

Alkin, Marvin C. (ed.) (2004). *Evaluation toots: tracing theorist' views and influences*, Thousand Oaks/Londres /Nueva Delhi: Sage Publications, Inc., 440 pp (ISBN: 0761928944)

EL ÁRBOL DE LA EVALUACIÓN

SANTOS LÓPEZ LEYVA

El libro reseñado emprende un estudio de la evaluación de la educación en Estados Unidos desde tres perspectivas: los usos que se le han dado a este proceso; los métodos utilizados y la valoración de los programas a través de los ejercicios de evaluación.

El coordinador de los trabajos y editor de la obra, Marvin C. Alkin, es profesor de métodos de investigación social en la División de Graduados en Educación de la Universidad de California, en Los Ángeles, y fundador del Centro de Evaluación de la misma Universidad.

La obra está compuesta de 26 trabajos donde participan, además de Alkin, 24 colaboradores, quienes exponen sus experiencias en torno de la evaluación de la educación y, en general, de programas sociales.

El primer capítulo, "Comparación de puntos de vista sobre la evaluación", delinea de manera concisa los propósitos del libro y define cómo serán abordados las teorías y el intento de agruparlos en tres vertientes denominadas ramas.

El segundo capítulo expone la metodología de trabajo, a través de lo que se designa el "árbol de la evaluación": en una de sus raíces es posible encontrar el control y la rendición de cuentas a la sociedad y en otra se ubica la investigación social. A su vez, el árbol cuenta con tres ramas: la primera, de izquierda a derecha, sostiene el uso de la evaluación; la segunda, considerada la más importante, soporta los métodos, y la tercera, sustenta los instrumentos e información para emitir una valoración del programa.

Santos López Leyva es investigador de la Universidad Autónoma de Sinaloa. Ángel Flores y Riva Palacio s/n, col. Centro, CP 80000 Culiacán, Sinaloa. CE: slleyva@hotmail.com

La primera variable de la raíz es la rendición de cuentas, donde el programa educativo manifiesta la voluntad y disposición de ser evaluado o examinado con la finalidad de atender la responsabilidad contraída con la sociedad, lo que encierra tres orientaciones: *a)* la disponibilidad de reportar lo realizado, es decir, abrirse a una mirada externa; *b)* justificar adecuadamente su quehacer; y *c)* reconocer las deficiencias y someterse a las observaciones y recomendaciones de los evaluadores. La segunda variable es la investigación social. La evaluación tiene como principal sustento un estudio sistemático sobre el comportamiento de grupos, individuos y diferentes fenómenos sociales, realizado por una variedad de métodos. Todos los autores analizados concuerdan que, ante todo, una evaluación es un proceso de investigación social aplicada.

De acuerdo con lo anterior, el objetivo central del libro es estudiar a los diversos teóricos de la evaluación en Estados Unidos, agrupándolos en las ramas señaladas.

Rama de métodos

Al considerar la evaluación como un ejercicio de investigación social aplicada es primordial la existencia de suficiente claridad en la metodología, disponiendo de los modelos más adecuados para cada caso de estudio; por estas razones, Alkin ubica la rama de los métodos en la parte central del árbol.

En esta rama incluye a Ralph Tyler, con su evaluación orientada hacia el cumplimiento de objetivos. Un programa debe definirlos y atenderlos bajo los siguientes pasos: *a)* formulación de una serie de objetivos, que se convierten en guía de acción y de metas; *b)* clasificación de los mismos, atendiendo diferentes variables; *c)* definición y explicación de cada uno de los objetivos en términos de comportamiento del programa; *d)* identificación de situaciones relacionadas con cada uno de ellos; *e)* selección y ejecución de métodos para el logro de los objetivos; y *f)* perfeccionamiento de los métodos para alcanzar mejores resultados.

Shadish denomina a Campbell el “metodólogo de la investigación social”. Durante el siglo XX fueron pocos los artículos que alcanzaran el impacto logrado por el trabajo de Campbell y Stanley, denominado *Experimental and quasi-experimental designs for research*. En ese trabajo destacan tres áreas mayores de investigación social: primera, los autores explican las condiciones necesarias para conducir un verdadero estudio experimental en el marco de variables aleatorias; segunda, son relevantes dos nociones clave,

que están articuladas y explicadas por la validación interna y externa del fenómeno; y la tercera, el experimento no es perfecto y no puede ser usado en todas las situaciones, cada uno tiene sus circunstancias especiales.

