

Capítulo 1

Introducción a la Evaluación

Stufflebeam, D. y Schinkfield, A. (1985). **Evaluación Sistemática. Guía teórica y práctica.** Colección Temas de Educación. Barcelona: Paidós.

(Compilación con fines de enseñanza)

INTRODUCCION A LA EVALUACION

La evaluación es una de las actividades fundamentales de los servicios profesionales más solventes. Los clientes de estos profesionales merecen una asistencia que satisfaga sus necesidades, que sea de una alta calidad, moderna y eficiente. Para que los profesionales se responsabilicen de satisfacer estos niveles, la sociedad debe someter regularmente a evaluación a los propios servicios profesionales. Una parte del trabajo evaluativo, la que consiste en la regulación y protección del interés público, debe, obviamente, ser dirigida por organismos independientes, como agencias gubernamentales y juntas acreditadas. Pero, fundamentalmente, las evaluaciones más importantes de los servicios profesionales están realizadas (o encargadas) por los propios profesionales.

Para mantener sus servicios al día y asegurarse de que, efectivamente, satisfacen las necesidades de sus clientes, los profesionales deben recabar continuamente información oportuna sobre el trabajo evaluativo. Este proceso incluye estudiar las necesidades de sus clientes, evaluar los métodos que proponen o utilizan otros, seguir muy de cerca el desarrollo del servicio, valorar los resultados a corto y largo plazo y buscar modos de que los servicios sean cada vez más eficientes y efectivos. La primera pregunta que todo grupo evaluativo exterior debe plantear es si el profesional realiza periódicamente un programa de autoevaluación que responde oportunamente, con las correspondientes pruebas y evidencias, a estas cuestiones.

Desafortunadamente, el alcance de las evaluaciones que realizan los servicios profesionales —por medio de los propios profesionales y sus agentes acreditados— ha sido, tradicionalmente, y en el mejor de los casos, desigual, y, en general, pobre. Algunos distritos escolares grandes y agencias gubernamentales mantienen sistemas de evaluación competentes y realizados por un personal adecuado, y los evaluadores de agencia han sabido ayudar a las orga-

nizaciones centrales a ser responsables con todos los implicados y a obtener la capacidad de planificar y administrar sus servicios. Pero incluso el más competente de estos sistemas evaluativos no puede asumir las responsabilidades evaluativas de todos los profesionales de la agencia. No debe olvidarse el hecho de que la evaluación es una responsabilidad tanto personal como institucional. Las oficinas de evaluación pueden cubrir las principales responsabilidades de una agencia y proporcionar preparación para el servicio y ayuda técnica al personal. Pero, a la larga, de todas maneras, cada profesional debe ser responsable de evaluar su propio trabajo.

Este libro ha sido escrito para ayudar a los educadores profesionales y a aquellos que desempeñen otros servicios profesionales, como servicios sociales o médicos, a consolidar sus conocimientos prácticos sobre la evaluación. Su intención primordial es la de un manual. Se puede consultar cada unidad, que es independiente de las demás, para obtener información sobre un tema en concreto. En general, los temas incluidos han sido escogidos para proporcionar información acerca de la naturaleza general de la evaluación, para facilitar un amplio análisis comparativo de ciertos métodos de evaluación alternativos y para exponer un profundo estudio de ocho modelos o planes alternativos para realizar evaluaciones. Cada unidad ha sido escrita como una guía autosuficiente que incluye objetivos de aprendizaje, información expositiva, preguntas y ejercicios que puede calificar usted mismo y unas pocas preguntas destinadas a la discusión.

Esta primera unidad, concebida como un panorama de la evaluación, proporciona una amplia definición de este concepto y pasa revista al desarrollo de la evaluación, especialmente desde 1930. Muestra cómo ha evolucionado el término *evaluación*, en lo que se refiere a la profundidad y amplitud de los significados que se le han atribuido. Las muchas definiciones y funciones que se le han aplicado a la evaluación, dependiendo de los cambios en el contexto político y social, también son descritas aquí. A la vez, se descubren ciertas novedades metodológicas clave y ciertas funciones que se han asociado con el trabajo evaluativo. Finalmente, se discuten los esfuerzos que se han realizado recientemente para profesionalizar el trabajo evaluativo, especialmente a través del desarrollo de criterios profesionales. En general, la intención de esta unidad es familiarizar al lector con la evaluación entendida como un campo de estudio y una práctica que puede diferenciarse de otros afines, tales como la investigación, la planificación y la administración.

Objetivos

Esta unidad se propone cinco objetivos de aprendizaje. Cuando haya estudiado a fondo el capítulo, usted debería ser capaz de realizar los siguientes objetivos:

1. Definir la evaluación, tal como se presenta en esta unidad, e identificar la base lógica que se ha proporcionado para esta definición.
2. Identificar varios métodos de evaluación alternativos.
3. Identificar los dos principales grupos de normas para programar evaluaciones y caracterizar los principales rasgos de las *Normas* del Joint Committee.
4. Identificar y describir los cinco períodos principales expuestos en esta unidad para caracterizar la evolución de la teoría y la práctica evaluativas.
5. Indicar y definir doce funciones principales del trabajo evaluativo.

Como en todas las unidades, al final de ésta encontrará una serie de preguntas y ejercicios. Le sugerimos que los haga y que compare sus respuestas con las soluciones correspondientes. Luego examine o discuta con sus colegas las preguntas sin respuesta. Aunque pueda alcanzar todos los objetivos estudiando sólo el material que se proporciona en esta unidad, le recomendamos que amplíe sus conocimientos sobre los conceptos estudiados consultando la bibliografía que se facilita al final de cada unidad.

¿Qué es la evaluación?

Se han dado muchas respuestas a la pregunta: ¿qué es la evaluación? Mientras algunas de ellas varían sólo en aspectos menores, otras mantienen diferencias sustanciales. Por ejemplo, algunas definiciones identifican esencialmente la evaluación con otros conceptos como la investigación o las encuestas, mientras otros distinguen claramente la evaluación de otras materias. Una definición importante, y que se viene dando desde hace mucho tiempo, afirma que la evaluación supone comparar objetivos y resultados, mientras que otras exigen una conceptualización más amplia, apelando a un estudio combinado del trabajo en sí y de los valores. Muchas de las distintas definiciones formuladas aparecen en los diferentes modelos de evaluación presentados en las unidades 3-10. Además, la unidad 2 proporciona un análisis comparativo de estos modelos y definiciones. Esencialmente, ese análisis examina todo lo que se ha argumentado contra la definición que damos en esta unidad.

Esta definición es consecuente con lo que dice comúnmente el diccionario y, de una manera más específica, es la que ha adoptado el Joint Committee on Standards for Educational Evaluation. La definición es la siguiente: *la evaluación es el enjuiciamiento sistemático de la valía o el mérito de un objeto.*

Esta definición se centra en el término *valor* e implica que la evaluación siempre supone juicio. Desde el momento en que la evaluación, en un sentido óptimo, emplea procedimientos objetivos para obtener una información segura e imparcial, ya está ligada a la valoración. En efecto, su meta esencial es deter-

minar el valor de algo que está siendo enjuiciado. De acuerdo con esta definición, si un estudio no informa de cuán buena o mala es una cosa, no se trata de una evaluación.

Algunos autores se han mostrado contrarios a esta posición. Creen que esta orientación valorativa conduce a interpretaciones radicales y subjetivas que conceden un poder inmerecido a los evaluadores e impiden el progreso. Afirman, por así decirlo, que «la belleza reside en los ojos del observador» y que las valoraciones de un grupo no deben tener precedencia sobre las de otro a la hora de determinar si algo es bueno o malo. También afirman que la evaluación es, demasiado a menudo, una fuerza negativa y, en general, mala psicología. Pero sus argumentos sólo nos convencen de que la mayoría de las evaluaciones no son determinantes y de que es difícil hacerlas bien; quizás existen muchas evaluaciones distintas, y posiblemente contradictorias en apariencia, de un objeto dado, dependiendo de la posición valorativa que se adopta. Pero, inevitablemente, la valoración debe desempeñar un papel. Y si sirve al progreso y se utiliza para identificar los puntos débiles y los fuertes, y para tender hacia una mejora, puede ser una fuerza positiva. Como puede verse, la evaluación es una empresa compleja.

Creemos que este aspecto de la complejidad no debe ser interpretado en el sentido de que estas evaluaciones sólo deben ser consideradas como procesos mentales que se producen de un modo natural, conduciéndonos así hacia una gran variedad de interpretaciones sobre el valor de algo. Al contrario: significa que hay que poner mucho cuidado no sólo cuando se recopile la información, que debe ser de calidad, sino también cuando se clarifique y proporcione una base lógica que justifique las perspectivas de valoración utilizadas para interpretar los resultados.

Como se puede ver, en la afirmación anterior incluimos la posibilidad de emplear muchas perspectivas de valoración. Consideramos que esta posibilidad es fundamental para los evaluadores y profesionales en general. A menudo, existen distintos públicos, que exigen o necesitan cosas distintas de un servicio profesional. Por ejemplo, a las familias acaudaladas quizá les importe más el hecho de que las escuelas públicas preparen a los estudiantes para entrar en colegios de élite y en universidades, mientras que los nuevos inmigrantes, que son pobres y hablan muy poco inglés, quizá rechacen las escuelas que no ofrecen una educación bilingüe, no atienden a los niños atrasados y no promueven la integración social. A la hora de evaluar un distrito escolar en concreto, un evaluador competente debe tener en cuenta, según nuestra definición, entre otras cosas, estas dos perspectivas de valoración. De hecho, podríamos afirmar que las necesidades que unen o separan a los clientes de un determinado servicio son las que deben considerarse como base para decidir qué tipo de información se debe recopilar y a qué criterios recurrir para determinar la valía o el mérito de un servicio.

Los términos *valía* y *mérito* también deben ser definidos para enriquecer

nuestro concepto de evaluación. SCRIVEN (1981) ya ha indicado las trascendentales diferencias que hay entre ellos y el importante papel que desempeñan a la hora de determinar el valor de algo. Por ejemplo, un programa para formar a profesores de Historia, ¿podrá producir profesores que enseñen Historia a otros de un modo efectivo y de manera que se pueda confiar en ellos? En general, ¿hará bien lo que se supone que debe hacer? Si lo hace, será meritorio, pero puede que no sea válido. Por ejemplo, si la sociedad ya no necesita más profesores de Historia porque ya tiene más de los que puede emplear, entonces aunque el programa sea muy bueno —y quizá sea mejor que otros parecidos—, continuar con el mismo nivel de expansión del programa quizá no sea la mejor manera de gastar sus escasos recursos. Aunque un programa meritorio puede no ser válido, debemos indicar que un programa que no sea meritorio no puede ser válido. Puede que no sea bueno formar malos profesores de Historia aunque exista una gran escasez de ellos.

Hasta aquí, esta discusión sobre qué bases de valoración deben ser tenidas en cuenta a la hora de evaluar algo, ha proporcionado tres importantes fuentes de criterios:

1. Las distintas expectativas de los *clientes*.
2. El *mérito* o excelencias del servicio en cuestión.
3. Hasta qué punto el servicio es *necesario* (es decir, su validez potencial).

