evaluador se reduce estrictamente a estimar la magnitud de los efectos de un tratamiento sobre el cual se había establecido previamente una hipótesis a modo de conjetura, y por tanto posee una finalidad claramente demostrativa (Kidder & Judd, 1986); en cambio, en la evaluación entendida como proceso, se pretende entender por qué se ha producido un determinado resultado como efecto de un tratamiento. La finalidad es ahora explicativa, y de aquí la necesidad de que el diseño del programa, en donde deberá planificarse la búsqueda de las relaciones que interesen, tenga una materialización adecuada en el tipo de análisis de datos que se lleve a cabo.

Respecto a esta cuestión, desde el fin de la década de los setenta, a partir de los trabajos de Bakeman & Dabbs (1976); Bakeman & Brown (1977); Bakeman (1978); Gottman (1978); Gottman & Notarius (1978); Sackett (1978, 1979, 1980); Castellan, 1979; Gottman & Bakeman (1979); Lehner; (1979) y Thomas & Malone (1979); hasta ahora, el camino recorrido es abismal si lo comparamos con trabajos realizados previamente.

Dado que en aras a la flexibilidad que impone la evaluación de programas, entendemos el diseño en su acepción no normativa (Anguera y Redondo, en prensa); en distintas ocasiones (Anguera, 1985, 1988, 1990) hemos propuesto, tan sólo a modo de pautas referenciales, dos ejes, referidos a la dicotomización idiográfico/nomotético, por una parte, y registro puntual/seguimiento, por otra, por las grandes posibilidades que ofrecen sus distintas combinaciones, y alenta-

mos al lector a que, en cada caso, en cada una de las evaluaciones de programas que se proponga llevar a cabo, haga el pequeño esfuerzo de adecuar este bagaje de recursos analíticos que brindan la metodología observacional a las peculiaridades propias y específicas del programa que evalúa.

No olvidemos que la humildad y la grandeza de la metodología observacional radican precisamente en la pobreza del dato único (y de ahí la necesidad de articular correctamente el proceso de recogida, optimización y análisis) y en su inmensa adaptabilidad a cualquier situación en que nos interese estudiar el comportamiento humano. En consecuencia, está ante nosotros el reto, ante la expansión presente —y previsible en los próximos años— de la evaluación de programas en los más diversos ámbitos, de saber poner en práctica todos los innumerables recursos que nos ofrece la observación sistemática.

6. CONCLUSIONES

Dado el estado actual de desarrollo de las dos disciplinas que aquí confluyen —metodología observacional y evaluación de programas—, como en tantas otras ocasiones no se produce por suerte una mera adición, sino una intersección que tiene importantes efectos multiplicadores no sólo a efectos de investigación, sino, lo que es mucho más importante, para el potencial usuario de programas de intervención. Estos programas, a diferencia de la aplicación de las técnicas de modificación de conducta —cuyo interés radica esencialmente en trastornos o disfunciones muy localizadas en un determinado sujeto—, tienen un alcance mucho mayor, se insertan en el dominio de la Psicología Comunitaria, y requieren de una infraestructura mínima que en la mayoría de las ocasiones precisa de recursos de considerable envergadura. También, en justa contrapartida, los beneficios para los potenciales usuarios llegan a tener importante envergadura.

Finalmente, y como mensaje explícito, debe quedar clara la necesidad y la voluntad de que, una vez más, los avances metodológicos se coloquen al servicio de nobles objetivos en los que el estudio científico del comportamiento humano es el eje indiscutible para su consecución.

REFERENCIAS

- Alvira, F. (1989). Los centros ocupacionales y los centros de recuperación de minusválidos físicos del Instituto Nacional de Servicios Sociales: Un ejemplo de evaluación externa. *Anales de Psicología* (Murcia), 5, 67-76.
- Anguera, M.T. (1979). Observational Typology. *Quality & Quantity. European-American Journal of Methodology*, 13, 449-484.
- Anguera, M.T. (1981). La observación (I): Problemas metodológicos. La observación (II): Situaciones naturales y de laboratorio. En R. Fernández Ballesteros y
-