Edward Suchman identifica cinco categorías que son necesarias en un proceso de evaluación: *a) esfuerzos*, que toman lugar en la cantidad y calidad con que se llevan a cabo las actividades de evaluación; *b) ejecución*, que se refiere al momento, la forma y los criterios para realizar y medir los resultados del esfuerzo; *c) adecuación de la ejecución*, que es el grado en que ésta se ajusta a las necesidades; *d) eficiencia*, que es medida generalmente en términos de costos monetarios, y *e) procesos*, que se refieren al cómo y por qué trabaja el programa educativo.

Hay que resaltar la preocupación de Thomas Cook por los factores contextuales de la evaluación y los caminos por los que ésta puede afectar los estudios relacionados con la sociedad. Consecuentemente, él escribió que los contratantes deberían trabajar junto con los evaluadores para decidir lo que una evaluación debería examinar, sin olvidar el contexto en que se desarrolla el programa y, por consiguiente, atender los factores que afectan la evaluación.

Robert Boruca es considerado el filósofo de la evaluación; para él, este proceso envuelve la aplicación sistemática de rigurosos diseños aleatorios para medir las implicaciones de los problemas sociales y evaluar su puesta en marcha, eficiencia relativa y el costo-efectividad de la ejecución de los programas.

Hemos sido testigos de la popularidad de Peter Rossi en los programas de educación en evaluación y, en general, en los cursos sobre esta temática que se imparten en las universidades de Estados Unidos. El gran interés de este autor ha sido captar grandes audiencias en el campo de la evaluación, de ahí podemos decir que su labor en la difusión de esta actividad ha sido intensa.

Huey-Tsyh Chen es quizá el teórico más preocupado y más influyente en el desarrollo del concepto y la práctica de la teoría de la “evaluación conducida”, gran parte de su trabajo lo desarrolló en forma coordinada con Rossi.

El gran mérito de Carol Weiss es haber introducido la perspectiva política a la evaluación, señala que se introduce en un programa por tres caminos: *a)* los programas son creados y mantenidos por fuerzas y motivaciones políticas; *b)* altos rangos del gobierno, los cuales toman decisiones de los

programas, se inscriben en organizaciones políticas; y 3) las variadas leyes de evaluación tienen connotaciones políticas.

Sin duda, uno de los gigantes de la metodología en el campo de la evaluación es Lee J. Cronbach. Sus contribuciones metodológicas incluyen el desarrollo del método de las *alfas*, tan usadas en el campo de las ciencias sociales. Siguiendo los trabajos de Weiss, Cronbach rechaza el modelo simplista que asume la decisión de “ir o no ir”. Considera la evaluación como una parte integral de una política de investigación enfocada sobre la elaboración y ejecución de políticas sociales.

Rama de valoración (juicio de valor)

Del tronco de la investigación social crece una rama de evaluadores cuya preocupación es establecer juicios de valor como producto del ejercicio de evaluación. Consideran que el principal atributo de un proceso de este tipo es el juicio que se puede hacer sobre el programa.

Scriven señala que: “lo malo es malo y lo bueno es bueno” y los evaluadores deben decidir cuál es cual. Este autor hace notar que el gran fracaso de algunos de ellos es, simplemente, dedicarse a proveer información a los tomadores de decisiones, pasando la responsabilidad de emitir el juicio definitivo a una persona no profesional. Los evaluadores deben atender todas las etapas del proceso, poniendo atención especial al momento final de establecer juicios sobre el programa.

La aproximación de Elliot Eisner insiste en la necesidad de que el evaluador participe con un sentido crítico que presenta tres aspectos: *a)* hay una descripción crítica donde el evaluador utiliza sus sentidos para puntualizar eventos, reacciones, interacciones y todo lo que alcanza a ver; el evaluador dibuja un retrato de lo que alcanza a percibir en el programa; *b)* la segunda dimensión de la crítica es la interpretación, donde el evaluador trata de entender y dar sentido a todo lo que ha observado; y *c)* hace un juicio final de los datos (u observaciones), incluso emite juicios en torno a las preguntas que se formuló, reconoce que todas las metas, resultados y observaciones son ilimitadas. Insiste en la evaluación por medio de datos cuantitativos, y la decisión se centra en la información cuantitativa que se tiene del programa.