En general, recomendamos que se tengan en cuenta los tres aspectos cuando se realice una evaluación. Pero creemos necesario considerar dos aspectos más: la *viabilidad* y la *equidad*. Aunque un servicio pueda ser de mucha calidad, atienda los diferentes objetivos de distintos grupos de clientes y se centre en un área que de verdad lo necesite, quizá falle en lo que se refiere a los valores de la viabilidad y la equidad.

Por ejemplo, puede que consuma más recursos de los que necesita o puede causar infinidad de problemas políticos. En cualquiera de los dos casos, el programa, en principio, debe ser modificado en esas áreas para hacerlo más viable. Obviamente, una buena evaluación del servicio debería mencionar esto y, en su caso, proporcionar una solución para que el servicio sea más sencillo, utilice el tiempo y los recursos de un modo más apropiado y sea políticamente viable. Este argumento en favor de la viabilidad es, en principio, aplicable a todas las sociedades.

La última base valorativa que vamos a mencionar aquí, la *equidad*, está predominantemente relacionada con las sociedades democráticas. Se exigen las mismas oportunidades para todo el mundo y se reclama la libertad para todos. En los Estados Unidos, una evaluación puede resultar incompleta si no tiene en cuenta que un servicio público debe estar pensado y ser asequible para todos los miembros de la sociedad. Es importante indicar que nuestro concepto de equidad es complejo. No es suficiente decir que los servicios pú-

blicos educativos deben poder ser requeridos y usados por todo el mundo. Como ha dicho KELLAGHAN (1982), para que exista verdaderamente la igualdad deben existir siete indicios de equidad:

1. Los servicios públicos educativos de una sociedad deben estar *pensados* para toda la población.
2. Todos los sectores de la sociedad deben tener el mismo tipo de *acceso* a los servicios.
3. Todos los grupos deben disfrutar de una *participación* equitativa en la utilización de los servicios.
4. Los niveles de *logro*, por ejemplo, los años en el sistema educativo, no serán significativamente distintos para los diferentes grupos.
5. Los *conocimientos* exigidos para alcanzar los objetivos del servicio no deben ser sustancialmente diferentes para los distintos grupos.
6. Las *aspiraciones* vitales que se persiguen deben ser similares en todos los grupos sociales.
7. El sistema del servicio deberá producir *resultados* similares en lo que se refiere a la mejora en el bienestar de todos los grupos sociales.

De la discusión anterior, se puede colegir que la definición de evaluación que hemos propuesto es engañosa en su aparente simplicidad. Cuando examinemos seriamente el término base *valor*, debemos inevitablemente tomar en consideración tanto las perspectivas de valoración como la información, la combinación de ambas con el fin de conseguir averiguar el valor de algo y, por último, la transformación de los resultados en conocimientos operativos que se utilizarán para suministrar o no suministrar (o para comprar o no comprar) ciertos servicios y para mejorar los que ya se están ofreciendo.

Antes de terminar con este tema de la definición de evaluación, debemos considerar otros tres aspectos de la evaluación. El primero se refiere a si la evaluación necesita ser comparativa; el segundo, a las principales utilidades de la evaluación; y el tercero se pregunta hasta qué punto los evaluadores necesitan calificaciones especiales.

Algunos autores (Stake, Scriven, Cronbach) han indicado que la evaluación puede centrarse en un solo producto o servicio, o compararlo con otras alternativas válidas. Por otra parte, han debatido los méritos de estas dos posturas, para acabar quedándose con una de ellas. Nosotros creemos que, dependiendo de las circunstancias, una evaluación puede ser, legítimamente, tanto comparativa como no comparativa.

Algo que se debe tener muy en consideración es la naturaleza de la audiencia y qué es lo que necesita saber acerca del objeto de la evaluación. Si la audiencia está compuesta por consumidores que desean escoger un producto o servicio, la evaluación debe ser comparativa y ayudar a los consumidores a saber qué alternativas son convenientes y cómo compararlas críticamente. Si,

por el contrario, la audiencia incluye promotores o consumidores que ya están comprometidos con el desarrollo o el uso de un programa determinado, es mejor que la evaluación se centre intensamente en la evolución del programa para intentar mejorarlo. A veces sucede, sin embargo, incluso si un grupo es totalmente devoto de cierto servicio o producto, que, al considerar otras opciones, obtiene mejores resultados de su proveedor o encuentra una alternativa mejor.

En general, creemos que la evaluación debe ser comparativa antes de que empiece el proceso o antes de suscribirse a un servicio, no comparativa durante el proceso (para que permita concentrarse en lo que resulte mejor para el desarrollo del servicio), y periódicamente comparativa después del proceso y/o la utilización continuada, para que el camino quede abierto ante posibles mayores avances.

Hablando en general, que una evaluación deba ser comparativa o no depende de lo que se quiera conseguir de esa evaluación, en cuanto a su utilización se refiere. Nosotros consideramos tres utilidades principales: perfeccionamiento, recapitulación y ejemplificación. La primera utilización se refiere a la información que se proporciona para asegurar la calidad de un servicio o para mejorarla. Hay que prestar mucha atención a la naturaleza y las necesidades de los consumidores, y a la relación existente entre el proceso y los resultados. En este punto, los evaluadores deberían estar muy en contacto con el personal del programa y servir como guías para las tomas de decisiones. El plan de evaluación necesita ser flexible y receptivo. Cuando la meta principal es mejorar algo, la evaluación debería parecerse más al estudio de un caso que a un experimento comparativo.

La segunda función más importante de la evaluación es hacer recapitulaciones o informes sumarios. Se trata de análisis retrospectivos acerca de proyectos ya terminados, programas ya hechos o productos ya acabados. La información no va dirigida al personal encargado de su desarrollo, sino a los patrocinadores y consumidores. Debe incluir una valoración de la eficacia de lo invertido, siendo también importantes las comparaciones con los más importantes competidores. La información que se deriva de estos profundos estudios y de las encuestas comparativas es muy interesante para la audiencia en este tipo de situaciones.

La tercera utilización es la ejemplificación. Básicamente, la evaluación y la investigación son actividades distintas. La primera se dedica a considerar todos los criterios que se utilizan en la determinación de un valor, mientras que la segunda se limita más bien al estudio de ciertas variables que tienen un interés teórico. Las evaluaciones suelen utilizar métodos subjetivos y no están tan estrechamente controladas ni sujetas a manipulación como las investigaciones. No obstante, los resultados obtenidos, en un cierto período de tiempo, evaluando servicios pueden facilitar información que luego será utilizada para crear y ensayar una teoría. Y, sin duda, los resultados de las evaluaciones a

menudo deberían conducir y conducen, de hecho, a unas conclusiones más concretas e interesantes en el terreno de la investigación. Por lo tanto, creemos que es importante que los evaluadores, al planificar sus estudios, tengan en cuenta que sus conclusiones pueden ser utilizadas en el contexto de una investigación que se proponga una comprensión más amplia de los fenómenos examinados en la evaluación. Con una planificación más minuciosa y cuidadosa, las evaluaciones pueden servir no sólo como guía de programas operativos, o como resumen de sus contribuciones, sino también para ilustrar investigaciones concretas o cuestiones teóricas.

Así pues, parece claro que la evaluación es un concepto altamente complicado. A la hora de planificar o dirigir una evaluación, deben tenerse en cuenta una serie de cuestiones:

- ¿Qué valores deben ser mencionados al juzgar los resultados?
- ¿Qué criterios deben ser adoptados y cuáles deben tener preferencia?
- ¿La evaluación debe ser comparativa?
- ¿La evaluación debe estar capacitada para guiar el desarrollo, responsabilizar a los profesionales, ordenar las opciones o promover nuevas intuiciones acerca de los fenómenos implicados?

Si se tienen en cuenta estas cuestiones con el fin de que la audiencia confíe en los resultados y los utilice, los evaluadores deben mantener ciertos criterios o normas en su trabajo. Como consecuencia, creemos que la distinción entre evaluaciones profesionales y evaluaciones informales es importante. Desde luego, todo el mundo recurre a las segundas cuando se trata de emplear, opinar, aceptar o comprar cosas. Pero ocurre demasiado a menudo que nuestras evaluaciones resultan poco o nada sistemáticas y en modo alguno convincentes, ni siquiera para nosotros mismos. En muchos casos, preferiríamos que nuestras conclusiones tuvieran más relevancia, rigor y objetividad. Estos requisitos, sin embargo, requerirían a menudo más tiempo, dinero y habilidad de los que nosotros tenemos.

Así, en muchas ocasiones, las buenas evaluaciones deben ser realizadas por un personal especializado. Aunque nuestra definición afirma que la evaluación es sistemática, muchas veces, sin duda, no lo es. Lo que intentamos proponer en este libro es una especie de evaluación que sea justificable. La clase de evaluación que tenemos en mente requiere un esfuerzo sistemático por parte de una o más personas que posean, en principio, un cierto nivel mínimo en lo que se refiere a las técnicas evaluativas que se exigen en estos casos. Pensamos que el concienzudo estudio que propone este libro puede ayudar al estudiante a alcanzar el nivel básico de habilidad que se necesita para llevar a buen término nuestro concepto de *evaluación sistemática* y para evaluar las evaluaciones realizadas por otros.

Métodos

Un aspecto que diferencia la evaluación sistemática de la evaluación informal es, sin lugar a dudas, el área metodológica. Cuando pasamos de las evaluaciones realizadas con juicios rápidos e intuitivos a las que suponen una rigurosa recopilación de resultados que deben ser considerados de un modo efectivo, debemos necesariamente enfrentarnos a las complejas áreas de la epistemología, las reglas empíricas, las ciencias informativas, los proyectos de investigación, las valoraciones, las estadísticas, la comunicación, etc. Muchos principios, instrumentos y estrategias de estas áreas son importantes para la evaluación sistemática. El evaluador bien preparado debe dominar convenientemente todas las técnicas de estas áreas y mantenerse informado acerca de los avances tecnológicos que puedan resultarle útiles. Estos evaluadores, que pueden convertirse en pioneros y contribuir al desarrollo de su profesión, suelen aumentar las críticas de los métodos ya existentes y la aparición de otros nuevos.

A lo largo de los años, muchos evaluadores han escogido, e incluso defendido, la utilización exclusiva de unas pocas técnicas. Algunos han equiparado la evaluación a sus métodos favoritos, como el diseño experimental, las pruebas estandarizadas o las visitas personales. Otros han atacado duramente la estrechez de miras respecto de la adecuación de los métodos y, en algunos casos, han abogado por sustituir su técnica favorita, como por ejemplo el estudio de un caso. Nosotros creemos que ambas posiciones son ineficaces, inadecuadas y, a veces, parciales.

A cambio, recomendamos un método ecléctico. Así, creemos que los evaluadores deberían conocer un amplio espectro de técnicas adecuadas y saber cómo aplicarlas en los distintos contextos evaluativos. De este modo, podrían valorar, en cada situación, qué técnicas son potencialmente aplicables y cuáles pueden servir de un modo más eficaz a los propósitos particulares de la evaluación en cuestión.