- J.A.I. Carrobbles (Eds.) *Evaluación conductual: Metodología y aplicaciones* (pp. 292-333, 334-363). Madrid: Pirámide.
- Anguera, M.T. (1983). *Manual de prácticas de observación*. México: Trillas.
- Anguera, M.T. (1985). *Metodología de la observación en las Ciencias Humanas* (3ª ed. ampliada). Madrid: Cátedra.
- Anguera, M.T. (1986a). Observación. En S. Molina (Dir.) *Diccionario Temático de Educación Especial* (pp. 446-483). Madrid: C.E.P.E., vol. 1.
- Anguera, M.T. (1986b). Niveles descriptivos en metodología observacional. *Apuntes de Psicología* (Sevilla), 16, 29-32.
- Anguera, M.T. (1988). *Observación en la escuela*. Barcelona: Graó .
- Anguera, M.T. (1989a). La observación de la conducta en el ámbito hospitalario: Principios, clases, ventajas y limitaciones. En A. Polaino-Lorente (Coord.) *Introducción a la modificación de conducta para profesionales de enfermería* (pp. 39-71). Barcelona: P.P.U.
- Anguera, M.T. (1989b). Innovaciones en la metodología de evaluación de programas. *Anales de Psicología* (Murcia), 5 , 13-42.
- Anguera, M.T. (1990). Metodología observacional. En J. Arnau, M.T. Anguera y J. Gómez. *Metodología de la investigación en Ciencias del Comportamiento* (pp. 125-236). Murcia: Secretariado de Publicaciones de la Universidad.
- Anguera, M.T. (Ed.) (1991). *Metodología observacional en la investigación psicológica*. Barcelona: P.P.U., Vol. I.
- Anguera, M.T. y Blanco, A. (1988-91). Evaluación de programas en Servicios Sociales: Alternativas metodológicas. Investigación subvencionada por la DGICYT (Proyecto PB87-0047).
- Anguera, M.T., Behar, J. Blanco, A., Carreres, M.V., Losada, J.L., Quera, V. y Riba, C. (en prensa). Glosario. En M.T. Anguera (Ed.) *Metodología observacional en la investigación psicológica*. Barcelona: P.P.U., Vol. II.
- Anguera, M.T. y Redondo, S. (en prensa). La evaluación de la intervención penitenciaria. *Delincuencia/Delinquency*.
- Arnau, J., Anguera, M.T. y Gómez, J. (1990). *Metodología de la investigación en Ciencias del Comportamiento*. Murcia: Secretariado de Publicaciones de la Universidad.
- Ato, M., Quiñones, E, Romero, A. y Rabadán, R. (1989).Evaluación de programas: Aspectos básicos. *Anales de Psicología*, 5, 1-12.
- Avia, M.D. (1981). La auto-observación. En R. Fernández Ballesteros y J.A.I. Carrobbles (Eds.) *Evaluación conductual: Metodología y aplicaciones* (pp.267-291). Madrid: Pirámide.
- Bakeman, R. (1978). Untangling streams of behavior: Sequential analyses of observation data. In G.P. Sackett (Ed.) *Observing Behavior. Vol. II: Data collection and analyses methods* (pp. 63-78). Baltimore: University of Park Press.
- Bakeman, R. & Brown, J.V. (1977). Behavioral dialogues: An approach to the assessment of mother-infant interaction. *Child Development*, 48, 195-203.
-

