

Construccionismo, postmodernismo y teoría de la evaluación. La función estratégica de la evaluación

Constructionism, postmodernism and evaluation theory. The strategic function of evaluation

Baltasar Fernández-Ramírez

Universidad de Almería

bfernan@ual.es

Resumen

La evaluación es una empresa lucrativa y un perfil profesional al que muchos se han sumado careciendo de las mínimas competencias y conocimientos exigibles. Evaluar es emitir juicios de valor sobre objetos diversos, incluidas las iniciativas de intervención social. El planteamiento convencional es racionalista (diagnóstico, planificación, intervención y análisis del impacto), a pesar de que los modelos racionalistas hayan perdido vigor en las ciencias sociales a favor de alternativas políticas, simbólicas o caóticas, tal como apreciamos en la historia del pensamiento organizacional. En una perspectiva construccionista, observamos además el poder reificador de los criterios escogidos en cada situación de evaluación. La evaluación sugiere o impone qué es lo correcto en cada caso; criterios que pasan a acotar, dar sentido y prioridad a determinado modo de entender las relaciones sociales, los objetivos o el modo de comportamiento personal y organizacional. Mi propuesta es elevar esta inevitable consecuencia práctica de las evaluaciones a rango de virtud política y ética, al servicio de los intereses y deseos de futuro de los grupos que participan en procesos evaluativos, primando su función estratégica sobre las tradicionales funciones de mejora y rendición de cuentas.

Palabras clave: Evaluación estratégica; Pensamiento organizacional; Postmodernismo

Abstract

"Testing" and "evaluating" are money-making activities that offer a rewarding professional career; perhaps it is no surprise that the field includes numbers of self-styled "evaluators" who provide services without the required training. Evaluation means the systematic application of given criteria for judging the merit or value of a variety of 'assessables' - capacities, powers and services, including social policy interventions. The conventional perspective is rationalist (in diagnosis, planning, intervention and impact analysis), even though rationalism have lost their pre-eminence in the social sciences, in favour of political, symbolic or chaotic alternatives, as the history of organizational thought reveals. Alternatively, from a constructionist perspective, we see the power of evaluation criteria to reify what they ostensibly assess. Evaluations suggest or impose what is right; criteria of value define, give meaning and prioritize just one way of understanding social relations, program goals or individual and organizational behaviour. The suggestion in this article is to make a positive virtue of this inevitable pragmatic consequence of evaluation: to openly embrace it as an agent of change, emphasizing a strategical function beyond the traditional purposes of enhancement and accountability.

Keywords: Social constructionism; Testing; Evaluation; Postmodernism; Organisational strategy

Han pasado más de cuarenta años desde que la demanda de evaluaciones se extendiera por muy distintos ámbitos profesionales y académicos norteamericanos. Cuarenta años desde la publicación de los primeros trabajos de Michael Scriven, E. Suchman, Daniel L. Stufflebeam o el popular libro de Donald T. Campbell y Julian C. Stanley sobre diseños cuasiexperimentales. En España, la primera generación de

evaluadores¹ académicos introdujeron el tema durante los años 80, por ejemplo, a través de la evaluación psicológica (Fernández-Ballesteros, 1995), la Psicología social (Reboloso, 1987) o la Sociología (Alvira, 1991). Nosotros, los que podemos considerarnos la segunda generación de los evaluadores españoles, hemos vivido en primera persona la impresionante expansión de los estudios de evaluación, la demanda profesional creciente y la incipiente institucionalización, con manuales, programas de formación y asociaciones profesionales (Fernández-Ramírez y Reboloso, 2006). Supongo que debemos sentirnos satisfechos. A estas alturas, ¿quién no ha escuchado e incluso ha participado en procesos de evaluación dentro de su organización, de su departamento o empresa, o no ha visto sometido su trabajo al escrutinio de un proceso evaluativo?

Sin embargo, este breve ensayo tiene como objetivo plantear una visión crítica de la práctica de la evaluación. La situación profesional en España dará pie a plantear el concepto de evaluación como juicio y a señalar las dificultades para responder a las demandas de evaluación desde una perspectiva científica convencional. Los modelos racionalistas (evaluaciones experimentales y evaluaciones sistemáticas) deben ser replanteados en coherencia con el pensamiento organizacional contemporáneo, el cual analiza las organizaciones como sistemas sociales políticos, simbólicos y caóticos. La evaluación de cuarta generación puede servir como alternativa, aunque quisiera ir un paso más allá y discutir, en la última parte del texto, sobre el poder que tienen los sistemas de evaluación para redefinir y construir la realidad organizacional en función de los intereses, intenciones y perspectivas que tienen los diferentes grupos que las componen. Asumiendo posiciones construccionistas críticas, terminaré reflexionando sobre un cambio en nuestro modo de plantear las tradicionales funciones de la evaluación, proponiendo que se entienda como una herramienta estratégica útil para dirigir el cambio social y organizacional en procesos democráticos de toma de decisiones.

La lógica de la evaluación

La realidad española continúa en la línea de la década pasada (Fernández-Ramírez y Reboloso, 2006). Muchos, incluso muchísimos, hablan de evaluación, pero muy pocos han recibido o han buscado la formación mínima que pudiera cubrir las competencias esenciales definidas por asociaciones de indiscutible renombre, por ejemplo, la Asociación Americana de Evaluación (King, Stevahn, Ghery y Minemma, 2001). He tratado directamente con muchos en los congresos de evaluación, en reuniones de trabajo con profesionales de diferentes campos e impartiendo clases en los distintos másteres que se ofertan en nuestro país. Muchos tienen responsabilidades de evaluación, pero buscan formación porque carecen de los fundamentos mínimos; otros saben lo que han aprendido en su campo específico de trabajo (cooperación para el desarrollo, gestión de la calidad, gestión ambiental, etc.), pero ignoran la gran literatura sobre el tema; muchos colegas en Psicología usan el término en sus estudios, pero con un sentido diferente (investigación sobre la dimensión evaluativa de la actitud). En fin, muchos reciben el

¹ Quisiera seguir la sugerencia de los editores de la revista y emplear un término que evitara el riesgo de sexismo contenido en el uso generalizado del masculino al tratar de colectivos o de una persona sin especificar. Quisiera también seguir los dictados de la Real Academia Española de la Lengua, y no caer en la incorrección gramatical de mencionar a cada paso fórmulas de compromiso tales como "los evaluadores y las evaluadoras" o barbaridades del tipo "los/las evaluadores/as". Comprendo el problema de la sensibilidad femenina, pero también entiendo que toda crítica, incluida la feminista, va dirigida contra una situación de poder injusta, y creo que el uso del masculino o femenino en este ensayo no debería ser considerado una amenaza para las evaluadoras españolas, que son muchas y cuya competencia no está en duda. Vaya por delante mi reconocimiento para ellas, y mi disculpa si alguien siente herida su sensibilidad.

encargo de evaluar, y hacen lo que pueden honestamente, si bien con resultados poco satisfactorios para un evaluador académico y, sobre todo, desconociendo el campo amplio de investigación al que internacionalmente denominamos Evaluación.

Esta situación me sugiere varias reflexiones posibles. Primero, que la demanda sigue yendo por delante de la especialización, de la diferenciación como perfil profesional, revelando una de las caras del asunto, a saber, el gran negocio que es el mundo de las evaluaciones. Muchos se apuntan sin disponer de las competencias requeridas, y medran porque tampoco quienes les contratan tienen conocimiento y criterio para distinguir lo que está o no correctamente planteado.