Entre las áreas técnicas que, según creemos, deben conocer a fondo los evaluadores profesionales se encuentran las siguientes: las entrevistas, los informes preliminares, el análisis de contenidos, la observación, el análisis político, el análisis económico, el examen investigativo, los informes técnicos, el estudio del caso, la evaluación sin metas, la escucha de opiniones contra puestas, las asociaciones de defensa, las listas de control, la elaboración de pruebas, el análisis estadístico, el diseño de la investigación, el análisis sistemático, la teorización y la administración del proyecto. Es recomendable consultar las fuentes de información general que, acerca de estas áreas técnicas, pueden encontrarse en SCRIVEN (1974); ANDERSON, BALL y MURPHY (1974); BRINKERHOFF y otros (1983) y SMITH (1981a y 1981b).

Normas

Una de las características de las prácticas profesionales es que sus miembros deben cumplir exigentes normas referentes al trabajo y al servicio que realizan. Por lo general, estas normas ya vienen formuladas por los miembros de la profesión, y, en algunos casos, por el gobierno. Ejemplos típicos de ello son las normas existentes en las áreas de la ley, la medicina y la contabilidad, con el fin de asesorar y valorar la práctica en cuestión, y, en algunos casos, llegar a conclusiones acerca de ciertos procedimientos ilegales. Hasta muy recientemente, los evaluadores de programas no disponían de normas explícitas para la evaluación, pero esta situación está cambiando.

Recientemente, se han publicado dos conjuntos de normas para las evaluaciones. Uno consiste en un grupo de treinta normas formuladas por el Joint Committee on Standards for Educational Evaluation. El otro es una serie de cincuenta y cinco normas formuladas por la Evaluation Research Society (ERS).

Aquí vamos a centrarnos en el primero por unas razones determinadas. Cuidadosas comparaciones realizadas por investigadores independientes han demostrado que casi todo lo que aparece en las normas de la ERS es examinado con mayor profundidad en las *Normas* del Joint Committee. Es más, la primera serie ha sido elaborada por un comité cuyos diecisiete miembros han sido nombrados por doce organizaciones profesionales que abarcan un total de alrededor de dos millones de miembros, mientras que las normas de la F han sido elaboradas por un grupo de alrededor de dos mil miembros. Otra diferencia clave es que las *Normas* del Joint Committee incluyen tanto al cliente como al que realiza los servicios evaluativos.

El texto *Standards for Evaluations of Educational Programs, Projects, and Material* (Joint Committee, 1981) fue elaborado en un período de cinco años. Este comité está formado por alrededor de doscientas personas relacionadas con la práctica profesional de la evaluación, que realizan la labor de crear, valorar y clarificar principios ampliamente compartidos con el fin de que sirvan como base para evaluaciones sobre programas educativos. El resultado son las treinta normas resumidas en la tabla 1-1. Es particularmente notable el continuo proceso que llevan a cabo reexaminando y revisando las normas. En esencia, el Joint Committee concluye afirmando que las treinta normas son necesarias y suficientes para asegurar que las evaluaciones sean útiles, factibles, precisas y válidas. Se aconseja al lector que estudie en profundidad ese libro de *Normas*.

En general, las *Normas* del Joint Committee aconsejan que los evaluadores y los clientes cooperen entre sí para que las evaluaciones puedan cumplir estas cuatro condiciones principales:

1. Una evaluación debe ser *útil*. Debe estar dirigida a aquellas personas y grupos que estén relacionados con la tarea de realizar aquello que se está evaluando o que sean sus más directos responsables. Debe ayudarles a identificar y examinar lo bueno y lo malo de este objeto. Debe poner muchísimo énfasis en plantearles las cuestiones de mayor importancia.

Debe proporcionarles informes claros de una manera oportuna. Y, en general, debe facilitar no sólo informaciones acerca de virtudes y defectos, sino también soluciones para mejorarlo todo.

2. Debe ser *factible*. Debe emplear procedimientos evaluativos que puedan ser utilizados sin demasiados problemas. Debe tomar en consideración y aplicar controles razonables sobre aquellas fuerzas políticas que puedan, de alguna manera, inmiscuirse en la evolución. Y deben ser dirigidas de un modo eficiente.
3. Debe ser *ética*. Debe estar basada en compromisos explícitos que aseguren la necesaria cooperación, la protección de los derechos de las partes implicadas y la honradez de los resultados. Además, debe proporcionar un informe equitativo que revele todas las virtudes y defectos del objeto.
4. Debe ser *exacta*. Debe describir con claridad el objeto en su evolución y en su contexto. Debe revelar las virtudes y defectos del plan de evaluación, de los procedimientos y de las conclusiones. Debe estar libre de influencias. Y debe proporcionar unas conclusiones válidas y fidedignas.

TABLA 1-1. *Resumen del texto «Standards for Evaluations of Educational Programs, Projects, and Materials»*

A	<i>Normas de utilidad</i>
	Las normas de utilidad están concebidas para asegurar que una evaluación proporcione las informaciones prácticas que necesita una audiencia determinada. Estas normas son:
A1	<i>Descripción de la audiencia</i>
	La audiencia relacionada con la evaluación o afectada por ella debe ser descrita con el fin de que todas sus necesidades puedan ser satisfechas.
A2	<i>Credibilidad del evaluador</i>
	Las personas que dirigen la evaluación deben ser honradas y a la vez competentes para poder llevarla a cabo, con el fin de que sus conclusiones alcancen el máximo de credibilidad y aceptación.

A3 *Alcance y selección de la información*

La información recopilada debe tener el suficiente alcance y haber sido seleccionada de tal manera como para incluir en sí misma las cuestiones pertinentes acerca del objeto de la evaluación y responder a las necesidades e intereses de audiencias concretas.

A4 *Interpretación valorativa*

Las perspectivas, procedimientos y bases lógicas utilizadas para interpretar los resultados deben ser cuidadosamente descritos, con el fin de que queden claras las bases para el juicio valorativo.

A5 *Claridad del informe*

El informe de la evaluación debe describir el objeto que se está evaluando y su contexto, a la vez que los propósitos, procedimientos y resultados de la evaluación, con el fin de que las audiencias comprendan en seguida qué se ha hecho, cómo se ha hecho, qué tipo de información se ha obtenido, qué conclusiones se han sacado y qué recomendaciones se hacen.

A6 *Difusión del informe*

Los resultados de la evaluación deben ser dados a conocer a los clientes y a otras audiencias interesadas o adecuadas, para que puedan valorar y utilizar las conclusiones.

A7 *Oportunidad del informe*

La difusión de los informes debe ser oportuna, con el fin de que las audiencias puedan hacer el mejor uso posible de la información recopilada.

A8 *Impacto de la evaluación*

Las evaluaciones deben ser planificadas y dirigidas de manera que animen a la audiencia a continuar investigando por sí misma.

B *Normas de viabilidad*

Las normas de viabilidad están concebidas para asegurar que una evaluación sea realista, prudente, diplomática y moderada. Son éstas:

B1 *Procedimientos prácticos*

Los procedimientos evaluativos deben ser prácticos, para que los problemas sean mínimos y pueda obtenerse la información necesaria.

B2 *Viabilidad política*

La evaluación debe ser planificada y dirigida conociendo las distintas posturas de los diferentes grupos de interés, con el fin de obtener su cooperación, y para que cualquier posible intento, por parte de alguno de estos grupos, de abreviar las operaciones evaluativas o de influir en los resultados o utilizarlos, pueda ser prevenido o contrarrestado.

B3 *Eficacia de los costes*

La evaluación debe producir información con el valor suficiente como para justificar lo que ha costado.

C *Normas de honradez*

Las normas de honradez están concebidas para asegurar que una evaluación sea realizada legal y éticamente, y, a la vez, con el debido respeto por el bienestar de todos los que intervienen en ella, así como de aquellos que están interesados en las conclusiones.

C1 *Obligación formal*

Las obligaciones de las partes formales de una evaluación (qué hay que hacer, cómo hacerlo, quién debe hacerlo y cuándo) deben ser plasmadas en un compromiso escrito, con el fin de que estas partes estén obligadas a cumplir todas las condiciones del compromiso o a renegociarlo formalmente.

C2 *El conflicto de intereses*

El conflicto de intereses, frecuentemente inevitable, debe ser afrontado abierta y honestamente, con el fin de que no interfiera en el proceso evaluativo y sus resultados.

C3 *Informes completos y claros*

Los informes orales o escritos de una evaluación deben ser francos, directos y honestos en la revelación de las conclusiones pertinentes, incluyendo las limitaciones de la evaluación.

C4 *El derecho del público a saber*

Las partes formales de una evaluación deben respetar y asegurar el derecho del público a saber, dentro de los límites de unos principios y estatutos determinados, como por ejemplo la seguridad del público y el derecho a la intimidad.

C5 *Los derechos humanos*

Las evaluaciones deben ser preparadas y dirigidas de tal manera que sean respetados y protegidos los derechos y el bienestar de los seres humanos.

C6 *Relaciones humanas*

Los evaluadores deben respetar la dignidad y la integridad humanas en su trato con otras personas relacionadas con la evaluación.

C7 *Información equilibrada*

La evaluación debe ser completa y honrada en la presentación de las virtudes y defectos del objeto de la investigación, para que las virtudes puedan ser mejoradas y los problemas solucionados.

C8 *Responsabilidad fiscal*

La cuota del evaluador y los gastos efectuados deben reflejar los procedimientos económicos del trabajo y, por otra parte, ser prudentes y éticamente responsables.

D *Normas de precisión*

Las normas de precisión están concebidas para asegurar que una evaluación revele y transmita una información técnicamente exacta acerca de las características del objeto que está siendo estudiado, para poder determinar su valor o su mérito. Estas normas son:

D1 *Identificación del objeto*

El objeto de la evaluación (programa, proyecto o material) debe ser examinado en profundidad, de manera que su(s) estructura(s) pueda(n) ser claramente identificada(s).

D2 *Análisis del contexto*

El contexto en el que está incluido el programa, proyecto o material debe ser examinado en su más mínimo detalle, con el fin de identificar sus posibles influencias sobre el objeto.

D3 *Descripción de intenciones y procedimientos*

Las intenciones y los procedimientos de la evaluación deben ser concebidos y descritos con todo detalle, de manera que puedan ser identificados y valorados.

D4 Fuentes de información fiables

Las fuentes de información deben ser descritas con todo detalle, para que pueda valorarse si la información es suficiente.

D5 Determinar la validez

Los instrumentos y procedimientos para la recopilación de información deben ser escogidos o creados y luego utilizados de manera que aseguren que la interpretación final es válida para su uso.

D6 Determinar la exactitud

Los instrumentos y procedimientos para la recopilación de información deben ser escogidos o creados y luego utilizados de manera que aseguren que la información obtenida es lo suficientemente exacta para ser utilizada debidamente.

D7 Control sistemático de los datos

Los datos recopilados, sistematizados y finalmente presentados en una evaluación deben ser revisados y corregidos, con el fin de que los resultados de la evaluación no sean inexactos.

D8 Análisis de la información cuantitativa

La información cuantitativa de una evaluación debe ser adecuada y sistemáticamente analizada con el fin de asegurar que las interpretaciones sean fidedignas.

D9 Análisis de la información cualitativa

La información cualitativa de una evaluación debe ser adecuada y sistemáticamente analizada con el fin de asegurar que las interpretaciones sean fidedignas.