- Bakeman, R. & Dabbs, J.M. (1976). Social interaction observed: Some approaches to the analysis of behavior streams. *Journal of Social and Personality Behavior*, 2, 335-345.
- Bakeman, R. & Gottman, J.M. (1987). Applying observational methods: A systematic view. In J.D. Osofsky (Ed.) *Handbook of infant development* (pp. 818-854). New York: Wiley / Sons.
- Bakeman, R. & Gottman, J. M. (1989). *Observación de la conducta interactiva: Introducción al análisis secuencial*. Madrid: Morata (edic. orig., 1986).
- Bickman, L. (1987). The functions of program theory. In L. Bickman (Ed.) *Using program theory in evaluation* (pp. 5-18). San Francisco: Jossey-Bass.
- Blanco, A. (1983). Análisis cuantitativo de la conducta en sus contextos naturales: Desarrollo de un modelo de series de datos para el establecimiento de tendencias, patrones y secuencias. Tesis Doctoral no publicada. Barcelona: Universidad de Barcelona.
- Blanco, A. (1986a, Mayo). Generalizabilidad en diseños de observación de la conducta. Comunicación presentada en la I Jornada Psicológica de la Sociedad Española de Psicología (Sección Cataluña). Barcelona.
- Blanco, A. (1986b, Junio). Generalizabilidad de la observación de la conducta. Trabajo inédito no publicado. Barcelona: Universidad de Barcelona.
- Blanco, A. (1986c, July). Problems of generalizability in Environmental Psychology. Paper presented at the 21st International Congress of Applied Psychology. Jerusalem, Israel.
- Blanco, A. (1987, Septiembre) Codificación y registro en evaluación. Comunicación presentada en el II Congreso de Evaluación Psicológica. Madrid.
- Blanco, A. (1989). Fiabilidad y generalización de la observación conductual. *Anuario de Psicología* (Barcelona), 45, 5-32.
- Blanco, A. y Anguera, M.T. (1984, Septiembre). Fiabilidad, precisión y validez de los registros observacionales. Comunicación presentada al XXIII Congreso Internacional de Psicología. México, Acapulco.
- Blanco, A. y Anguera, M.T. (1991). Sistemas de codificación. En M.T. Anguera (Ed.) *Metodología observacional en la investigación psicológica* (pp. 193-239). Barcelona: P.P.U., Vol. I.
- Blanco, A., Losada, J.L. y Anguera, M.T. (1991, Mayo). Descripción de errores de medida en estudios observacionales de evaluación conductual. Comunicación presentada en el II Congreso Internacional "Latini Dies". Sitges, Barcelona.
- Blanco, A., Losada, J.L. y Anguera M.T. (en prensa). Estimación de la precisión en diseños de evaluación ambiental. *Evaluación Psicológica / Psychological Assessment*.
- Castellan, N.J. (1979). The analysis of behavior sequences. In R.B. Cairns (Ed.) *The analysis of social interactions* (pp. 81-116). Hillsdale, N.J.: Lawrence Erlbaum Associates.

- Cerro, P. (1980). Aplicación de las técnicas de moldeamiento y auto-regulación al aprendizaje de saltos de aparatos en Educación Física. Tesis de Licenciatura. Madrid: Universidad Complutense.
- Clemente, M. y Gil, F. (1985). Prevención y tratamiento del comportamiento delictivo: Contribución desde una perspectiva psicosocial. En J.F. Morales, A. Blanco, C. Huici y J.M. Fernández Dols (Eds.) *Psicología social aplicada* (pp. 315-335). Bilbao: Desclee de Brouwer.
- Cordray, D.S., Bloom, H.S. & Light, R.J. (Eds.) (1987). *Evaluation Practice in Review*. San Francisco: Jossey-Bass.
- Culyer, A.J. (1982, Mayo). Marco para la evaluación multidisciplinaria de los Servicios Sociales. Ponencia presentada en las Primeras Jornadas de Economía de los Servicios Sociales, organizadas por el Departamento de Sanidad y Seguridad Social, Colegio de Economistas y Caja de Ahorros de Barcelona. Barcelona.
- Chacón, F. (1989). Necesidad social y Servicios Sociales. *Papeles del Psicólogo* (Madrid), 41/42, 18-24.
- Chen, H. & Rossi, P.H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, 7, 238-302.
- Del Valle, A. (1987). El modelo de Servicios Sociales en el Estado de las Autonomías, Madrid: Siglo XXI.
- Ericsson, K.A. & Simon, H.A. (1984). *Protocol analysis. Verbal reports as data*. Cambridge: The M.I.T. Press.
- ERS Standards Committee (1982). Standards for practice. In P.H. Rossi (Ed.) *Standards for evaluation practice* (pp. 7-20) San Francisco: Jossey-Bass.
- Farré, Ll. (1990). Dinámica emocional en l'assistència terapèutica al malalt neoplàsic. Anàlisi de les ansietats i recursos defensius mitjançant l'observació sistemàtica directa i indirecta. Tesis Doctoral no publicada. Barcelona: Universidad de Barcelona.
- Fassnacht, G. (1982). *Theory and practice of observing behaviour*. New York: Academic Press.
- Fernández Ballesteros, R. (1987, Septiembre). Ciencia, ideología y política en evaluación de programas. Conferencia pronunciada en el curso "Evaluación de programas y servicios para el bienestar social: Un enfoque pluridisciplinario". Santander: Universidad Internacional "Menéndez Pelayo".
- Fernández Ballesteros, R., Díaz, P. Izal, M. y González, J.L. (1987). Evaluación de una residencia de ancianos y valoración de intervenciones ambientales. En R. Fernández Ballesteros (Coord.) *El ambiente. Análisis psicológico* (pp. 227-248). Madrid: Pirámide.
- Fernández-Ballesteros, R., Izal, M. Díaz, P., González, J.L. & Souto, E. (1988). Training of conversational skills with institutionalized elderly: A preliminary study. *Perceptual & Motor Skills*, 66, 923-926.
- Fernández-Ballesteros, R., Montorio, I., Izal, M. Guerrero, M.A., Llorente, M.G. & Hernández, J.M. (1990, July). Improving memory among the elderly. XXII International Congress of Applied Psychology. Kyoto.