En segundo lugar, que se realizan evaluaciones en campos profesionales muy diferentes (por ejemplo, el diagnóstico y la prueba clínica en contextos de salud, la evaluación de impacto en las obras públicas, la valoración de candidatos en la selección de personal, la revisión y el cambio de sistemas de gestión, etc.), que implican a profesionales también muy diferentes (de la Medicina, de la Química, de la Ingeniería, de la Psicología, de la Gestión, de la Economía, de la Educación, etc.), sin que exista conciencia de estar desarrollando un trabajo similar, y sin conocer ni reconocer la existencia de un campo general de conocimientos transversales (transdisciplinares, en acertada expresión de Michael Scriven (1991), al que denominados Evaluación. En la práctica, les resulta suficiente conocer algunos conceptos teóricos y algunas soluciones metodológicas para resolver los problemas prácticos de las demandas que encuentran en su campo, pero no salen de este ni tienen nada que aportar a nuestro campo común.

Y tercero, que se carece de una concepción común sobre lo que significa evaluar y cuáles son los servicios que presta un evaluador profesional. *Evaluar es valorar* –otra vez siguiendo el magisterio de Michael Scriven, 1994, y la opinión de la inmensa mayoría de los especialistas (Stufflebeam, Rossi y Freeman, Stake, Weiss, etc.)–. Evaluar es juzgar el mérito o la valía de un objeto, sea el que fuere (un producto de consumo, un programa social, un proceso organizacional, el desempeño de un empleado, etc.); evaluar es afirmar, concluir, utilizando procedimientos aceptables² de investigación, sobre el valor o la calidad del objeto, sobre si es bueno o malo, adecuado o no, apropiado, deseable o correcto para el fin que se persiga. Nadie cuestiona esta posición. Si acaso, se podría matizar, con Michael Q. Patton (1996), que *evaluar es valorar, y muchas otras cosas*, dada la complejidad de las tareas que afronta un profesional ante la demanda de evaluación. E incluso aquí, me atrevería a llevarle la contraria: evaluar es valorar, además de muchas otras cosas *que se realizan con el objetivo último de valorar* si el objeto es válido o no desde algún punto de vista, si es recomendable su uso (o su implantación, en el caso de los programas de intervención social) y si debe ser mejorado en algún aspecto (Reboloso, Fernández-Ramírez y Cantón, 2003, 2008a).

Esta definición no es discutible –lamento afirmarlo con tal rotundidad ante mis colegas españoles–, al menos, si queremos vincular nuestro trabajo e investigaciones con la gran corriente de la literatura internacional que en los últimos cuarenta años se ha desarrollado bajo la etiqueta de Evaluación. Es un presupuesto básico, sin el cual, el resto de esta exposición carece de sentido. Supongo que la lectura de este texto no ayudará mucho a convencerles, y me resultaría harto extraño que incluso un texto como este estuviera entre sus manos. Aún así, les rogaría que esperaran al final del mismo antes de decidir –valorar, juzgar, evaluar– si les parece o no satisfactorio, válido o útil como un concepto marco que nos

² Estaba tentado de usar la expresión “procedimientos científicos de investigación”, pero no lo haré más que en esta nota a pie de página, puesto que las discusiones sobre qué es o no científico son muchas, y veo justificada la utilización de métodos que usualmente no se consideran científicos, cuando la ocasión lo requiera.

diferencia de la investigación científica tradicional (positivista, operacionista, objetivista, libre de valores, etc.)³.

Michael Scriven (1994) sintetiza el procedimiento básico para la realización de las evaluaciones en lo que ha venido a ser reconocido de manera extendida como la *lógica de Scriven* o la *lógica de la evaluación*, tal ha sido el éxito de la propuesta. Le sirven de modelo los análisis comparativos que periódicamente publican distintas asociaciones o editoriales independientes para recomendar la mejor opción de compra a los consumidores de una gama concreta de productos (automóviles, ordenadores, productos de consumo en general). La cuestión estriba en escoger un conjunto reducido de criterios de valor que definan qué es lo mejor, lo deseable en el producto en cuestión, traducir cada criterio en indicadores (mensurables⁴), y medir después cada uno de los productos disponibles, para concluir finalmente con una puntuación relativa que indica la mejor compra posible. El procedimiento es aplicable a cualquier objeto de evaluación que se nos ocurra (un alimento, un servicio hotelero, un profesor, un proceso de curación, etc.); la diferencia radica en que los criterios de valor relevantes y el conocimiento exigible al evaluador para determinarlos serán acordes a las características del objeto.

Las dificultades quedan bien resumidas en dos preguntas: ¿quién decide cuáles son los criterios de valor que se utilizarán? y ¿cómo traducir cada criterio en indicadores mensurables que cubran satisfactoriamente nuestros criterios de validez? Con un ejemplo, si evaluamos la función docente de un profesor, ¿decidirán los criterios el claustro de profesores del centro escolar en cuestión, sus alumnos, la delegación de educación o serán derivados de alguna teoría de educación en boga?, ¿serán decididos por técnicos ajenos al centro en su función de evaluadores externos, se dejarán en manos de los propios implicados, o se buscará algún tipo de solución intermedia? Cada una de estas opciones dará un resultado más o menos diferente. Al respecto, se derivan una serie de interesantes discusiones, que intentaré retomar en un apartado posterior, cuando hable de una visión política y construccionista de la evaluación.

En cuanto al problema de la validez, la solución pasa por encontrar indicadores que cuenten con el consenso y la aprobación de los grupos implicados, y que sean definidos siguiendo las directrices incluidas en los manuales de métodos y técnicas de investigación y evaluación; con el gran escollo, sin embargo, de que caeremos casi sin darnos cuenta en la tradicional y aburrida polémica entre los llamados métodos cuantitativos y cualitativos. Y conste que digo aburrida, no porque carezca de interés y de profundidad, que los tiene, sino porque la interpretación más extendida se ha visto reducida a una simplificación, según la cual, los métodos cuantitativos cumplen mejor con los criterios de validez –desde la propia perspectiva positivista convencional que les da sustento epistemológico, evidentemente–, y son preferidos para la comprobación empírica, quedando los otros como meras estrategias menores para realizar análisis exploratorios, en espera de que las ideas y los datos estén en condiciones de pasar por un filtro cuantitativo. Mucho se puede hablar al respecto; bastará con indicar que los evaluadores han

³ Aquí entramos ya en terrenos pantanosos. Igual que en otras ciencias sociales, la Evaluación cuenta con cierta diversidad de orientaciones epistemológicas y metodológicas, entre las que podemos incluir una versión convencional identificada con las características mencionadas (evaluación orientada por los objetivos, evaluación de corte experimental centrada en el análisis de los resultados de la intervención), pero no es la única.

⁴ Usualmente, se equipara el concepto de indicador con alguna variable mensurable o cuantificable, y así es en la mayoría de los casos, pero no es esta mi intención. La idea, tal como yo la entiendo, es más bien encontrar elementos de reflexión, cuantitativos o no, que sirvan como indicio para concluir sobre el estado del objeto en el criterio del que tratemos. No se olvide que todo indicador cuantitativo no es más que una simplificación con la que tratamos de captar los matices cualitativos del objeto, que se violentan cuando el discurso se organiza en torno a las cualidades del número (su comportamiento dentro de una matriz de covarianza, por ejemplo), y no a las del objeto.

discutido sobre cuáles son los criterios de validez aplicables en nuestro campo, y las respuestas no coinciden por completo con los criterios de validez clásicos, que, en el mejor de los casos, son aceptados dentro de un conjunto mayor de criterios relevantes para considerar que una evaluación está bien realizada (Reboloso, Fernández-Ramírez y Cantón, 2008a)⁵.