Thomas Owens y Robert Wolf, en forma separada, desarrollaron la categoría de evaluación “adversaria” o, podríamos traducir, “por oposición”. Centran su atención en la manera de juzgar los programas a través de

evaluaciones simultáneas. Diferente de las aproximaciones de Scriven o Eisner, donde la evaluación es hecha por un solo analista o equipo, en la antagonista existen dos conjuntos diferentes de evaluadores atendiendo el mismo programa y buscando los mismos datos, ambos conjuntos asumen posiciones antagónicas.

Sin duda, el teórico que pudo abarcar las diferentes aristas de la evaluación es Robert Stake, por tal motivo es el más difícil de categorizar en una rama del árbol. El autor introdujo una serie de elementos que son esenciales en los actuales procesos de evaluación, entre ellos: *a)* la afirmación de que no existe un valor verdadero para todo, es decir, la relatividad en el comportamiento de los programas; *b)* la necesidad de que la perspectiva de los contratantes de la evaluación deba ser considerada o integrada en la misma, y *c)* la creencia de que los casos de estudio constituyen el mejor método para representar las creencias y valores de los involucrados y para integrar el reporte de evaluación.

Barry MacDonald sitúa al evaluador como un negociador de perspectivas. Describe la evaluación como un proceso que puede ocurrir en tres contextos diferentes: burocrático, autocrático y democrático. En el primero, las valoraciones las hacen quienes están en el poder, en este contexto el evaluador acepta las condiciones y marco que le ponen. En estilo autocrático, el analista no tiene ninguna independencia, se convierte sólo en un consejero, pero no puede modificar las reglas. En cambio, en el modelo democrático, el evaluador considera la opinión de los involucrados, en general de la comunidad evaluada, atiende el principio de derechos “a conocer”, “a participar”, “a ser tomado en cuenta”. Es un modelo de tipo participativo.

En la opinión de Ernest House, la evaluación tiene como propósito proveer de información a los tomadores de decisiones para que puedan llevar a cabo una legítima distribución de los recursos. House menciona la existencia de “falacias éticas” dentro del proceso: el *clientelismo*, *contractualismo*, *administracionismo*, *metodismo*, *elitismo* y *relativismo*.

En mi opinión, el trabajo más importante de Egon Guba e Yvonna Lincoln es haber desarrollado el modelo de cuarta generación de la evaluación; iniciado en 1989, está basado en un paradigma constructivista, que toma lugar en la existencia de una simple realidad donde los individuos construyen sus percepciones sobre ella. La visión constructivista busca compaginar los intereses de todos los involucrados, procura que los grupos afectados en forma negativa sean lo más reducido posible.

Rama de los usos de la evaluación

El grupo de autores ubicados en esta rama insisten en que lo fundamental es el uso de la evaluación, responden a la pregunta ¿evaluación para qué?, qué componentes del programa se pretenden mejorar; en una palabra, lo importante es definir los propósitos de la evaluación, de ello dependen sus características y sus métodos

Daniel Stufflebean, junto con Guba, inicialmente desarrollaron el modelo CIPP, que es una aproximación a la evaluación enfocada en la toma de decisiones. CIPP es el acrónimo de cuatro tipos de evaluación: *contexto*, *insumos*, *procesos* y *productos*. La primera envuelve la identificación de las necesidades de los contratantes para decidir sobre los objetivos del programa y, en general, el marco en que éste se desenvuelve; *insumos* conduce una evaluación de recursos, estrategias y diseños; *proceso*, definición de la operación del programa y, finalmente, la evaluación de *producto* mide los resultados y encamina la evaluación hacia la redefinición del proyecto.

Malcolm Provus especifica cuatro etapas que son requeridas para el desarrollo de la evaluación, más una opcional. Las cuatro primeras son: *a) definición* (intenta especificar las metas, los procesos, recursos, etc.) *b) instalación* (identifica las discrepancias en la operación del programa); *c) proceso* (determina los mecanismos para alcanzar los objetivos); *d) el producto* (lo que se alcanza). La quinta etapa opcional, pero que resulta de suma importancia para la evaluación, es un análisis de costo beneficio.