- Fernández del Valle, J. (1991a). Evaluación de contextos en Centros de Protección de Menores. Tesis Doctoral no publicada. Madrid: Universidad Autónoma de Madrid.
- Fernández del Valle, J. (1991b). El absentismo escolar en el municipio de Oviedo. Un estudio desde la perspectiva de los Servicios Sociales. Oviedo: Concejalía de Bienestar Social del Ayuntamiento de Oviedo y Universidad de Oviedo.
- Ferrero, C. (1990). Programa de detección, registro e intervención coordinada interredes de malos tratos en la infancia. *Infancia y Sociedad* (Revista de la Dirección General de Protección Jurídica del Menor), 2, 105-109.
- Freeman, H. & Rossi, P. (1981). Social experiments. *Health and Society: Milbank Memorial Fund Quarterly*, 59, 340-373.
- Fitz-Gibbon, C.T. & Morris, L.L. (1987). *How to design a program evaluation*. Beverly Hills: Sage.
- Forns, M. i Gómez, J. (1990). *Incidencia dels Programes d'Inmersió en el desenvolupament cognitiu-lingüístic i en el nivell d'adquisicions escolars en nens de preescolar. Informe de seguiment de dos anys del Programa d'Inmersió. Pàrvuls 4-5 anys <Incidencia de los Programas de Inmersión en el desarrollo cognitivo-lingüístico y en el nivel de adquisiciones escolares en niños de preescolar. Informe de seguimiento de dos años del Programa de Inmersión. Párvulos 4-5 años. SEDEC (Servei d'Ensenyament del Catalá). Barcelona: Departament d'Ensenyament de la Generalitat de Catalunya.*
- Generalitat de Catalunya i Ajuntament de Barcelona (1989, Julio). Programa d'actuació contra la pobresa i la marginació a Barcelona programa de actuación contra la pobreza y la marginación en Barcelona >. Publicación interna.
- Gottman, J.M. (1978). Nonsequential analysis techniques in observational research. In G.P. Sackett (Ed.) *Observing Behavior: Data collection and analysis methods* (pp. 45-61). Baltimore: University of Park Press.
- Gottman, J.M. (1980). The consistency of nonverbal affect and affect reciprocity in marital interaction. *Journal, of Consulting and Clinical Psychology*, 48, 711-717.
- Gottman, J.M. & Bakeman, R. (1979). The sequential analysis of observational data. In M.E. Lamb, S.J. Suomi & G.R. Stephenson (Eds.) *Social interaction analysis. Methodological issues* (pp. 185-206). Madison: University of Wisconsin Press.
- Gottman, J.M. & Notarius, C. (1978). Sequential analysis of observational data using Markov chains. In T.R. Kratochwill (Ed.) *Single subject research. Strategies for evaluating change* (pp. 237-285). New York: Academic Press.
- Hodgson, R.J. & Rachman, S. (1977). Obsessional-compulsive complaints. *Behavior Research and Therapy*, 15, 389-395.
- Howard, G.S., Schmeck, R.R. & Bray, J.H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement*, 16, 129-135.