Pensamiento racional y pensamiento político

La evaluación es una empresa racional, tal como afirmaban Peter H. Rossi y Howard E. Freeman en su bien conocido manual (Rossi, Lipsey y Freeman, 2004); es decir, aplica una teoría racional en la toma de decisiones en el contexto de los programas de intervención social: dado cierto objetivo de intervención (reducir una necesidad social, proveer de un servicio), pensamos en las distintas alternativas de acción disponibles, analizamos (valoramos) sus ventajas y desventajas, ponemos en práctica la opción preferida en una situación controlada (ensayo piloto), analizamos (evaluamos) los resultados y extendemos el programa si estos fueron satisfactorios. El planteamiento conecta directamente con las reflexiones de Kurt Lewin (1946) sobre la intervención social, de quien Peter Rossi y Howard Freeman se reconocen deudores. Visto en términos amplios, también forma parte de una tradición en la que se reunirían la gestión científica del trabajo, la gestión y evaluación por objetivos, la planificación racional o la creencia en las posibilidades de realizar una política social científicamente dirigida y experimentalmente comprobada (Campbell, 1969).

La intervención racionalista (tecnocrática) ha sobrevivido a diversos movimientos intelectuales y políticos a lo largo del siglo XX. Aún da muestras de su potencia, precisamente en la expansión de la evaluación como empresa que se justifica por su función informativa (obtener mejores datos para fundamentar mejores decisiones), o en modas tan extendidas como los sellos de calidad como paradigma del buen gobierno de todo tipo de organizaciones, desde las culturales y educativas hasta las sanitarias, las judiciales o las de producción hortofrutícola.

La cima del pensamiento racionalista en nuestro campo llega con las denominadas evaluaciones sistemáticas, que aúnan satisfactoriamente teoría de la gestión (planificación estratégica, gestión por objetivos, gestión por procesos) y racionalidad científica (decisión basada en la evidencia, diseños experimentales y cuasiexperimentales), superando las estrecheces que ambas encuentran en sus respectivos campos de aplicación originales. Siguiendo la senda trazada por Kurt Lewin (1946), las evaluaciones sistemáticas están representadas en las propuestas de Peter Rossi y Howard Freeman (1989), el CIPP de Daniel Stufflebeam (2001, 2002), y muchos otros autores que plantean variaciones más o menos originales sobre el modelo, entre los cuales, modestamente, nos encontramos (Reboloso, Fernández-Ramírez y Cantón, 2003, 2008a). En fin, la evaluación no se realiza en un momento puntual al finalizar la intervención (evaluación de resultados), sino que se extiende y acompaña a cada una de las fases del proceso global de la intervención social, proponiendo actividades evaluativas específicas en

⁵ Frente a la opción convencional, identificada con los trabajos de Donald T. Campbell y sus colegas (Campbell y Stanley, 1963; Cook y Campbell, 1979), Egon G. Guba e Yvonne S. Lincoln (1989) sugieren un conjunto de criterios acordes con su orientación constructorista. En un intento de integración, Huey-Tsyh Chen (1990) propone los criterios de responsividad (oportunidad, relevancia, amplitud de consecuencias, igualdad), objetividad (fiabilidad, ausencia de sesgos), confianza (validez interna, certeza) y generalizabilidad (validez externa). Desde hace años, Stufflebeam propone un amplio conjunto de estándares para la práctica profesional, agrupados bajo de las normas de utilidad, viabilidad, honradez y precisión. Quizá la diferencia más marcada respecto de una visión convencional de la ciencia es la consideración del criterio de utilidad con el mismo peso que los conceptos tradicionales de validez científica.

cada una de ellas. Por ejemplo, se evalúan las necesidades de la población en desventaja, se evalúa si el plan de la intervención es viable y coherente, se realiza el seguimiento de la intervención –evaluación del proceso– y se evalúan los resultados e impactos de segundo orden originados como consecuencia de la misma intervención. Asimismo, persigue los tradicionales objetivos de mejora, rendición de cuentas y desarrollo del conocimiento de manera integrada y ordenada (es decir, sistemática). Cientifista, integral, contingente y heurística serían algunas de sus características (Reboloso, Fernández-Ramírez y Cantón, 2008a).

El racionalismo forma parte del auge de las ciencias positivas durante el siglo XIX, pero la historia intelectual del siglo XX, desde cierto punto de vista, es la historia de la depreciación del racionalismo como modelo universalmente válido en cualquiera de los campos del saber y del hacer humano⁶. La relatividad, la incertidumbre, el giro lingüístico, el desorden y el caos son conceptos extendidos y asentados en nuestro bagaje teórico actual (Munné, 1993; Jiménez Burillo, 1997). Hablemos, por ejemplo, sobre el cambio en el pensamiento organizacional a lo largo del siglo, tan relevante para nosotros en las cuestiones de la planificación y la implantación de los programas de intervención social (Ramió y Ballart, 1993; Reboloso, Fernández-Ramírez y Cantón, 2003).

Hasta la década de los veinte, aproximadamente, la burocracia weberiana y la gestión científica del trabajo son la clave para comprender y gestionar las organizaciones. La organización se define por una estructura de grupos de trabajo, con relaciones mutuas, objetivos y requerimientos de tarea establecidos racionalmente. La división especializada y la estructura formal serían los elementos clave. A partir de estas fechas, el auge de las escuelas de relaciones humanas y recursos humanos introducen un nuevo factor clave que no puede ser obviado: la persona. Las personas que forman parte de la organización tienen objetivos, motivos y valores propios que pueden ser cruciales para un funcionamiento eficiente, más allá de los problemas de la estructura organizacional. Más aún, las personas forman parte de grupos que no son definidos ni controlados por la dirección y cuya influencia puede incluso generar conflictos (la estructura informal; son miembros de un sindicato, una cofradía, colegas de una asociación o miembros de una misma familia, por ejemplo). La teoría organizacional asume entonces que el asunto es más complicado, puesto que hay que conjugar las estructuras formales e informales, y preocuparse por ciertos factores de índole psicológica, si bien todavía puede ser gestionado racionalmente si disponemos de suficiente información, gracias a los departamentos de personal o de recursos humanos, por ejemplo. Y más; en los años cincuenta, hace su aparición y triunfa con claridad la teoría de sistemas (Katz y Kahn, 1966) que, entre otras cuestiones, introduce la idea de que la organización es un sistema abierto, es decir, un elemento más que interacciona, depende e influye sobre otros elementos similares que forman el macrosistema de un sector económico o de servicios (sanitario, educativo, financiero...). Comprender y gestionar una organización, en una visión puramente ecológica (sistémica), exige comprender los elementos, las relaciones y los cambios que se producen en las demás organizaciones e instituciones relevantes dentro de su sector. Nada que, sin embargo, no pueda ser controlado si se

⁶ En cuanto teoría sobre la toma de decisiones, el racionalismo es útil a la especie humana como práctica de dominación o de intervención en la naturaleza (es pragmático, no epistémico), pero derivar de aquí que la naturaleza social es racional supone un salto difícilmente asumible. Que mi pensamiento racionalista me permita decidir cómo comportarme ventajosamente en un medio social, no convierte a éste en un medio racional. Sería confundir los términos. Además, gran parte de la investigación cognitiva en psicología y psicología social demuestra precisamente cuánto nos equivocamos, los atajos que usamos y de qué modo nos engañamos a nosotros mismos cuando razonamos. Más que el animal racional, es más adecuado el conocido símil del *indigente* o el *perezoso cognitivo* (Rodríguez, 1995).

dispone de buena información, por ejemplo, gracias a diversos departamentos del *staff* de apoyo (investigación y desarrollo, análisis de mercado, comercialización).