D10 Conclusiones justificadas

Las conclusiones a las que se llegue en una evaluación deben ser explícitamente justificadas, para que las audiencias puedan valorarlas.

D11 Objetividad del informe

Los procedimientos de las evaluaciones deben proporcionar las garantías suficientes para proteger los resultados y los informes de la evaluación contra la distorsión de los sentimientos e inclinaciones personales de cualquiera de las partes formales.

Las *Normas* del Joint Committee pueden aplicarse en cualquier etapa del proceso evaluativo: deben ser tenidas en cuenta, por ejemplo, a la hora de decidir si hay que realizar la evaluación o no, a la hora de planear la evaluación y llevarla a cabo, al presentar los resultados y al aplicar las conclusiones. Por lo tanto, la utilización de las *Normas* no sólo sirve para controlar la calidad de un informe ya terminado, sino también como asesoramiento durante la realización de la evaluación. Un segundo punto que hay que tener en cuenta es que las distintas normas son más o menos importantes en cada etapa de la evaluación. Así, quienes utilicen las normas deben estudiarlas y analizarlas comparativamente con el fin de poder aplicarlas juiciosamente en cada una de las etapas de la evaluación. El resumen de las treinta normas presentado en la tabla 1-1 es un punto de partida, pero recomendamos al lector que consiga y estudie el texto completo.

El trabajo del Joint Committee se ha basado en dos premisas básicas: que la evaluación es una actividad humana tan esencial como inevitable, y que la evaluación solvente debe proporcionar una comprensión más amplia y una mejora de la educación. Además, el comité se guió por la creencia de que una serie de normas concienzudamente elaboradas puede desempeñar un papel vital a la hora de mejorar la práctica de la evaluación educativa. En otras palabras: un papel metaevaluativo.

Otro punto trascendental con respecto a las *Normas* del Joint Committee es que se trata de un documento vivo. Un comité permanente, formado por doce organizaciones y respaldado por el Centro de Evaluación de la Universidad de Western Michigan, controla la utilización de las normas y carga con la responsabilidad de ponerlas al día y perfeccionarlas periódicamente. El comité anima a los usuarios para que les proporcionen información basada en críticas y, especialmente, en los usos más actuales de las normas.

La historia¹

Ninguna introducción a la evaluación entendida como una práctica profesional estará completa si no se presta la debida atención a la evolución histórica del tema. Cualquier profesión, para poder satisfacer las necesidades de sus clientes, debe evolucionar de acuerdo con los cambios de las necesidades sociales y teniendo en cuenta todos los avances teóricos y técnicos. A menos que los miembros de una profesión elaboren y sostengan una perspectiva histórica sobre su trabajo, probablemente seguirán teniendo siempre una concepción limitada de su función social, sin poder estimular la innovación ni

¹ Este apartado sobre la historia de la evaluación está basado en un análisis de George Madaus y Daniel Stufflebeam, que aparece resumido en George MADAUS, Michael SCRIVEN y Daniel STUFFLEBEAM, *Evaluation Models*, Boston: Kluwer-Nijhoff, 1983.

contribuir a ella. Se dice a menudo que los que no conocen su propia historia están condenados a repetirla. Nosotros facilitaremos únicamente un breve esbozo histórico con el fin de informar al lector sobre los acontecimientos más significativos ocurridos en el campo de la evaluación.

Nuestro análisis histórico se basa en un trabajo clave de Ralph W. Tyler (véase unidad 3), que a menudo es considerado el padre de la evaluación educativa. Utilizando sus contribuciones iniciales como principal punto de referencia, hemos establecido cinco períodos básicos: 1) el período pre-Tyler, que abarca hasta 1930; 2) la época tyleriana, que va desde 1930 hasta 1945; 3) la época de la «inocencia», desde 1946 hasta 1957; 4) la época del realismo, que cubre el período 1958-1972; y 5) la época del profesionalismo, que incluye la evolución experimentada desde 1973 hasta el presente.

El período pre-tyleriano

La evaluación sistemática no era desconocida antes de 1930, pero tampoco era un movimiento reconocible. El hecho de evaluar individuos y programas apareció, como mínimo, en el año 2000 a.C., cuando algunos oficiales chinos dirigieron unas investigaciones de los servicios civiles, y en el siglo V a.C., cuando Sócrates y otros maestros griegos utilizaron cuestionarios evaluativos como parte de su metodología didáctica. En el siglo XIX, en Inglaterra, se utilizaron comisiones reales para evaluar los servicios públicos. En 1845, en los Estados Unidos, Horace Mann dirigió una evaluación, basada en tests de rendimiento, para saber si las escuelas de Boston educaban bien a sus estudiantes. Entre 1887 y 1898, Joseph Rice estudió los conocimientos en ortografía de 33000 estudiantes de un amplio sector escolar y sacó la conclusión de que la gran insistencia en la enseñanza de la ortografía, que entonces estaba en boga, no había producido avances sustanciales en el aprendizaje. Este estudio está generalmente reconocido como la primera evaluación formal sobre un programa educativo realizado en América. También en los últimos años del siglo pasado, empezó a crearse el movimiento para la acreditación de instituciones educativas y programas en los Estados Unidos; desde entonces, el proceso de acreditación se ha convertido en uno de los principales medios para evaluar los servicios educativos. Una aproximación aún mayor a la evaluación se produjo, en los primeros años de este siglo, con la aparición de los tests estandarizados. Este acontecimiento formaba parte de un gran esfuerzo que pretendía que la educación fuera más eficiente para que un mayor número de estudiantes pudiera utilizar los recursos entonces disponibles en las escuelas.

El breve resumen anterior viene a confirmar que la evaluación sistemática tiene una larga historia. Mientras que sólo muy recientemente se ha llegado a considerar la evaluación como una práctica profesional buena parte del tra-

bajo evaluativo moderno continúa basándose en ideas y técnicas que ya se aplicaban hace mucho tiempo, como por ejemplo las encuestas, las comisiones, la acreditación y las comparaciones experimentales de varios programas.

La época tyleriana

En los primeros años de la década de los treinta, Ralph Tyler acuñó el término *evaluación educacional* y publicó una amplia y renovadora visión del currículo* y la evaluación. En un período de alrededor de quince años, desarrolló sus intuiciones hasta que conformaron un método que suponía una clara alternativa a otras perspectivas.

La principal característica de su método era que se centraba en unos objetivos claramente fijados. De hecho, definía la evaluación como algo que determina si han sido alcanzados ciertos objetivos. Como consecuencia de esta definición, los evaluadores debían ayudar a quienes elaboraban currículos a clarificar el rendimiento de los estudiantes, que había sido precisamente determinado a través de la realización de un currículo. Los objetivos resultantes, referidos al rendimiento, proporcionaban entonces la base para elaborar un currículo y un test de desarrollo. La elaboración del currículo quedaba así influida tanto por los contenidos que debían aprenderse como por la evolución del rendimiento de los estudiantes. La tecnología de los tests de desarrollo iba a servir de base para la elaboración de tests relacionados con los objetivos, y también a los relacionados con las diferencias individuales y con las normas nacionales o del estado.

Durante la década de los treinta, los Estados Unidos, como el resto del mundo, cayó en las profundidades de la Gran Depresión. Las escuelas, como otras instituciones públicas, se estancaron en una ausencia total de recursos y optimismo. Justo cuando Roosevelt intentaba con su New Deal sacar de este abismo a la economía americana, John Dewey y algunos otros intentaron contribuir a que la educación se convirtiera en un sistema dinámico e innovador, que a la vez se fuera renovando a sí mismo. Con el nombre de Educación Progresiva, este movimiento reflejó la filosofía del pragmatismo y empleó los instrumentos de la psicología conductista.

Tyler se vio metido de lleno en este movimiento cuando le encargaron dirigir la parte de investigación del hoy famoso Eight-Year Study (SMITH y TYLER, 1942). Este estudio fue pensado para examinar la efectividad de ciertos currículos renovadores y estrategias didácticas que se estaban empleando en treinta escuelas a lo largo de toda América. El estudio resulta notable por-

* Dado el significado amplio que tiene la palabra «curriculum» en inglés, utilizada en general para designar todo aquello que gira en torno a la actividad docente y no sólo al plan de estudios, hemos preferido conservarla, una vez castellanizada, a lo largo de toda la obra. [T.]

que ayudó a Tyler a difundir, ensayar y demostrar su concepción de la evaluación educacional de una manera inmediata.

A través de este programa de ámbito nacional, Tyler pudo publicar lo que él veía como un método que tenía claras ventajas sobre los demás. Desde el momento en que la evaluación tyleriana incluía comparaciones internas entre los resultados y los objetivos, ya no necesitaba recurrir a costosas y problemáticas comparaciones entre grupos de control y grupos experimentales. El método se centra en determinar el grado de éxito, por lo que se opone a métodos indirectos que determinan factores como la calidad de la enseñanza, el número de libros que hay en la biblioteca, los materiales y la participación de la comunidad. Y además, las evaluaciones tylerianas no se ven demasiado afectadas por la fiabilidad de las diferencias entre las puntuaciones individuales de los alumnos ya que generalmente cubren una franja más amplia de variables exógenas que las cubiertas por los tests tipificados. Todos estos argumentos fueron bien recibidos por el sector de la educación americana, y a mediados de los años 40 Ralph Tyler ya había conseguido la categoría suficiente como para ejercer una gran influencia sobre el panorama educacional en los próximos veinticinco años.

La época de la «inocencia»

Para la sociedad americana, los últimos años 40 y la década de los 50 fue un tiempo para olvidar la guerra, dejar atrás la Depresión, crear y desarrollar aptitudes, adquirir recursos, y lograr disfrutar de la «buena vida». Podríamos haber calificado esta época como el período de expansión, pero hay que tener en cuenta que también existía una amplia complacencia a la hora de enjuiciar los graves problemas de aquella sociedad. Como consecuencia, creemos que es mejor referirnos a esta época como al Período de «inocencia» o incluso de la *irresponsabilidad social*.

Además de ser una época de plenitud, este período fue un tiempo de pobreza y desesperación en las ciudades del interior y en las áreas rurales, pero al parecer casi nadie, excepto los afectados, se dio cuenta. Fue un período de profundos prejuicios raciales y de segregación pero gran parte de la población blanca parecía inconsciente de esta enfermedad. Fue un período de exorbitado consumo y de gran despilfarro de los recursos naturales, sin tener en cuenta que un día estos recursos iban a agotarse. Fue un período de un abrumador desarrollo de la industria y la capacidad militar, pero se tomaron muy pocas precauciones contra los perjuicios que iban a sufrir el Paisaje y las futuras generaciones.

Más que de la evaluación educacional se trató de una expansión de las ofertas educacionales, del personal y de las facilidades. Se construyeron nuevos edificios. Aparecieron nuevos tipos de instituciones educativas, como los co-

legios comunitarios. Los pequeños distritos escolares se unieron a otros con el fin de proporcionar un más amplio abanico de servicios educativos, como ya lo hacían los sistemas escolares más grandes, incluyendo servicios para la salud mental y física, consejos escolares, comedores, enseñanzas musicales, amplios programas deportivos, educación comercial y técnica y educación social. Se incrementaron desmesuradamente las inscripciones en los programas educativos para maestros y, en general, las inscripciones en los colegios aumentaron espectacularmente.