- Kahn, A.H. & Kamerman, S.B. (1980). *Social Services in international perspective*. London: Transaction Books.
- Kidder, L. & Judd, C.M. (1986) *Research methods in social relations*. New York: Helt, Rinehart & Winston.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills: Sage.
- Lehner, P.N. (1979). *Handbook of ethological methods*. New York: Garland Press.
- Levinson, W., Gordon, G. & Skeff, K. (1990). Retrospective versus actual pre-course self-assessments, *Evaluation & the Health Professions*, 13 445-452.
- Lipsey, M.W. , Crosse, S., Dunkle, J., Pollard, J. & Stobart, G. (1985). Evaluation: The state of the art and the sorry state of the science. In D.S. Cordray (Ed.) *Utilizing prior research in evaluation planning* (pp. 7-27). San Francisco: Jossey-Bass.
- López Blasco, A. (1984). Proyectos de evaluación de la praxis en Servicios Sociales. En *Actas del Simposium Nacional de evaluación en Servicios Sociales* (pp. 9-31). Jaén: Patronato de Bienestar Social, Diputación Provincial de Jaén.
- Luque, O. (1989). Los Servicios Sociales como ámbito de intervención de la Psicología. *Papeles del Psicólogo* (Madrid), 41/42, 34-44.
- Mahoney, M.J., Moura, N.G. & Wade, T.C. (1973). The relative efficacy of self-reward, self-punishment and self-monitoring techniques for weight loss. *Journal of Consulting and Clinical Psychology*, 40, 404-407.
- Markoff, J., Shapiro, G. & Weitman, S. (1975). Toward the integration of content analysis and general methodology. In D. Heise (Ed.) *Sociological Methodology 1975* (pp. 1-58). San Francisco: Jossey-Bass.
- Moss, R.H. (1982). Improving social settings by social climate measurement and feedback. In R.F. Muñoz, L.R. Snowden & J.G. Kelly (Eds.) *Social and psychological research in community settings* (pp. 145-182), San Francisco: Jossey-Bass.
- Moreno, R. (1984). Métodos y técnicas de investigación: Diferenciaciones terminológicas. *Apuntes de Psicología* (Sevilla), 6, 9-11.
- Muñoz, J. (1984). Metodología de intervención en el seguimiento y evaluación de los programas de ayuda social a la infancia. En *Actas del Simposium Nacional de evaluación en Servicios Sociales* (pp. 99-125). Jaén: Patronato de Bienestar Social, Diputación Provincial de Jaén.
- Norris, S.P. (1981). A concept of observation statements. *Philosophy of Education*, 132-142.
- Norris, S.P. (1984). Defining observational competence. *Service Education*, 68, 129-142.
- OMS (Organización Mundial de la Salud) (1981). *Evaluación de los programas de salud*. Ginebra: OMS, Serie "Salud para todos" n° 6.
- Palumbo, D.J. (1987). Politics and evaluation. In D.J. Palumbo (Ed.) *The politics of program evaluation* (pp. 12-46). Beverly Hills: Sage.