De este modo, durante cien años, se han venido ampliando el número de los elementos que deben ser considerados para comprender y gestionar las organizaciones, pero sin variar el planteamiento racionalista de fondo, según el cual, hay que orquestar mecanismos para disponer de suficiente información sobre estos múltiples elementos, con la cual planificar mejor los procesos de trabajo y la distribución de recursos para lograr los fines estratégicos de la empresa (Mintzberg, Lampel, Quinn y Ghoshal, 2006). Una perspectiva de evaluación sistemática, como la descrita más arriba, encaja perfectamente con este planteamiento, por cuanto que en ambos casos se utilizan del mismo modo los conceptos de diagnóstico, planificación, seguimiento y evaluación de los resultados, como herramientas adecuadas para la gestión, bien sea de la organización mercantil, la institución de servicios o los programas de intervención social.

Todo esto ha cambiado desde los años setenta, aproximadamente, en los que nuevos modelos teóricos se suman a los anteriores y, en parte, suponen un corte radical (véase el capítulo primero de Reboloso, Fernández-Ramírez y Cantón, 2003). Creo que, aún pecando mucho de simplificador, podríamos reducirlos a tres. Primero, la organización es un sistema político, en el que diversos grupos persiguen objetivos propios (el poder es un objetivo capital) y utilizan las estrategias y recursos disponibles para imponer su perspectiva sobre cómo debe ser la organización (Pfeffer, 1998). De ahí, la interesante observación de que, frente al clásico aserto de que “las metas preceden a la estructura organizacional”, ahora se defiende que sólo pueden perseguirse las metas que resultan viables dentro de esta estructura de grupos, intereses y presiones; la estructura es previa a las metas de la organización. O la acertada impresión del profesor Andrés Rodríguez, al definir la organización (pública, en este caso) como una “anarquía organizada” donde el poder es difuso, las metas se negocian antes de traducirse técnicamente, y hay que pensar en un mosaico de influencias y controles, más que en un organigrama jerárquico presidido por una cúpula central (Rodríguez y Ardid, 1996). Segundo, que la organización tiene una suerte de realidad virtual (es realidad social construida), cuya existencia se identifica con un conjunto de significados (prácticas, valores, creencias), con los que cada grupo define, entiende y afronta su posición respecto de los restantes grupos y elementos presentes (Weick, 1969). La organización sería, por tanto, un sistema psicosocial complejo, que es interpretado, construido y asumido por cada grupo de manera peculiar, en función de sus intereses, creencias previas, normas, valores y factores similares. Es decir, no es que cada grupo construya una visión particular ante una organización que existe objetivamente fuera de ellos: esas construcciones son la organización. ¿Quién, qué grupo, en este contexto, podría afirmar que tiene la explicación o la imagen “verdadera” de la organización, sin convertirse automáticamente en sospechoso de intentar imponer una visión interesada del asunto? Y tercero, desde los años noventa, se vienen revisando algunas propuestas clásicas de la teoría de sistemas, fundamentalmente el concepto de “equilibrio”, llegando a la idea de que una situación equilibrada es algo completamente excepcional y momentáneo (una instantánea que sólo existe en el momento de la fotografía, si se me permite el símil), siendo más correcto afirmar que la organización es un sistema dinámico alejado del equilibrio (entrópico, desordenado, caótico), en el que distintas fuerzas producen cambios continuos que forman y transforman la realidad organizacional en modos no previsibles (Munné, 2007; Navarro, 2005). De ahí, en mi opinión, que sea acertado el énfasis actual en la innovación y la flexibilidad que se escucha en los círculos de gestión, frente a un modelo organizacional estático (equilibrado) que no soportaría los vaivenes del mercado o del sector económico y social

correspondiente. De ahí también la defensa de la flexibilidad frente a la clásica rigidez burocrática de las administraciones públicas (de nuevo Rodríguez y Ardid, 1996).

Retomemos el discurso en este punto. La empresa evaluativa cuenta con un arsenal de conceptos y métodos racionalistas, que podríamos llamar convencionales, entre los que incluiríamos los diseños experimentales y los modelos sistemáticos. Y cuenta, en línea con el espíritu de nuestro tiempo, con modelos alternativos, no racionalistas, sino políticos y construccionistas (postmodernos, podríamos apuntar, si bien no todos se sentirían cómodos con la etiqueta). La rancia (en el doble sentido castizo de la expresión) polémica de los métodos cualitativos y cuantitativos ilustraría las dos posiciones, aunque personalmente intentaría ir más allá, para explorar nuevos planteamientos rupturistas. Sin duda, entre los modelos más conocidos y aceptados, la evaluación de cuarta generación es la que mejor se aleja de una perspectiva racionalista convencional, con su crítica radical en el terreno ontológico (no existen realidades objetivas, sino construcciones sociales), epistemológico (el conocimiento se construye en la interacción entre el observador y el objeto) y metodológico (defensa de la hermenéutica para construir nuevas realidades en el proceso de evaluación) (Guba y Lincoln, 1989). Todo un desafío al pensamiento dominante en gran parte de nuestras ciencias sociales y entre muchos de nuestros colegas. Por cierto, una propuesta que, en la práctica, guarda un estrecho parecido con los sistemas de acreditación, al menos en el procedimiento de creación del sistema a través de un extenso período de discusión a múltiples bandas entre diferentes grupos de interés, donde el evaluador recoge, clarifica, organiza y difunde entre cada grupo la perspectiva de los demás, con vistas a la creación de un conjunto final de valores e indicadores que servirán como estándares de calidad válidos para todas las organizaciones participantes (Philips, 2001).

Las nuevas funciones de la evaluación

Prácticamente, de todos es conocida la distinción entre la función formativa y sumativa de la evaluación. Son muchos menos los que conocen la polémica del año 1996, en la que el mismo Michael Scriven, creador de ambos conceptos, renuncia a su utilización más que como elementos para la reflexión, a los que no se debe dar mayor importancia. Máxime, teniendo en cuenta la confusión existente en su uso. La idea de que los datos evaluativos sirven, bien para dar forma o reformar el programa, en su fase de definición o de implementación, o bien para informar a los grupos interesados sobre las actuaciones y logros del programa, se ha complicado y matizado con los años. Ahora sería necesario dedicar un capítulo completo para exponer las diferentes posibilidades con las que se han enriquecido los usos e impactos potenciales de la evaluación. Por ejemplo, la enorme diferencia entre el concepto de mejora, propio de un discurso gestorlista y racionalista, y el concepto de fortalecimiento, sobre todo cuando se utiliza en contextos de intervención para el desarrollo, donde se alimenta de tradiciones políticas e ideológicas de izquierda, que lo acompañan de un discurso sobre el potencial *liberador* y *emancipador* de la intervención social (Blanco y Valera, 2007). Eleanor Chelmsky (1997) sintetizó un buen número de estas posibilidades en tres categorías de funciones, en las que coincide a grandes rasgos con las propuestas anteriores de Daniel Stufflebeam (Stufflebeam, 2001) y Peter Rossi y Howard Freeman (1989). Así, las evaluaciones se realizan para cumplir con el requisito de demostrar que la responsabilidad asumida por el programa (la institución, el personal técnico o quien fuere) se está cumpliendo correctamente; para colaborar en el desarrollo de la intervención, empleando desarrollo en su doble acepción de puesta en práctica adecuada y de maduración o crecimiento, como cuando

entendemos que los evaluados han aumentado sus competencias o habilidades gracias a su participación en el proceso, y por tanto *han crecido* como profesionales; y para conocer mejor sobre los problemas que afronta la intervención, sobre el funcionamiento de las organizaciones o sobre el modo de resolver los problemas.