Este panorama general de la sociedad y la educación se reflejó también en la evaluación educacional. Mientras se producía esta gran expansión de la educación, la sociedad no parecía tener un gran interés en formar profesores competentes, en localizar y solucionar las necesidades de los menos privilegiados o en identificar y solucionar los problemas del sistema educativo. Los educadores escribían sobre la evaluación y sobre la recopilación de los datos más importantes, pero, al parecer, no traducían estos esfuerzos en intentos de mejorar los servicios educativos. Esta falta de objetivos provocó también el retraso de los aspectos técnicos de la evaluación. Hubo un considerable desarrollo de los instrumentos y estrategias aplicables a los distintos métodos evaluativos, como los tests, la «experimentación comparativa» y la «coincidencia entre resultados y objetivos». Como consecuencia, los educadores dispusieron de nuevos servicios de tests y de nuevas maneras de calificarlos, de algoritmos para designar objetivos de comportamiento, de taxonomías de objetivos, de nuevos modelos experimentales y de nuevos procedimientos estadísticos para analizar los datos educativos. Pero estas contribuciones no se derivaban de un análisis de la información necesaria para valorar y perfeccionar la educación, ni representaban un desarrollo de la experiencia escolar.

Durante este período, las evaluaciones educacionales dependían, como ya lo habían hecho antes, sobre todo, de los límites de los distritos escolares locales. Las escuelas, pues, dependían del interés y la habilidad locales para poder hacer o no una evaluación. Las agencias federales y estatales aún no estaban demasiado introducidas en el campo de la evaluación de programas. Los fondos destinados a las evaluaciones provenían de las arcas locales, de fundaciones o de organizaciones profesionales. Esta ausencia de ayudas exteriores y de apoyos para la realización de evaluaciones en todos los niveles de la educación iba a concluir con la llegada del próximo período de la historia de la evaluación.

La época del realismo

La época de la inocencia en la evaluación tuvo un brusco final con el auge, a finales de los años 50 y principios de los 60, de las evaluaciones de proyectos de currículos a gran escala, financiadas por estamentos federales. Los educadores se encontraron durante este período, con que ya no podían hacer o no

hacer las evaluaciones a su antojo, y que los futuros progresos de las metodologías evaluativas deberían estar relacionados con conceptos como utilidad y relevancia. Su violento despertar, en este período, marcó el final de una era de complacencia y ayudó a emprender profundos cambios, guiados por el interés público y dependiendo del dinero de los contribuyentes para su financiación, lo cual condujo a que la evaluación se convirtiera en una industria y en una profesión.

Como consecuencia del lanzamiento, por parte de la Unión Soviética, del Sputnik I en 1957, el gobierno federal respondió promulgando la National Defense Education Act de 1958. Entre otras cosas, esta acta proporcionó nuevos programas educativos en matemáticas, ciencias e idiomas, y creó servicios asesores y programas de encuestas en los distritos escolares. Aparecieron, así, un cierto número de nuevos proyectos nacionales para el desarrollo de currículos, especialmente en las áreas de ciencias y matemáticas. Finalmente, se asignaron fondos para evaluar estos programas.

Los cuatro métodos evaluativos expuestos hasta aquí estaban ya representados en las evaluaciones que se realizaron en este período. Primero, el método Tyler fue usado para definir los objetivos de los nuevos currículos y para valorar el grado en que estos objetivos quedaban, más tarde, alcanzados. Segundo, los nuevos tests estandarizados a nivel nacional sirvieron para reflejar mejor los objetivos y los contenidos de los nuevos currículos. Tercero, el método del criterio profesional fue usado para valorar las propuestas y verificar periódicamente los resultados. Finalmente, muchos evaluadores se comprometieron a evaluar los resultados del desarrollo del currículo a través de pruebas concretas.

A principios de los años 60, algunos de los más importantes evaluadores educacionales se dieron cuenta de que su trabajo y sus conclusiones no ayudaban demasiado a los que elaboraban currículos, y no respondían a las preguntas sobre los programas que planteaban aquellos que necesitaban comprobar su grado de efectividad. La «flor y nata» de la comunidad evaluativa educacional participó en estos esfuerzos para evaluar los nuevos currículos; estaban adecuadamente financiados y aplicaban cuidadosamente la tecnología que había ido desarrollándose durante la pasada década o incluso antes. A pesar de todo esto, empezaron a darse cuenta de que sus esfuerzos no tenían éxito.

Esta valoración negativa quedó perfectamente reflejada en un histórico artículo de CRONBACH (1963: véase unidad 5). Examinando los esfuerzos evaluativos del pasado reciente, criticó duramente los conceptos en los que se habían basado las evaluaciones por su falta de relevancia y utilidad, y aconsejó a los evaluadores que abandonaran sus inclinaciones de entonces y se dedicaran a hacer evaluaciones basadas en comparaciones de los resultados obtenidos en tests tipificados, realizados por grupos experimentales y de control. Cronbach recomendó a los evaluadores que reconceptualizaran la evaluación no como si se tratara de una carrera de caballos entre programas compe-

tidores, sino como un proceso consistente en recoger y formalizar información que pueda ayudar a quienes elaboran los currículos. Cronbach argüía que el análisis y la relación de los resultados de un test podrían ser, probablemente, de más utilidad para los profesores que informarles sobre el promedio del total de las calificaciones. Cuando fue publicado por primera vez, el consejo de Cronbach y sus recomendaciones pasaron inadvertidos durante mucho tiempo, exceptuando un pequeño círculo de especialistas en evaluaciones. De cualquier manera, el artículo era básico y contenía hipótesis sobre la conceptualización y la dirección de evaluaciones que iban a considerarse totalmente válidas al cabo de muy pocos años.

En 1965, conducida por la clarividencia del senador Hubert Humphrey, el carismático liderazgo del presidente John Kennedy y la gran habilidad política del presidente Lyndon Johnson, se desencadenó la Guerra contra la Pobreza. Estos programas invirtieron billones de dólares en reformas que pretendían igualar y aumentar las oportunidades para todos los ciudadanos de los Estados Unidos a través de una amplia serie de servicios sanitarios, sociales y educativos. La expansión de la economía permitió que el gobierno federal financiara estos programas, y se consiguió un gran apoyo de toda la nación para desarrollar lo que el presidente Johnson llamó la Gran Sociedad.

Junto a este esfuerzo masivo para ayudar a los necesitados, se produjo, en ciertas regiones, una cierta preocupación por el posible despilfarro de las inversiones si no se exigían algunos requisitos en la contabilidad. En respuesta a estas preocupaciones, el senador Robert Kennedy y algunos de sus colegas del congreso promulgaron el Acta de la Educación Elemental y Secundaria de 1965 que incluía requisitos específicos para la evaluación. Como resultado, el primer artículo de esa acta (que aspiraba a proporcionar una educación compensatoria a los niños menos favorecidos) exigía específicamente que cada distrito escolar recibiera fondos, basándose en ese artículo, para que evaluara anualmente hasta qué punto los proyectos que había realizado al amparo del artículo 1 habían alcanzado sus objetivos, todo ello utilizando los datos de tests debidamente estandarizados. Este requisito, con su referencia específica a las encuestas y a la valoración de la congruencia entre resultados y objetivos, refleja cómo se encontraba la evaluación educacional en aquella época. Y, lo que es más importante, el requisito obligó a los educadores a que trasladaran sus preocupaciones, en lo que se refiere a la evaluación educacional, desde el campo de la teoría y las suposiciones al de la práctica y el perfeccionamiento.

Cuando los distritos escolares empezaron a responder a los requisitos evaluativos del artículo 1, comprobaron rápidamente que los instrumentos y estrategias que empleaban entonces los evaluadores, resultaban totalmente inapropiados para aquel trabajo. Los tests estandarizados que se habían elaborado con el fin de determinar el promedio del grado de conocimientos de los estudiantes no sirvieron para casi nada cuando hubo que diagnosticar las necesi-

dades y valorar los avances de los niños menos favorecidos, cuyo desarrollo educativo se encontraba muy por detrás del de sus compañeros de clase media. Además, dichos tests eran relativamente inadecuados para discernir las diferencias entre escuelas y/o programas, principalmente a causa de sus propiedades psicométricas y el alcance de sus contenidos. En lugar de determinar los resultados directamente relacionados con una escuela o un programa en concreto, estos tests determinaban, en el mejor de los casos, el grado de aprendizaje, centrándose casi en los mismos rasgos que los tests de aptitudes (KELLAGHAN, MADAUS y AIRASIAN, 1980).

La utilización de tests estandarizados planteó otro problema: un método de evaluación como éste se oponía a los preceptos del método tyleriano. Desde el momento en que Tyler reconocía y enfatizaba las diferencias entre los objetivos de un lugar u otro, a este modelo le resultaría difícil adaptar a todo el ámbito nacional sus programas de tests estandarizados. Para ser comercialmente viables, estos programas regularizados debían olvidar ciertos objetivos bastante extendidos e importantes para ciertos lugares, en favor de otros que eran importantes para la mayoría de los distritos. Es más, había una escasez de información acerca de las necesidades y niveles de éxito de los niños menos afortunados, lo cual limitaba a los profesores a la hora de fijar objetivos de comportamiento que fueran significativos para este grupo de alumnos.

Los intentos de aislar los efectos del artículo 1 a través de la utilización de planes establecidos por grupos de control o experimentales fallaron, debido, principalmente a la incapacidad de cumplir los requisitos de estos planes. Y las inspecciones *in situ* realizadas por expertos —aunque fueron muy utilizadas por responsables gubernamentales— no resultaron aceptables como primera estrategia evaluativa porque se consideró que este método estaba falto de la «objetividad» y el «rigor» estipulados en la legislación de la ESEA. Cuando se informaba de que no se había llegado a ninguna conclusión lo cual ocurría frecuentemente se daba muy poca información acerca de los procedimientos empleados y, a menudo, no se proporcionaban datos sobre cómo se habían llevado a cabo. Frecuentemente el énfasis puesto en las puntuaciones de los tests desviaba la atención del examen de los procedimientos o de su ejecución.

Como resultado de la creciente inquietud producida por los esfuerzos realizados en las evaluaciones con resultados siempre negativos, el Phi Delta Kappa creó el National Study Committee on Evaluation (STUFFLEBEAM y otros, 1971). Tras examinar el panorama, este comité concluyó que la evaluación educacional «era víctima de una grave enfermedad» y recomendó el desarrollo de nuevas teorías y métodos de evaluación, así como nuevos programas para preparar a los propios evaluadores. Al mismo tiempo, empezaron a aparecer nuevas conceptualizaciones sobre la evaluación PROVUS (1971), HAMMOND (1967), EISNER (1967) y METFESSEL y MICHAEL (1967) propusieron una reforma del modelo Tyler. GLASER (1963), TYLER (1967) y POPHAM (1971) recomen-

daron los tests basados en criterios como una alternativa a los tests basados en normas. COOK (1966) pidió la utilización del método de análisis sistemático para la evaluación de programas. SCRIVEN (1967), STUFFLEBEAM (1967, 1971) y STAKE (1967) crearon nuevos modelos de evaluación que se alejaban radicalmente de los métodos anteriores. Estas conceptualizaciones reconocían la necesidad de evaluar las metas, examinar las inversiones, y analizar el perfeccionamiento y la prestación de los servicios, así como determinar los resultados que se desean (o no se desean) obtener del programa. También insistieron en la necesidad de enjuiciar el mérito o el valor del objeto de la evaluación. Los últimos años de la década de los 60 y los primeros de la de los 70 vibraron con las descripciones, discusiones y debates acerca de cómo debía ser concebida la evaluación. Las unidades que siguen tratan en profundidad los métodos alternativos que empezaron a tomar forma en este período.