Joseph Wholey insiste en la organización y sus restricciones, reconoce que obtener información para la evaluación implica costos. De tal modo, propone un proceso de cuatro fases con la idea de disminuir estos costos y menciona etapas secuenciales de información. La primera “valorando la evaluación”, consiste en realizar un análisis inicial que nos indica los términos de viabilidad para conducir la evaluación en términos de organización y los puntos que serán examinados. Las siguientes etapas en el proceso secuencial son: definición de una forma rápida para obtener información, cómo alimentar de datos al proceso de evaluación; la tercera es definir de qué manera medir el comportamiento de los productos, usualmente comparar los resultados esperados con los encontrados; y, por último, la cuarta etapa se refiere a la evaluación intensiva, que usa grupos control o estima la efectividad de las actividades del programa y causas de los resultados observados.

Michael Patton tiene el punto de vista basado en que no es suficiente pensar en la evaluación como información para la toma de decisiones, sino que es necesario que los evaluadores piensen y apoyen a los contratantes en los tipos de decisiones que son posibles tomar con la información disponible. Este punto de vista es que el evaluador sea proactivo.

A pesar de que Marvin Alkin desarrolla un modelo similar al CIPP, reconoce que las evaluaciones de proceso y de producto tienen las dimensiones *sumativa* y *formativa*; un proceso se va reportando mediante la documentación permanente del programa, esto es la dimensión sumativa, mientras la formativa se refiere al mejoramiento del producto. En su modelo más reciente trabaja el punto de vista de evaluación orientada hacia el usuario.

Bradley Cousins instrumenta lo que denomina “Evaluación práctica participativa”, que es definida como una “investigación social aplicada que envuelve a personal entrenado en evaluación y tomadores de decisiones que, basados en sus respectivas prácticas, trabajan en sociedad”.

El trabajo de Hallie Preskill continúa en la rama cuyo foco es el aprendizaje y el desarrollo organizacional. Ella insiste en que durante el proceso de evaluación ocurre un uso sustancial de la investigación y el empleo de un conjunto de herramientas de aprendizaje transformativo.

La teoría de King también se corresponde con la creación de un ambiente de participación a través del proceso de evaluación. Argumenta que para el éxito de evaluación participativa será necesario: *a)* un aceptable poder de la estructura; *b)* compartir significados y experiencias entre los participantes; *c)* la existencia de voluntarios y líderes; *d)* bastante tiempo; *e)* suficiente recursos, y *f)* un alto grado de iteración interpersonal y confianza organizacional.

John M. Owen se centra, primero, en el dominio de la evaluación, el desarrollo y cambio organizacional. Sustenta que la participación de los contratantes es en múltiples etapas, pero considera que las más críticas son la de planeación y negociación. Afirma: “Los evaluadores deben estar preparados para reconocer el marco de interés de los clientes, entonces este conocimiento permite conocer qué tipo de decisiones se pueden tomar”. Describe numerosas técnicas, guías y marcos de trabajo para una planeación y negociación flexible entre contratantes y evaluadores.

El enfoque de David Fetterman es hacia lo que se denomina el “empoderamiento de la evaluación”, describe dos formas generales que tienen

muy poca diferencia: la principal distinción entre ambas es la extensión en que el evaluador participa en el proceso. En el primer caso, éste enseña a los participantes a conducir las evaluaciones de su propio programa, depositando en ellos una mayor seguridad, pero no interviene en el proceso; en el segundo caso sirve como entrenador que ayuda a los otros a conducir sus propias evaluaciones, pero participando.

Experiencias para México

El estudio de los diferentes autores nos deja un conjunto de experiencias para el caso mexicano; en primer lugar, debemos tomar conciencia de que la práctica de la evaluación es la realización de investigaciones sociales acerca de las instituciones y los programas educativos. No es un mero ejercicio desde la perspectiva administrativista, enfocada a contabilizar recursos y productos, sino que constituye un proceso de autorreflexión del programa donde participan todos los actores en la definición de las posibles trayectorias.

La evaluación es un ejercicio que encierra una gran dosis de contenido político a través del cual corre la toma de un conjunto de decisiones que afectan, de distinta manera, a los sujetos involucrados.

Tomando en cuenta todo lo anterior es necesario contar con personas certificadas como evaluadores, por lo que es primordial que las instituciones de educación superior instrumenten programas encaminados a la formación de recursos humanos en esta área.

Asimismo, es prioritario que los académicos participemos en la instrumentación de las políticas de evaluación y con ello evitar que se conviertan en meros ejercicios burocráticos.