Personalmente, la idea de que la evaluación sirve para aumentar el conocimiento siempre me ha parecido un argumento para justificar la realización de evaluaciones que podríamos denominar académicas; cuyos resultados pueden tener un gran interés, y no las critico por eso, pero que son ajenas a la realidad de los programas y a un planteamiento verdaderamente aplicado de la investigación científica. Lo que algunos llaman ciencia aplicada, no es más que ciencia normal desarrollada en contextos no usuales o alejados del entorno controlado de los laboratorios experimentales y los departamentos universitarios. Por otra parte, es evidente que todo estudio hace que nuestro conocimiento aumente, lo hagamos con la idea de demostrar la responsabilidad cumplida o con el objetivo de obtener información que sirva para la mejora continua de las operaciones del programa. No las critico, insisto, ni les niego utilidad. Creo que son pertinentes dentro de la discusión sobre ciencia básica o aplicada, que los académicos conocemos tan bien, pero que es ajena al problema de las funciones que cumple la evaluación en el contexto de los programas de intervención. La dejaremos a un lado, por lo tanto, para centrarnos en las otras dos funciones. (Para aquellos interesados, remitimos al capítulo del profesor José Ramón Torregrosa (1996) para un replanteamiento de la discusión sobre lo aplicado, superados los argumentos manidos y simplistas que nuestros colegas menos informados suelen utilizar.)

La irrupción de la teoría neoinstitucional en el análisis sociológico de las organizaciones (DiMaggio y Powell, 1983; Powell y DiMaggio, 1991), y su extensión hacia el análisis de las administraciones públicas y la propia evaluación, da pie para redefinir las tradicionales funciones de responsabilidad y mejora, de un modo que resulta más coherente con los planteamientos construccionistas presentados más arriba. El estudio de Thomas Y. Choi y Karen Eboch (1998), sobre la implantación de modelos de calidad, muestra cómo un supuesto uso perverso de la evaluación explica su éxito actual y su imparable extensión en todo tipo de organizaciones, públicas y privadas, asumiendo aquí como sinónimo evaluación, sellos de calidad y procesos de acreditación, que no son sino evaluaciones específicas de las actividades de gestión que cada organización pone en práctica (ver Stufflebeam, 2001, y Reboloso, Fernández-Ramírez y Cantón, 2005). El verdadero objetivo que persiguen estas organizaciones es la legitimación social, ofrecer una imagen pudiente o puntera ante la sociedad, su clientela, los organismos políticos de supervisión o los potenciales socios empresariales. La rendición de cuentas y la mejora, funciones tradicionales de la evaluación, quedan relegadas a un segundo plano –puede que ocurran, pero no es imprescindible– hasta el punto de que no es necesario que los planes de mejora se ejecuten, no importa que su ejecución se dilate en el tiempo hasta perder el valioso sentido de la oportunidad, e incluso que se abandonen definitivamente porque la alta dirección decide un cambio de modelo de evaluación o de agencia acreditadora en función de factores ajenos al proceso de mejora. (Tenemos ejemplos cercanos de estos problemas en el sistema de evaluación de la calidad que utilizan nuestras universidades públicas, tal como mostramos en Reboloso, Fernández-Ramírez y Cantón, 2009).

El deseo de legitimación justificó inicialmente la implantación de un sistema de evaluación o de calidad, y justificará su posterior abandono y recambio si no logra este crucial objetivo. La mejora y la rendición de

cuentas quedan entonces como una suerte de retórica justificatoria al servicio de la introducción de cambios y de la imagen social de la organización⁷.

Entendámoslo bien; no es sencillamente un efecto perverso de un modelo racionalista de gestión que puede mantener sus “verdaderos” objetivos inamovibles e impolutos al margen⁸. Es la consecuencia de introducir una nueva práctica de gestión (evaluación y acreditación) dentro de un sistema social, complejo y abierto, en el que la continuidad y el éxito de cada organización depende de cómo establezca sus relaciones con las restantes organizaciones y grupos de interés relevantes en su sector correspondiente (Powell y DiMaggio, 1991). Se trata de una cuestión ecológica y política, no de epistemología y metodología. El problema epistemológico y metodológico aparece después, y plantea a los teóricos de la evaluación un gran desafío, acorde con la altura de las reflexiones y retos que caracterizan a la ciencia postmoderna: ¿qué teoría de la evaluación –epistemología– debemos proponer y qué prácticas evaluativas –metodología– debemos desarrollar para hacer que la evaluación sea una empresa con significado social, ético y profesional, en este complejo contexto en el que ha venido a implantarse? Mi respuesta, aunque reflexionada, asume importantes riesgos, pero estoy obligado a intentarlo, y le dedicaré los restantes párrafos de este breve ensayo.

La función estratégica de la evaluación

Peter Dahler-Larsen (2007), anterior presidente de la Asociación Europea de Evaluación, argumentaba recientemente sobre la capacidad que las listas de indicadores utilizadas en las evaluaciones tienen para definir o construir parcelas de realidad en el tema o el objeto sobre el que se aplican⁹. Cuando proponemos un criterio de evaluación (por ejemplo, que lo correcto o lo mejor es obtener cierto resultado empresarial, o ser muy citado en los circuitos de las revistas de difusión científica de nuestro área de especialización, o ser empático y negociador con nuestros alumnos o clientes), estamos fijando la atención sobre un aspecto de nuestro comportamiento o de nuestra relación con los demás, convirtiéndolo en lo relevante, en el aspecto principal de nuestro comportamiento. Si además asumimos una actitud competitiva frente a nuestros colegas, o si estamos obligados a demostrar nuestra valía para defender la continuidad de nuestro contrato profesional, negaremos entonces valor a otros posibles comportamientos y resultados, convirtiendo al indicador (o la lista de indicadores) en la única opción válida que merece ser perseguida. Los demás comportamientos no son relevantes, no son válidos, o sencillamente *no son*, pierden su categoría de posibilidad, y quien se atreve a proponerlos se coloca en una posición de desafío ante el status, ante la visión del grupo dominante, y sus argumentos son fácilmente despreciados y reducidos a la categoría de marginalidad, rareza o locura, contrarios a la “evidente realidad” del planteamiento oficial¹⁰.

⁷ Hasta tal punto que a los miembros del personal que participará en los comités de trabajo sin apenas recibir compensación alguna, se les intenta embarcar apelando al argumento del bien común, de la mejora para todos que traerán los sistemas de evaluación. Nuevamente, la universidad nos sirve como ejemplo cercano de estas prácticas retóricas.

⁸ Todo el modelo racionalista está sujeto a crítica porque no tiene capacidad para asumir una realidad ecológica que le trasciende.

⁹ Las expresiones “reificar” y el muy castellano “cosificar”, me parecen acertadas en este contexto.