La época del profesionalismo

Aproximadamente en 1973, el campo de la evaluación empezó a cristalizar y a emerger como una profesión diferenciada de las demás, relacionada con (pero también distinta de) sus antecedentes de investigación y control. El campo de la evaluación ha avanzado considerablemente como profesión, aunque es instructivo considerar su desarrollo en el contexto del período anterior.

En aquellos momentos, como ya hemos explicado, los evaluadores se enfrentaban a una crisis de identidad. No sabían muy bien cuál era su papel, si debían ser investigadores, administradores de tests, profesores, organizadores o filósofos. Cualquier calificación especial que se les pudiera dar no resultaría clara. No existían organizaciones profesionales dedicadas a la evaluación entendida como especialidad, ni tampoco publicaciones especializadas a través de las cuales los evaluadores pudieran intercambiar información acerca de su trabajo. Esencialmente, no existía literatura acerca de la evaluación educacional, excepto trabajos inéditos que circulaban únicamente entre círculos de especialistas. Había muy pocas oportunidades de prepararse para ejercer un servicio evaluativo. Las normas claras para una buena práctica se veían con finadas a las encuestas educacionales y psicológicas. El campo de la evaluación era amorfo y fragmentado. Muchas evaluaciones estaban dirigidas por personal sin preparación, otras por investigadores metodológicos que intentaban sin éxito adaptar sus métodos a las evaluaciones educativas (GUBA, 1966). Los estudios evaluativos estaban cargados de confusión, ansiedad y hostilidad. La evaluación educacional como especialidad tenía poca talla y ninguna dimensión política.

Con este telón de fondo, el progreso realizado por los evaluadores educacionales para profesionalizar su especialidad durante los años 70 es totalmente destacable. Algunas publicaciones, como *Educational Evaluation and Policy*

Analysis, Studies in Evaluation CEDR Quarterly, Evaluation Review, New Directions for Program Evaluation, Evaluation and Program Planning y Evaluation News, aparecieron en aquella época, y luego han demostrado ser excelentes vehículos para recoger y difundir información acerca de las diversas facetas de la evaluación educacional. A diferencia de lo que ocurría quince años atrás, numerosos libros y monografías tratan ahora exclusivamente de la evaluación. De hecho, el problema actual no es encontrar literatura sobre la evaluación, sino seguir el ritmo de lo que se publica. El grupo May 12, la Division H de AERA, la Evaluation Network y la Evaluation Research Society han proporcionado excelentes oportunidades para que pueda establecerse un intercambio profesional entre personas relacionadas con la evaluación y otros programas.

Muchas universidades han empezado a ofrecer por lo menos un curso de metodología evaluativa (distinguiéndola de la metodología investigativa) Unas pocas —como la Universidad de Illinois, la Universidad de Stanford, el Boston College, la UCLA, la Universidad de Minnesota y la Universidad de Western Michigan— han elaborado programas para graduarse en evaluación. Durante siete años, la Oficina de Educación de los Estados Unidos patrocinó un programa nacional de preparación para el servicio evaluativo destinado a educadores especiales (BRINKERHOFF y otros, 1983), y otras organizaciones profesionales han ofrecido talleres e institutos que tratan distintos temas relacionados con la evaluación. Se han fundado centros para la investigación y el desarrollo de la evaluación, entre los que se encuentran la unidad de evaluación del Northwest Regional Educational Laboratory, el Center for the Study of Evaluation de la UCLA, el Stanford Evaluation Consortium, el Center for Instructional Research and Curriculum Evaluation de la Universidad de Illinois, el Evaluation Center de la Universidad de Western Michigan y el Center for the Study of Testing, Evaluation, and Educational Policy del Boston College. El estado de Louisiana ha establecido un sistema y un programa para la titulación de evaluadores, Massachusetts está trabajando actualmente en un programa evaluativo similar y Dick JOHNSON (1980) ha elaborado un primer borrador de una guía telefónica de evaluadores y agencias de evaluación.

De un modo creciente, la especialidad está considerando la metaevaluación (SCRIVEN, 1975; STUFFLEBEAM, 1978) como un medio para asegurar y comprobar la calidad de las evaluaciones. Como ya se ha indicado, un comité (Joint Committee, 1981), formado por doce organizaciones profesionales, ha publicado una clarificadora serie de normas para juzgar las evaluaciones de los programas educativos proyectos y materiales y ha establecido un mecanismo (Joint Committee, 1981) para repasar y revisar las *Normas* y facilitar su uso. Además, se han publicado muchos otros grupos de normas importantes para la evaluación educacional (*Evaluation News*, mayo de 1981). Y han aparecido muchas técnicas nuevas, tal y como se describe a lo largo de este libro.

Este considerable desarrollo profesional de la evaluación educacional ha producido diversos resultados. Primero, al producirse un indudable aumento y mejora de la comunicación en la especialidad, ha aparecido también una enorme cantidad de «cháchara» (CRONBACH, 1980). Segundo, aunque la preparación y titulación de evaluadores ha mejorado con el fin de que las instituciones puedan disponer de servicios realizados por personas cualificadas, algunos lamentan que este desarrollo pueda reducirse a una especie de club cerrado y exclusivista (STAKE, 1981). Tercero, la cooperación entre organizaciones profesionales relacionadas con la evaluación educacional, promovida por el Joint Committee on Standards for Educational Evaluation, es un prometedor pero frágil compromiso para fomentar la ejecución y utilización de trabajos evaluativos de alta calidad. Cuarto, mientras la creación de nuevas organizaciones profesionales ha incrementado la comunicación y reducido la fragmentación de la especialidad evaluativa, aún persiste una profunda división entre la División H de AERA, la Evaluation Network y la Evaluation Research Society. Finalmente, aunque se ha incrementado la comunicación entre los partidarios de los métodos positivistas/cuantitativos y los que proponen métodos fenomenológicos/cualitativos, existe, actualmente, el peligro de que se produzca una polarización entre estos dos campos.

A pesar de la creciente búsqueda de métodos adecuados, del incremento de la comunicación y el entendimiento entre los principales teóricos, y del desarrollo de nuevas técnicas, la práctica actual de la evaluación ha cambiado muy poco en la mayoría de los casos. Está claro que es necesario renovar los esfuerzos para educar a los evaluadores en las nuevas técnicas, para poner a prueba e informar de los resultados conseguidos mediante estas nuevas técnicas y para desarrollar técnicas adicionales. En todos estos esfuerzos, debe ponerse el énfasis en conseguir que la metodología responda a las necesidades de los profesionales y de aquellos a quienes sirven.

Hemos presentado la evaluación educacional como una profesión dinámica y aún en desarrollo. Aunque se trata de una profesión aún inmadura, no hay duda de que poco a poco se va convirtiendo en una parte identificable del amplio aparato profesional y gubernamental de educación, salud y bienestar. La predicción que se solía hacer en los años 60, en el sentido de que la evaluación de programas formalizada era sólo una moda que pronto desaparecería, ha demostrado ser falsa, y hay fuertes indicios de que esta especialidad continuará creciendo en importancia, complejidad y valor. Los logros alcanzados desde el final de la década de los sesenta son importantes, pero aún existen muchas deficiencias obvias y son necesarias una investigación y evolución constantes.

Reparto de papeles en el trabajo evaluativo

Hasta aquí hemos presentado y discutido una definición básica de la evaluación, descrito algunos de los métodos relacionados con el trabajo evaluativo, señalado las normas de esta profesión y proporcionado un panorama histórico. Hecho esto, concluiremos echando un vistazo a los papeles que se desempeñan en una evaluación.

Tanto si una evaluación es dirigida por una persona como si lo es por muchas, deben desempeñarse unos cuantos papeles, y cuando se considera el trabajo relacionado con apoyar y promover la evaluación como profesión, este abanico de papeles se convierte en mucho más amplio. Todos los que están relacionados con la evaluación, ya sea en general o como especialistas, deben tener una amplia visión de todos estos papeles con el fin de que su participación y su colaboración se integren en la labor de equipo. Quizá nada sea tan destructivo con respecto a la potencial contribución en una evaluación como el hecho de pensar que esa evaluación es una investigación independiente realizada por un solo investigador. Por el contrario, la evaluación alcanza su mejor momento cuando se desarrolla como un trabajo de equipo.

Con la intención de proporcionar un amplio panorama del trabajo relacionado con las evaluaciones individuales y con la profesión evaluativa en general, hemos descrito doce papeles básicos. Aunque otros autores describan un panorama distinto y le asignen funciones diferentes, nosotros creemos que nuestra clasificación de papeles abarca muchas, si no todas, de las responsabilidades laborales que los evaluadores pueden llevar a cabo. Es más, en nuestra experiencia, esos doce papeles se han asociado con la actual distribución del trabajo. Desde luego, no siempre existe una correspondencia unívoca entre papel y trabajo, pero creemos que esos doce papeles pueden resultar útiles para el lector, no sólo para que reflexione sobre la amplia serie de actividades evaluativas existentes sino también para que tenga en cuenta los posibles estudios que se pueden realizar en torno a la evaluación. Lo que sigue, pues, son unas breves definiciones de los doce papeles básicos que pueden desempeñarse en una evaluación.

El cliente de la evaluación

Cualquier evaluación de utilidad implica la plena participación de un cliente. Esto es cierto tanto si el estudio está dirigido por una persona para servir a sus propios propósitos como si lo está por muchas. El cliente es la persona o grupo que utilizará los resultados para algún propósito, como la selección o el perfeccionamiento de un programa. El grupo de los clientes incluye tanto a quien haya encargado la evaluación en primer lugar, como a

todos aquellos que esperan los resultados para hacer uso de ellos. De una forma u otra, el tiempo y los recursos requeridos para una evaluación son proporcionados por el cliente o por una o más personas pertenecientes a un grupo de clientes. Los clientes son, también, el origen crucial de las cuestiones planteadas y de los criterios empleados para interpretar los resultados; también tienen la responsabilidad primordial de aplicar las conclusiones; y, a menudo, una evaluación no puede realizarse si el cliente no lo dispone todo con mucho cuidado y trata de asegurar de que existirá una plena cooperación entre todos los que intervienen en el trabajo.