- Pancer, S.M. & Westhues, A. (1989). A developmental stage approach to program planning and evaluation. *Evaluation Review*, 13, 56-77.
- Pleite, A. y Pardo, C. (s/f). *Programa de atención especializada para el tratamiento de conductas adictivas en sujetos internados en los Centros Penitenciarios de Cataluña*. Barcelona: Departamento de Justicia de la Generalitat de Catalunya, Dirección General de Servicios Penitenciarios y de Rehabilitación.
- Posavac, E.J. & Carcy, R.G. (1985). *Program evaluation. Methods and case studies*. Englewood Cliffs, N.J.: Prentice-Hall.
- Quera, V. (1991). Muestreo y registro observacional. En M.T. Anguera (Ed.) *Metodología observacional en la investigación psicológica* (pp. 241-329). Barcelona: P.P.U., Vol. I.
- Redondo, S. (en prensa). Entorno penitenciario y reinserción social. En S. Redondo. *Evaluar e intervenir en prisiones*. Barcelona: P.P.U.
- Redondo, S., Pérez, E., Agudo, F., Roca, M. i Azpiazu, M. (Dirs.) (1990). *Programes de rehabilitació a les presons* <Programas de rehabilitación en las prisiones>. Barcelona: Direcció General de Serveis Penitenciaris i de Rehabilitació. Departament de Justicia de la Generalitat de Catalunya.
- Redondo, S., Roca, E., Sánchez, A. y Deumal, E. (en prensa). Diseño ambiental de una prisión de jóvenes: Cinco años de evaluación. En S. Redondo. *Evaluar e intervenir en prisiones*. Barcelona: P.P.U.
- Riba, C. (1991). Metodología observacional: Decisiones básicas y objetivos. En M.T. Anguera (Ed.) *Metodología observacional en la investigación psicológica*. (pp. 29-114). Barcelona: P.P.U., Vol.I.
- Richart, M. y Cabrero, J. (1988). *La psicología en los centros de salud comunitaria. Programa de Promoción y Educación para la Salud*. Valencia: Promolibro.
- Rodrigo, M.J. (1982). Las posibilidades del análisis de tareas como técnica para el estudio de los procesos mentales. *Infancia y Aprendizaje*, 19-20, 159-173.
- Rossi, P.H. & Freeman, H.E. (1985). *Evaluation. A systematic approach*. Beverly Hills: Sage.
- Rueda, J.M. (1989). Buscando un esquema conceptual referencial y operativo. *Papeles del Psicólogo* (Madrid), 41/42, 25-33.
- Rutter, M. & Giller, H. (1988). *Delincuencia juvenil*. Barcelona: Martínez Roca.
- Sackett, G.P. (Ed.). *Observing Behavior: Data collection and analysis methods*. Baltimore: University Park Press, vol. II.
- Sackett, G.P. (1979). The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In J.D. Osofsky (Ed.) *Handbook of infant development* (pp. 623-649). New York: Wiley.
- Sackett, G.P. (1980). Lag sequential analysis as a data reduction technique in social interaction research. In D.B. Sawin, R.C. Hawkins, L.O. Walker & J.H. Penticuff (Eds.) *Exceptional infant. Psychosocial risks in infant-environment transactions* (pp. 300-340). New York: Brunner/Mazel.
- Santoyo, C. y Espinosa, C. (1987). Un sistema de observación conductual de Interacciones Sociales. *Revista Mexicana de Análisis de la Conducta*, 13, 235-253.
-

- Shadish, W.R., Cook, T.D. & Leviton, L.C. (1991). *Foundations of program evaluation*. Beverly Hills: Sage.
- Smith, M.F. & Bobrow, L.B. (1991). Study of the effectiveness of a health-risk reduction program. *Evaluation & the Health Professions*, 14, 88-99.
- Stufflebeam, D.L. (1978). Meta-evaluation: An overview. *Evaluation & the Health Professions*, 1, whole numer.
- Thomas, E.A. & Malone, T.W. (1979). On the dynamics of two-person interaction. *Psychological Review*, 86, 331-360.
- Vency, J.E. & Kaluzny, A.D. (1984). *Evaluation and decision making for health services program*. Englewood Cliffs, N.J.: Prentice-Hall.
- Vizcarro, C. y García Marcos, J.A. (1981). Los auto-informes. En R. Fernández Ballesteros y J.A.I. Carrobles (Eds.) *Evaluación conductual: Metodología y aplicaciones* (pp. 233-266). Madrid: Pirámide.
- Weick, K.E. (1968). Systematic observational methods. In G. Lindzey & E. Aronson (Eds.) *Handbook of social psychology* (pp. 357-451). Reading, Mass.: Addison-Wesley, vol. II.
- Wertheimer, M., Barclay, A.G., Cook, S.W., Kiesler, C.A., Koch, S., Riegel, K.F., Rorer, L.G., Senders, V.L. Smith, M.B. & Sperling, S.E. (1978). Psychology and the future *American Psychologist*, 33, 631-647.
- Zeiler, M.D. (1986). Behavioral units: A historical introduction. In T. Thompson & M.D. Zeiler (Eds.) *Analysis and integration of behavioral units* (pp. 1-12). Hillsdale, N.J.: Lawrence Erlbaum Associates.