¹⁰ Peter L. Berger y Thomas Luckmann (1968) señalaban un proceso similar en su teoría de las instituciones, ilustrado en la triada “una-la-nuestra”. Después que se propone “una” opción posible para un suceso (un nombre, un uso, una explicación, un mito), el grupo olvida en generaciones posteriores (a veces, en la misma generación) los

Peter Dahler-Larsen (2007) señala que somos conscientes de que el indicador es una construcción social, y por tanto imperfecto o mejorable, mientras que aceptamos la existencia incuestionable del supuesto fenómeno que tratamos de medir o valorar con el indicador, cuando para muchos de nosotros es evidente que éste también es una construcción social que puede ser criticada, matizada o redefinida por completo. Es decir, no es que exista un objeto concreto y palpable que intenta ser valorado mediante un conjunto de indicadores, más o menos acertados; los indicadores *crean* (definen) el objeto, obviando que la situación podría ser definida y construida de maneras diferentes si ello nos interesara y si nuestro grupo adquiriera la influencia suficiente para imponer una nueva versión del objeto, de nuestra situación o de la propia organización. Por otra parte, aquí es interesante recordar cómo Michael Scriven (1995) defendía el carácter científico de la lógica de la evaluación trazando un paralelismo con el problema de la construcción de conceptos en Psicología. Así, por ejemplo, se acepta que los items que componen una escala son una buena traducción operacional de los supuestos constructos que intentamos analizar con ellos. Igual que aceptamos que la lógica de los constructos teóricos es válida (científica), debemos aceptarlo para los constructos de valor, es decir, las listas de indicadores con las que los evaluadores intentan traducir operacionalmente los criterios de valor escogidos para realizar la evaluación. El problema es que, en ambos casos, los indicadores operacionalizados se convierten en el constructo, del cual no tenemos más referencia que la propia lista de indicadores.

Con un ejemplo. Un modelo de calidad aplicado en una institución educativa (o de cualquier otro tipo) sugiere cuáles son las áreas y prácticas de gestión relevantes en el desempeño organizacional, generalmente relacionadas con el modo en que se toman las decisiones. El modelo propone un conjunto de indicadores útiles para determinar si las prácticas de gestión al uso en la institución coinciden, se ajustan o se alejan del ideal propuesto en el modelo (el criterio de valor es el ajuste; Reboloso, Fernández-Ramírez y Cantón, 2005), y ofrece un procedimiento de evaluación para determinar el ajuste y generar supuestos planes de mejora (en realidad, planes para cambiar en la línea propuesta por el propio modelo; que eso sea o no una mejora, es cuestión aparte). El “modelo” (que no es más que una metáfora, un símil; una propuesta, y no una descripción) se impone como la única posibilidad, apoyada en la presión que realizan los responsables políticos e institucionales para que las evaluaciones se realicen “correctamente” (es decir, como el modelo dice que deben ser hechas) y así obtener el sello, la acreditación y las supuestas recompensas asociadas al mismo. Las personas que participan en los comités de autoevaluación, actuando de buena fe (o no), tenderán a asumir que el modelo es una descripción válida de la realidad organizacional, y a tachar como desviaciones y errores las visiones alternativas; la mera presión grupal dentro del comité acallará las voces discordantes, reclamando el ajuste al procedimiento como una obligación ineludible¹¹.

Llevando el argumento hasta el extremo, lo que pretendemos afirmar son dos cuestiones. Primero, que no existe ningún objeto de conocimiento que pueda ser mantenido más allá de su valor como construcción social útil en determinadas condiciones. Los conceptos científicos (y los criterios de valor, y los conceptos mundanos, por extensión) son constructos, definiciones operacionales que, valiéndose de su potencial como metáforas, intentan fijar una porción de la realidad, posibilitando la discusión técnica sobre el objeto y sus relaciones con otros objetos (elaborados discursos en los que se ponen en relación

orígenes de la opción, para convertirse en “la” única opción, en la verdad. Finalmente, el grupo la asume como “nuestra” forma de hacer o entender las cosas, así que criticarla se convierte en un desafío a la identidad y a la mitología que legitima al grupo.

¹¹ En sí, aceptar el modelo como una descripción de la realidad organizacional no es reprochable. Como veremos más adelante, depende de si ello favorece nuestros intereses y valores personales e institucionales.

unos y otros constructos, todos ellos con una referencia independiente de realidad difusa, en el mejor de los casos). Paul Ricoeur (1995) piensa que la metáfora es útil porque nos permite mirar –denominar, fijar nuestra atención– hacia parcelas de realidad que antes pasaban desapercibidas; en mi opinión, la metáfora permite continuar con el diálogo técnico, pero nada nos garantiza que la supuesta realidad denominada no sea más que una profecía que se cumple a sí misma¹². En cualquier caso, nadie en el pensamiento postmoderno niega el carácter metafórico y narrativo del discurso científico.

Y en segundo lugar, que el éxito de un discurso, de un conjunto de conceptos o indicadores, de un modelo teórico –que, para el caso, son los mismo–, se deriva de la capacidad que tienen algunos de los grupos implicados para imponer su versión de las cosas. Los evaluadores disponen del acertado término inglés de *stakeholders* –aquellos que tienen un interés en juego– para denominar a estos grupos. Y a ninguno nos cabrá duda, supongo, de que cada grupo organizacional o cada grupo social organizado perseguirá de manera primaria que sus intereses se resguarden y ganen como resultado de esta dinámica política organizacional de imponer el modelo de trabajo –lo bueno y lo malo, lo deseable, lo relevante, lo real, los criterios de valor– que la institución habrá de seguir en adelante. En términos más directos: el grupo de interés en el poder impondrá su versión de las cosas (parte del proceso de acceder al poder estriba en esta imposición, puesto que, pensarán los demás, si ellos tienen la razón, es lógico que su opción triunfe; ¡de sentido común!) y propondrá un modelo a seguir, una versión de cómo debe ser la organización en adelante, reforzándose el modelo en la medida en que todos vayan modificando sus campos de actuación para parecerse al mismo¹³. Posteriormente, con una ingenuidad que me sorprende, y a veces incluso resulta sospechosa, nuestros colegas expertos, profesionales o académicos, aceptan y bendicen esta versión interesada de la realidad -que originariamente es sólo una entre otras posibles (Berger y Luckman, 1968)- mediante el proceso de aplicar sobre ella y sus conceptos asociados la lógica de la investigación científica¹⁴. La capacidad crítica de la ciencia es nula en este caso, puesto que, como apunta Andrés Rodríguez (2005), los profesionales asumen como válidas las estrategias y los criterios de valor definidos por el grupo organizacional en el poder, de tal modo que legitiman su versión interesada de cómo son y cómo deberían ser las cosas, contribuyendo al mantenimiento del *statu quo* y haciendo inviable cualquier visión alternativa. La pregunta es, naturalmente, ¿por qué la versión interesada del grupo en el poder, y no la de cualquier otro de los grupos de *stakeholders* en liza?, ¿por qué no una versión negociada?

Volvamos ahora a nuestro problema con los criterios de valor y los indicadores de evaluación. No nos dicen cómo es la realidad del programa o de la organización, sino cómo pensamos que debería ser.

¹² No es mi deseo defender posiciones nihilistas. Los clásicos argumentos del sentido común, no por simples, son menos contundentes. (Thomas Reid: cualquiera que niegue que esta sea mi mano izquierda, está loco; o golpea una pared con la cabeza, tu dolor dirá si es una construcción social. Perdón por las simplezas.) No niego la realidad, sino la posibilidad de un conocimiento técnico-científico representacional, como muchos otros hacen de manera más brillante que yo (Ibáñez, 2001, por ejemplo). Modelos teóricos, conceptos y constructos no son imágenes en espejo de la realidad, sino metáforas que nos ayudan a comprender y tomar decisiones. Son válidas porque resultan útiles (son coherentes dentro de un discurso establecido, o nos ofrecen ideas para emprender acciones provechosas), no porque sean verdades. Rizando el rizo, no sería una posición nihilista, sino una suerte de *solipsismo intersubjetivo utilitarista*. Juegos de palabras, al fin y al cabo, para continuar nuestro diálogo con el mundo.