Puede encontrarse una gran variedad de clientes en las organizaciones educativas, sociales y de la salud. Esto incluye, por ejemplo, inspectores de distritos escolares, administradores hospitalarios y directores de agencias; directores de colegios, presidentes de departamento y administradores de becas; y maestros, psicólogos y médicos. Si la evaluación consiste en contribuir al máximo a la hora de proporcionar servicios meritorios y válidos a la gente, deberán intervenir y aprovecharse toda esta serie de clientes. Además, creemos que deben estar activamente comprometidos en el encargo y la planificación de las evaluaciones, con el fin de obtener la información que necesitan para hacer bien sus trabajos. Por lo tanto, para satisfacer estas expectativas, deben estar adecuadamente preparados, tanto como todas aquellas personas que van a realizar los otros once papeles. Queremos subrayar que, quizá, la misión más importante de los programas de preparación para una evaluación —tanto para antes del servicio como para su realización— es enseñar a los clientes de la evaluación a cumplir con sus responsabilidades. Si no tienen un concepto sólido de la evaluación, probablemente no intervendrán en ella, se mostrarán pasivos, no sabrán utilizarla para conseguir sus propósitos y/o considerarán que los resultados tienen poco interés. En cambio, los clientes entendidos pueden ayudar a asegurar que la evaluación sea un instrumento útil y poderoso para garantizar la calidad de los servicios profesionales.

El diseñador de la evaluación

Uno de los aspectos más fascinantes del trabajo evaluativo es que se trata de una empresa creativa. Diseñar una evaluación para satisfacer la demanda de información de un cliente es algo más que un proceso mecánicamente rutinario. Entre otras responsabilidades, debe conceptualizar las cuestiones que serán abordadas, examinar el panorama político, trazar una recopilación de datos y un plan de análisis, idear los informes y el proceso informativo, en cargarse del aspecto financiero y del personal y determinar cómo deben ser empleadas las normas profesionales de la evaluación. Para solucionar todo este complejo entramado de decisiones de planificación, no bastan modelos

de otras evaluaciones o planes ya publicados. Por el Contrario, el diseño típico de una evaluación consiste en un continuo proceso conceptual

Para diseñar una buena evaluación, se debe ser un observador sensible de los ambientes sociales, saber entrevistar y escuchar, y estudiar con voluntad y capacidad lo esencial del programa que va a ser examinado. El evaluador debe ser también experto en un amplio campo de técnicas comunicativas y de encuestas, y capaz de reunir las distintas piezas de las premisas y la información en planes de trabajo que sean prácticos coherentes y que sirvan como guías racionales.

Debemos considerar el papel del diseñador de la evaluación como el más importante del trabajo evaluativo, junto con el del cliente. Aquellos que se convierten en expertos en diseño evaluativo, a menudo trabajan todo el tiempo como evaluadores y están muy solicitados como asesores para ayudar a los clientes de una evaluación a planificar y encargar los estudios evaluativos.

Las conceptualizaciones generales sobre la evaluación presentadas en las unidades 3 a 10 proporcionan un importantísimo material de estudio a todos aquellos que deseen adquirir mayor facilidad para conceptualizar planes de evaluación. Además de estudiar las contribuciones creativas de otros, subrayamos que el desarrollo de la capacidad para diseñar evaluaciones requiere práctica; se necesita haber participado en la planificación de muchas evaluaciones. Creemos que esta práctica debe incluir algún trabajo individual con un cliente. Además, se puede aprender mucho colaborando con grupos encargados del diseño de evaluaciones. Igualmente, el estudio de un gran número de diseños evaluativos preparados por otros proporciona muchos buenos ejemplos e intuiciones acerca de esta compleja área de trabajo. Creemos que estos programas de preparación para evaluaciones deben incluir ejercicios y otras formas de guías prácticas para la planificación de evaluaciones.

El coordinador de la evaluación

Muchas evaluaciones requieren los esfuerzos de varios participantes, y la utilización de considerables recursos y servicios para múltiples clientes. Por lo tanto, se hace necesario coordinar de un modo eficaz todo este trabajo. Esta necesidad se amplía cuando una determinada agencia tiene en preparación Varios proyectos de evaluación distintos. No es sorprendente, por lo tanto, que la profesión incluya muchos trabajos evaluativos con nombres como ayudante del supervisor de la evaluación, coordinador de proyectos evaluativos director del departamento de evaluación o director de personal del proyecto de evaluación.

Este papel implica que las personas especializadas en evaluaciones deben también prepararse para funciones administrativa y también que aquellos a quien se les asigne un trabajo administrativo relacionado con la evaluación

deben poseer o adquirir conocimientos básicos sobre la planificación de evaluaciones. Finalmente, creemos que las más importantes instituciones educativas, sanitarias y sociales, deben a menudo dirigir sus esfuerzos hacia la creación de grupos administrativos competentes.

Asistentes de la evaluación

Frecuentemente, uno de los papeles más importantes —a menudo independientes— de la evaluación es el del asistente. Este papel supone contactos periódicos con los clientes, revisiones de los documentos del programa, inspecciones de las actividades del programa, hacer entrevistas y recoger otro tipo de informaciones acerca de las partes, redactar informes y asumir la principal responsabilidad a la hora de presentar los informes. A veces, el asistente es el encargado de realizar la planificación del estudio completo. En otros casos, él o ella tiene una responsabilidad menor. Pero siempre estará encargado de colaborar con el personal del programa y de cooperar en la realización de la planificación del estudio.

Para desempeñar el papel de asistente se requieren tanto habilidad técnica como capacidad para las relaciones humanas. Además, las personas asignadas a este puesto deben poseer o ser capaces de desarrollar una gran credibilidad ante los clientes y ante todos aquellos que participan en la realización del programa.

Respondentes de la evaluación

Se trata de un papel al que demasiado a menudo no se le concede mucha importancia. Son las personas que rellenan los *formularios*, responden a los tests, contestan a las entrevistas y permiten que les observen en su trabajo. En muchos casos, la calidad de la información obtenida depende casi por completo de su buena voluntad para cooperar y dar lo mejor de sí mismos. Es improbable que realicen cualesquiera de estas tareas si no están informados acerca de su papel y convencidos de la importancia del estudio. Como consecuencia, quienes están al frente de las evaluaciones deben hacer un esfuerzo especial para asegurar que el estudio servirá de algo a estos colaboradores, tal y como ocurre con los clientes. También deben cuidar que las actividades que promuevan les motiven y preparen para responder apropiadamente.

Especialistas en ayuda técnica

La mayor parte de las evaluaciones suponen una cierta cantidad de trabajo técnico que requiere la ayuda de expertos especializados. Por ejemplo, puede

ser necesario supervisar y anotar los resultados de unas entrevistas; crear, proporcionar y utilizar instrumentos para una recopilación de datos estructurada; analizar tanto la información cuantitativa como la cualitativa; y realizar presentaciones audiovisuales de los resultados.

En las evaluaciones a pequeña escala, un simple evaluador puede, a menudo, hacer todas estas tareas sin asistencia especializada. El agente evaluador está obligado a coordinar y realizar todos los trabajos que sean necesarios, así como a dominar todas las técnicas especializadas que se requieran. Este tipo de agente necesita una amplia preparación que le introduzca en las más importantes de estas especialidades. A menudo, esta persona debe ser capaz de coordinar la asistencia de los especialistas cualificados,

De todas maneras es inevitable que los estudios a gran escala contraten especialistas para llevar a cabo estas tareas con el fin de que los asistentes puedan concentrar sus esfuerzos en su propio campo de trabajo: los clientes y los respondentes. Este grupo de expertos puede incluir especialistas en la elaboración de pruebas, especialistas en muestreos, especialistas en el procesamiento de datos, estadísticos, especialistas en el estudio de casos y redactores técnicos.

Especialistas en información

Un papel técnico que hemos escogido para dedicarle una atención especial es el del especialista en información. Y lo hemos hecho porque ese papel supone una continua necesidad de información relacionada con otras evaluaciones y no con problemas técnicos determinados. El especialista en información es especialmente importante en las agencias que disponen de departamentos de evaluación. Las evaluaciones dirigidas desde estos departamentos deben tomar en consideración el panorama institucional y su contexto histórico, examinar los descubrimientos del pasado y de la actualidad para determinar las tendencias y evitar una innecesaria pluralidad de esfuerzos en lo que se refiere a la recopilación de datos. Es necesario, como mínimo, un buen sistema para ordenar y archivar los informes de manera que la información que contienen pueda ser fácilmente consultada en futuras evaluaciones. Además, muchas agencias se beneficiarían de una recopilación regular de información que las ayudaría a estar totalmente al corriente de entradas de datos, procesos y producciones, así como de la actitud de sus componentes hacia los servicios de la agencia.

El desempeño del papel de especialista en información requiere que el evaluador, o uno o más miembros del equipo de evaluación, conozca bien las ciencias informativas. Eso supone conocimientos y técnicas en áreas como la Comunicación, los sistemas de información, el proceso de datos, la contabilidad financiera y la planificación de gran alcance. Las personas que piensan espe-

cializarse en evaluación deben adquirir un dominio general de las especialidades de la información, mientras que un cierto tipo de trabajo correspondiente a esta área resulta importante incluso para las evaluaciones más sencillas. Además, los departamentos de evaluación deben, a menudo, asignar una o más personas para el cumplimiento de esta importante función.

Especialistas en comunicación

La evaluación bien hecha es una comunicación efectiva acerca del mérito y el valor de algún objeto. Esto quiere decir que los problemas de los clientes guían el trabajo de los evaluadores, y las conclusiones de éstos llegan hasta los clientes para informarlos. Ninguno de estos objetivos se hace posible sin una comunicación efectiva entre el evaluador y el cliente.

Ocurre demasiado a menudo, de todas formas, que la comunicación efectiva se pierde en el trabajo evaluativo. Los evaluadores y/o los clientes asumen que los primeros conocen su oficio y, de algún modo, obtendrán y facilitarán las conclusiones adecuadas; en lugar de eso, los evaluadores a menudo se ocupan de las cuestiones equivocadas. Por otra parte, los evaluadores deben hacer un buen trabajo consiguiendo que los clientes ayuden a centrar la evaluación y a recopilar una información que resulte sólida y potencialmente útil, para luego «pasarles la pelota» y que sean ellos los que, sin ayuda, interpreten y apliquen la información. Desearíamos dejar totalmente claro que la evaluación efectiva supone no sólo recopilar y analizar la información apropiada, sino también asegurar, a través de una efectiva tecnología comunicativa, que la información sea presentada de una forma útil y utilizada en consecuencia. Sabemos que los evaluadores no deben ser considerados responsables de todas las falsas interpretaciones y empleos erróneos que se han hecho de sus informes, pero sí, entre otras cosas, de asegurarse de que sus informes sean relevantes para los intereses de sus audiencias y de que puedan entenderlos.

Todo esto requiere que los evaluadores tengan conocimientos sobre tecnologías de la comunicación. Básicamente, deben ser capaces de preparar informes escritos con claridad y que estén dirigidos a —y resulten comprensibles para— la audiencia o audiencias que los han solicitado. También deben saber cómo comunicar conclusiones a través de medios distintos del material impreso: por ejemplo, deben saber cómo hablar en público, utilizar técnicas audiovisuales y servirse de los medios públicos. Finalmente, queremos señalar que deben poseer conocimientos sobre el análisis político y la solución de conflictos, ya que, a menudo, las evaluaciones son objeto de controversia.