¹³ El tipo de cambio isomórfico que describen Paul J. DiMaggio y Walter W. Powell (1983), mero mimetismo sin transformación interna, como si las prácticas organizacionales fueran intercambiables de lugar a lugar, de grupo a grupo, sin necesidad de adaptación.

¹⁴ Los modos en que el lenguaje sirve a estos fines de reificación, normalización, consolidación y mantenimiento del poder, son sutiles y numerosos. La obra de Michel Foucault es una referencia obligada en este punto. La crítica feminista no cesa de denunciarlo desde hace años en su parcela de interés.

Como no podemos disociarnos del juego político y del acceso al poder, decidamos conscientemente de parte de quiénes queremos estar¹⁵ o promovamos procesos de negociación para acercar posiciones en una ambiente democrático, tal como sugería Egon G. Guba e Yvonne S. Lincoln (1989) o como se practica en el proceso de creación de un sistema de acreditación (Philips, 2001). Ayudemos a concretar qué tipo de organizaciones queremos, y diseñemos las herramientas para conseguirlas. Nuestra ciencia evaluativa (y toda la ciencia, si es que aceptamos como hipótesis de trabajo los argumentos de párrafos anteriores) es, por tanto, prescriptiva –condiciona el futuro–, y no descriptiva. Asumamos ya, de una vez, la realidad política, simbólica, caótica y socialmente construida de las organizaciones para las que trabajamos, y asumamos que la empresa científica y evaluativa no es pura, sino una organización más como las restantes, siendo críticamente conscientes de las implicaciones políticas y éticas de los modelos que nosotros mismos defendemos.

En otros trabajos (Reboloso, Fernández-Ramírez y Cantón, 2008b, 2009), hemos propuesto el concepto de metacriterio y la metaevaluación como un procedimiento útil para decidir cuáles son los elementos clave que definen hacia dónde queremos orientar la organización y reformar o dar forma –*shaping*– a las dinámicas y estructuras organizacionales para crear sinergias en la línea adecuada. Así, la evaluación, en un planteamiento postmoderno, trasciende las clásicas funciones de responsabilidad y mejora, que quedan como mera retórica legitimista, para dirimirse entre sus nuevas funciones legitimadora y estratégica (DiMaggio y Powell, 1983). No evaluamos para mejorar, ni para conocer, sino para legitimar nuestras decisiones ante los grupos de interés, y para apoyar la puesta en práctica de las estrategias grupales y organizacionales. La evaluación se definiría así como constructorista, narrativa, fortalecedora, política y crítica, con capacidad para apoyar a los grupos organizacionales en sus intereses y para cuestionar las realidades impuestas desde los niveles directivos. De aquí también la utilidad de un concepto de evaluación como asignación del mérito (valoración, juicio), porque desvelar, concretar y negociar los valores e intereses en juego (los criterios de valor, lo deseable, lo relevante) forma parte de la tarea necesaria para democratizar las organizaciones, apoyar intereses múltiples y dar voz a quienes usualmente se limitan a asumir lo que viene impuesto.

En definitiva, mi propuesta no es un nuevo modelo, sino redefinir la evaluación como una función estratégica en apoyo de los intereses declarados por los distintos grupos que conforman la organización. Conocidos, negociados y hechos explícitos sus intereses y deseos de futuro, la evaluación selecciona y define conjuntos de criterios de valor e indicadores que sirven para juzgar o valorar si los cambios que se introducen están en línea con la visión de futuro deseada. Evidentemente, la decisión sobre qué futuro desea o conviene a la organización, depende completamente de sus miembros o de los grupos que la forman, y no del evaluador. Al evaluador, técnico o científico, le corresponde un rol de asesor y facilitador del cambio, y no el rol tecnocrático y paternal del supuesto experto que tiene la clave para dictar lo que le interesa o no a las personas. Gestión democrática por gestión tecnocrática.

La evaluación de cuarta generación (Guba y Lincoln, 1989) mantenía gran parte de sus preocupaciones en intentar defender la validez de su propuesta frente a la opción positivista, es decir, apenas trasciende el planteamiento de la polémica sobre la validez de los métodos cualitativos como opciones válidas para generar conocimientos científicamente aceptables. Asumiendo buena parte de sus postulados, mi

¹⁵ Por ejemplo, Joseph S. Wholey (1986) sugiere que trabajemos para los técnicos de rango medio, ayudándoles con los problemas prácticos de la intervención; Michael Scriven (1994), que seamos sustitutos informados del consumidor; Donna M. Mertens (1995), que seamos portavoces de los sin voz. En muchas organizaciones, la función de evaluación se sitúa en estrecha dependencia de los grupos directivos, apoyando el statu quo con sutiles mecanismos de control, o sencillamente limitando la posibilidad de realizar evaluaciones críticas (Sonnichsen, 2000).

propuesta no se centra en la discusión sobre cómo desarrollar conocimientos válidos (desconfío de todos, con perdón), sino en plantear la evaluación como una herramienta útil para orientar y apoyar el cambio social deseado. Cuestiones como conocimiento, validez, ciencia o verdad, están fuera de lugar aquí, y será un placer discutir las en una ocasión posterior, con ánimo de criticar su valor como elementos o procedimientos retóricos que están siendo utilizados para mantener posiciones de poder cuya legitimidad puede y debe ser cuestionada.

Referencias

- Alvira, Francisco. (1991). *Metodología de la evaluación de programas*. Madrid: CIS.
- Berger, Peter L. y Luckmann, Thomas. (1968). *La construcción social de la realidad*. Buenos Aires: Amorrortu.
- Blanco, Amalio y Valera, Sergi. (2007). Los fundamentos de la intervención social. En A. Blanco y J. Rodríguez Marín (Coords.), *Intervención psicosocial* (pp. 3-44). Madrid: Pearson-Prentice Hall.
- Campbell, Donald T. (1969). Reforms as experiments. *American Psychologist*, 24 (4), 409-429.
- Campbell, Donald T. y Stanley, Julian C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton-Mifflin.
- Chelimsky, Eleanor. (1997). The political environment of evaluation and what it means for the development of the field. En Eleanor Chelimsky y William R. Shadish (Eds.), *Evaluation for the 21st century. A handbook* (pp. 53-71). Thousand Oaks, CA: Sage.
- Chen, Huey-Tsyh. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Choi, Thomas Y. y Eboch, Karen. (1998). The TQM paradox: Relations among TQM practices, plant performance, and customer satisfaction. *Journal of Operations Management*, 17 (1), 59-75.
- Cook, Thomas D. y Campbell, Donald T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally.
- Dahler-Larsen, Peter. (2007). ¿Debemos evaluarlo todo ? O de la estimación de la evaluabilidad a la cultura de la evaluación. *ICE. Revista de Economía*, 836, 93-104.
- DiMaggio, Paul J. y Powell, Walter W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Fernández-Ballesteros, Rocío. (1995). *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Fernández-Ramírez, Baltasar y Reboloso, Enrique. (2006). Evaluation in Spain: Concepts, contexts, and networks. *Journal of MultiDisciplinary Evaluation*, 5, 134-152.
- Guba, Egon G. y Lincoln, Yvonne S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Ibáñez, Tomás. (2001). *Municiones para disidentes. Realidad-Verdad-Política*. Barcelona: Gedisa.
- Jiménez Burillo, Florencio. (1997). *Notas sobre la fragmentación de la razón*. Madrid: Universidad Complutense.
- Katz, Daniel y Kahn, Robert L. (1966) *The social psychology of organizations*, New York, Wiley.
- King, Jean A.; Stevahn, Laurie; Ghore, Gail; y Minemma, Jean. (2001). Toward a Taxonomy of Essential Evaluator Competencies. *American Journal of Evaluation*, 22 (2), 229-247

-
- Lewin, Kurt. (1946). Action research and minority problems. *Journal of Social Issues*, 2(4), 34-46.
- Mertens, Donna M. (1995). Identifying and respecting differences among participants in evaluation studies. En William R. Shadish, Dianna L. Newman, Mary Ann Scheirer y Christopher Wye (Eds.), *Developing the guiding principles. New Directions for Program Evaluation* (Volumen 66, pp. 91-97). San Francisco, CA: Jossey-Bass.
- Mintzberg, Henry; Lampel, Joseph; Quinn, James B.; y Ghoshal, Sumantra. (2003) *The strategy process: Concepts, contexts and cases* (4ª ed.). London, Prentice Hall.
- Munné, Frederic. (1993) La teoría del caos y la psicología social. En Inmaculada Fernández Jiménez de Cisneros y Manuel F. Martínez (Comps.) *Epistemología y procesos psicológicos básicos* (pp. 37-48). Sevilla: Eudema.
- Munné, Frederic. (2007). ¿Qué es la complejidad? *Encuentros en Psicología Social*, 3 (2), 6-17.
- Navarro, José. (2005). La psicología social de las organizaciones desde la perspectiva de la complejidad. Revisión y estado actual de una relación prometedora. *Encuentros en Psicología Social*, 3 (2), 78-87.
- Patton, Michael Q. (1996). A world larger than formative and summative. *Evaluation Practice*, 17 (2), 131-144.
- Pfeffer, Jeffrey. (1998) Understanding organizations: Concepts and controversies. En Daniel T. Gilbert, Susan T. Fiske y Gardner Lindzey (Eds.), *The handbook of social psychology* (pp. 733-777). Boston, MA: McGraw-Hill.
- Philips, Thomas R. (2001). A history of the quality review program of the European Consortium of Innovative Universities (ECIU/IQR). En A. Sursock, *Towards accreditation schemes for higher education in Europe? Final project report*. Geneva, Switzerland: CRE Association of European Universities.
- Powell, Walter W. y DiMaggio, Paul J. (1991). *The new institutionalism in organizational analysis*. University of Chicago Press.
- Ramió, Carles y Ballart, Xavier. (Comps.) (1993). *Lecturas de teoría de la organización*, vol. 1. Madrid: MAP.
- Reboloso, Enrique. (1987). La investigación de evaluación vista a través de los "Evaluation Studies Review Annual". *Revista de Psicología Social*, 3-4 (2), 183-224.
- Reboloso, Enrique; Fernández-Ramírez, Baltasar; y Cantón, Pilar (2003). *Evaluación de programas en el ámbito organizacional*. Granada: Grupo Editorial Universitario.
- Reboloso, Enrique; Fernández-Ramírez, Baltasar; y Cantón, Pilar. (2005). The influence of evaluation on change in management systems in educational institutions. *Evaluation*, 11(4), 465-481.
- Reboloso, Enrique; Fernández-Ramírez, Baltasar; y Cantón, Pilar. (2008a). *Evaluación de programas de intervención social*. Madrid: Síntesis.
- Reboloso, Enrique; Fernández-Ramírez, Baltasar; y Cantón, Pilar. (2008b). Responsibility of Educational Institutions for Strategic Change. *Journal of MultiDisciplinary Evaluation*, 5(10), 5-20. Disponible en: http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/181/197
- Reboloso, Enrique; Fernández-Ramírez, Baltasar; y Cantón, Pilar. (2009). *Crítica y mejora de los sistemas de evaluación universitaria*. Almería: Education & Psychology I+d+i.
- Ricoeur, Paul. (1995). *Autobiografía intelectual*. Buenos Aires: Nueva Visión. 1997.
- Rodríguez, Armando. (1995). La imagen del ser humano en la Psicología social. *Psicothema*, Vol. 5, Suplem.1, pp. 65-79.
- Rodríguez, Andrés. (2005). Psicología de las organizaciones. *Encuentros en Psicología Social*, 3 (1), 157-163.
-

- Rodríguez, Andrés y Ardid, Carmen. (1996). Psicología social y políticas públicas. En José Luis Alvaro, Alicia Garrido y José Ramón Torregrosa (Coords.), *Psicología social aplicada* (pp. 471-474). Madrid: McGraw-Hill.
- Rossi, Peter H. y Freeman, Howard E. (1989). *Evaluación. Un enfoque sistemático para programas sociales*. México: Trillas.
- Rossi, Peter H.; Lipsey, Mark W.; y Freeman, Howard E. (2004). *Evaluation. A systematic approach* (7ª ed.). Thousand Oaks, CA: Sage.
- Scriven, Michael. (1991). Beyond formative and summative evaluation. En G.W. McLaughlin y D.C. Phillips (Eds.), *Evaluation and education: At quarter century* (pp. 19-64). Chicago, IL: University of Chicago Press.
- Scriven, Michael (1994). Product evaluation - The state of the art. *Evaluation Practice*, 15(1), 45-62.
- Scriven, Michael. (1995). The logic of evaluation and evaluation practice. En Deborah M. Fournier (Ed.), *Reasoning in evaluation: Inferential links and leaps. New Directions for Evaluation* (Volumen 68, pp. 49-70). San Francisco, CA: Jossey-Bass.
- Sonnichsen, Richard C. (2000). *High Impact Internal Evaluation*. Thousand Oaks, CA: Sage.
- Stufflebeam, Daniel L. (2001). *Evaluation models. New Directions for Evaluation* (Volumen 89, pp. 7-97). San Francisco, CA: Jossey-Bass.
- Stufflebeam, Daniel L. (2002). CIPP evaluation model checklist. En *Evaluation Checklist Project*. Kalamazoo, MI: The Evaluation Center, Western Michigan University.
- Torregrosa, José Ramón. (1996). Concepciones del aplicar. En José Luis Álvaro, Alicia Garrido y José Ramón Torregrosa (Coords.), *Psicología social aplicada* (pp. 39-56). Madrid: McGraw-Hill.
- Weick, Kart E. (1969). *The social psychology of organizing*. Reading, MA: Addison-Wesley.
- Wholey, Joseph S. (1986). Managing for high performance. The role of evaluation. En David S. Cordray y Mark W. Lipsey (Eds.), *Evaluation Studies Review Annual* (Vol. 11, pp. 194-204). Beverly Hills, Ca.: Sage.

Historia editorial

Recibido: 13/08/2008

Aceptado: 28/04/2009

Formato de citación

Fernández-Ramírez, Baltasar (2009). Construccionismo, postmodernismo y teoría de la evaluación. La función estratégica de la evaluación. *Athenea Digital*, 15, 119-134. Disponible en <http://psicologiasocial.uab.es/athenea/index.php/atheneaDigital/article/view/559>.



Este texto está protegido por una licencia [Creative Commons](#).

Usted es libre de copiar, distribuir y comunicar públicamente la obra bajo las siguientes condiciones:

Reconocimiento: Debe reconocer y citar al autor original.

No comercial. No puede utilizar esta obra para fines comerciales.

Sin obras derivadas. No se puede alterar, transformar, o generar una obra derivada a partir de esta obra.

[Resumen de licencia](#) - [Texto completo de la licencia](#)