El papel del especialista en comunicación tiene considerables implicaciones para la preparación de los evaluadores, su utilización de los especialistas y la elección de personal en los departamentos de evaluación. Sin duda, los evaluadores deben adquirir preparación en una amplia serie de técnicas comu-

nicativas. En su práctica, deben, como mínimo, someter los borradores de sus informes al juicio de los editores. Y los departamentos de evaluación, a menudo, deben emplear a especialistas en comunicación para que asesoren la planificación de las evaluaciones y la presentación de las conclusiones. Estos especialistas también desempeñan un importante papel a la hora de evaluar y ayudar a mejorar el exhaustivo trabajo de un departamento de evaluación, comunicando los resultados a sus clientes.

El análisis de los papeles realizado hasta este punto ayuda a clarificar la complejidad del trabajo evaluativo. Una gran agencia, que deba someter continuamente a evaluaciones a sus exhaustivos servicios, puede que necesite un personal completo de especialistas en evaluación. La figura 1-1 es un ejemplo de cómo debe estar organizado el departamento de evaluación de una agencia para que cumpla a la perfección las diversas funciones de una evaluación. Como puede verse en ese esquema organizativo el coordinador de la evaluación tiene la responsabilidad de satisfacer las necesidades evaluativas de la agencia y de sus componentes. Para ayudarle en el desempeño de este papel, el coordinador de la evaluación dispone de los esfuerzos de distintos especialistas. Que los servicios del departamento sean técnicamente sólidos e influyentes, depende de hasta qué punto su trabajo esté coordinado de un modo efectivo. Al mismo tiempo, queremos subrayar que la evaluación debe, a menudo, estar realizada por un solo evaluador. Para hacer bien su trabajo, debe llegar a desenvolverse bien en toda la amplia gama de papeles ya descritos. Probablemente, la mejor manera de encontrar y aprender el modo de conocer bien toda la gama de papeles de una evaluación es dirigiendo evaluaciones de una sola persona bajo la tutela de un experto en evaluaciones. Participar en equipos de evaluación también proporciona excelentes oportunidades de aprender. Los evaluadores deben ser inventivos a la hora de utilizar la ayuda de los especialistas, que pueden resultarles baratos e incluso no costarles nada. Hasta el Llanero Solitario consideraba necesario y beneficioso colaborar con un socio leal y capaz.

Llegados a este punto en el análisis de los papeles en una evaluación, sólo hemos examinado aquellos que corresponden a una evaluación efectiva. Ahora vamos a ver otros cuatro que actúan como papeles secundarios en la profesionalización de la evaluación.

Preparador de la evaluación

A lo largo del análisis precedente, nos hemos ido refiriendo a la necesidad de una buena preparación por parte de aquellos que participan en una evaluación. Estas necesidades incluyen tanto la orientación hacia unos planes de evaluación específicos como el desarrollo de unas técnicas y conocimientos más generalizados y, más concretamente los conocimientos que se adquieren



Figura1-1. Esquema organizativa

examinando una práctica determinada. Además, la necesidad de una preparación corresponde tanto a los especialistas como a los clientes y a los respondentes.

La evaluación, como cualquier otra profesión, debe asegurarse de que los que participan en ella disfruten de una amplia gama de sólidas oportunidades para prepararse bien. Esto incluye programas de preparación especializada que culminan en una titulación; cursos auxiliares que están concebidos para asesorar a todos aquellos que se estén preparando para una serie de funciones profesionales —como por ejemplo supervisores, asistentes sociales, profesores y médicos— sobre la evaluación de sus propios servicios profesionales; continuas oportunidades educativas para evaluadores y especialistas, y sesiones orientativas sobre estudios específicamente evaluativos destinados a clientes y respondentes. Este tipo de preparación debe presentar un contenido actual que resulte válido para los problemas concretos que deberán enfrentar los aprendices. Sin duda, la preparación es un papel crucial para mantener y hacer progresar la calidad de los servicios evaluativos.

Investigador evaluativo

La misión de cualquier profesión es servir a sus clientes de la mejor manera posible, y la realización de este imperativo se ve siempre limitada por el estado actual de los conocimientos acerca de la especialidad. Por lo tanto, todas las profesiones deben tener un carácter investigativo. Los profesionales deben

estudiar las necesidades relacionadas con la profesión y los problemas de los clientes, examinar sus prácticas a la luz de los conocimientos y principios de las disciplinas más relevantes, analizar su trabajo en su contexto histórico, teorizar sobre los servicios profesionales y hacer avanzar los conocimientos sobre su especialidad a través de un análisis estructurado de sus principales premisas e hipótesis.

Los programas de investigación más extensos se dan en las profesiones más veteranas, como la medicina y el derecho, o los negocios y la industria. En el primer caso, la investigación, bajo la influencia de la ética profesional, se utiliza para rendir el mejor servicio posible. En el segundo, la principal motivación es la obtención de beneficios y la supervivencia en un mundo competitivo. Ambas áreas son muy conscientes de que un mejor servicio y la competitividad dependen casi totalmente de un programa de investigación dinámico y adecuado.

Aunque la evaluación es tan antigua como la civilización, sólo desde hace muy poco tiempo existe como práctica formal. Únicamente en las dos últimas décadas se han realizado serios esfuerzos para recopilar la historia de la especialidad, teorizar sobre la práctica evaluativa y planificar estudios que resulten adecuadamente descriptivos y se basen en hipótesis apropiadas. Pero existe, y debe existir, una fuerte tendencia hacia el establecimiento y mantenimiento de una base investigativa para la especialidad. Esta tendencia es especialmente evidente en la gran cantidad de literatura aparecida sobre la especialidad evaluativa, aunque gran parte de este material no puede calificarse como investigativo. En cualquier caso, queremos subrayar que la investigación evaluativa es una función vital para asegurar que los evaluadores mejoren progresivamente sus servicios y sus resultados.

Promotor de la evaluación

Se trata de otro de los papeles vitales de la evaluación, Tiene como base la función de ayudar colectivamente a los evaluadores a alcanzar y mantener el *status* de una profesión. El significado de este papel puede encontrarse en la definición de un diccionario que afirma que una profesión es «un oficio que requiere un conocimiento especializado y, a menudo, una preparación larga e intensiva, incluyendo el dominio de técnicas y métodos, así como también de las nociones científicas históricas y eruditas en las que se basan estas técnicas y métodos, manteniendo, mediante la organización o la coordinación de conceptos, altos niveles de ejecución y éxito, y sometiendo a sus miembros a un estudio continuo y a una concepción del trabajo en la que lo que más importa es el ofrecimiento de un servicio público».

La especialidad de la evaluación está empezando ya a presentar las características de una profesión tal y como se promulgan en esta definición. Ya

existen programas de preparación que terminan en doctorados y *masters*, así como también una extensa y boyante literatura sobre el tema. La literatura sobre evaluación recomienda contenidos y experiencias para ser incluidas en los programas de preparación, así como análisis sobre cómo deben elaborarse los programas a partir de disciplinas afines, como la filosofía, la psicología, la sociología y la economía. Muchas organizaciones profesionales y dos amplios conjuntos de normas profesionales ya están a disposición de los evaluadores. Y, desde luego, el gobierno, el público y las organizaciones de servicios continúan requiriendo evaluaciones de los servicios profesionales, así como fondos que los financien. Mantener estos avances hacia la profesionalización de la evaluación requiere iniciativas nacionales y estatales, además del apoyo de los miembros de la profesión. Cada una de las personas que cobren por trabajar como evaluadores están obligadas a apoyar los distintos esfuerzos que se están haciendo para que la profesión siga adelante. Entre las funciones específicas que pueden cumplirse se encuentran las de editor de una publicación sobre la evaluación, elaborador de normas, educador de evaluadores y especialistas y funcionario de una sociedad evaluadora. Compartiendo las responsabilidades de estas funciones, los evaluadores pueden contribuir a asegurar al público que los servicios profesionales que ellos ofrecen son modernos y de gran calidad.

El metaevaluador

El último papel que queremos subrayar, el del metaevaluador, requiere mucha astucia y penetración. Está relacionado con las funciones de evaluar evaluaciones, pero también supone enjuiciar el valor y el mérito de todo lo que representa y hace la profesión: servicios de evaluación, utilización de las evaluaciones, preparación de la evaluación, investigación evaluativa y desarrollo organizativo. La metaevaluación utiliza las normas aceptadas de la profesión y trata de asegurar que son cumplidas. También se ocupa de examinar las propias normas con el fin de decidir en qué sentido deben ser revisadas.

De una manera u otra, la metaevaluación necesita incorporarse a cada una de las actividades de la profesión. A veces se trata de una actividad de autoasesoramiento, como cuando los evaluadores, los preparadores, los investigadores o los funcionarios de sociedades evaluativas examinan sus planes, la ejecución de su trabajo o los resultados, comparándolos con las normas de la profesión. En otros casos supone contratar a un agente independiente para que valore los servicios. El objetivo de la metaevaluación es asegurar la calidad de los servicios evaluativos, evitar o enfrentarse a las prácticas ilegales o a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la empresa evaluativa. Desde luego, todos los evaluadores profesionales deben mantener al día sus conocimientos de las normas de la profesión, incorporarlas a su tra-

bajo, estar dispuestos a actuar como metaevaluadores de los trabajos de otros profesionales y actuar profesionalmente a la hora de evaluar e intentar mejorar las contribuciones que la evaluación hace a la sociedad.

Con este breve comentario sobre la metaevaluación, terminamos nuestro repaso de los papeles que se desempeñan en una evaluación. Hemos analizado minuciosamente, hasta cierto punto, la función del evaluador profesional y hemos echado un vistazo a las divisiones que presenta. Creemos que, sin duda, todo ello ha contribuido a apoyar la idea de que la evaluación es a menudo un trabajo de equipo. Pero sacar la conclusión de que las evaluaciones necesitan de los esfuerzos de un equipo sería una falsa interpretación: por el contrario, la evaluación debe ser una parte necesariamente integrante de las funciones de todo profesional. Nuestra conclusión es que estas funciones (o papel) son tan difíciles como importantes. Y la conclusión más obvia y fundamental es que todos los profesionales deben desarrollar sus responsabilidades evaluativas, puedan o no conseguir la ayuda de un equipo y aunque tengan que desempeñar ellos solos todas las funciones de la evaluación.

Ya daríamos un gran paso en la dirección adecuada si consiguiéramos una mayor facilidad para conceptualizar los problemas evaluativos y los métodos adecuados para solucionarlos. El resto de este libro ha sido pensado para llevar a cabo este proceso. Presenta y examina en profundidad varias conceptualizaciones alternativas de la evaluación.

Test de conocimientos sobre la unidad 1

Puede comprobar que cada unidad finaliza con tres tipos distintos de cuestiones o ejercicios; cada uno de ellos, a su manera, pretende saber cuál es su grado de comprensión de la unidad. Se encontrará con preguntas que presentan múltiples respuestas para que escoja una, y con explicaciones que aclaran cuáles de las respuestas alternativas son correctas y cuáles no. Tanto la pregunta como las respuestas deben permitirle saber si usted ha alcanzado un nivel satisfactorio en relación con los objetivos que se proponía la unidad en lo que a conocimientos se refiere. En cada caso, rodee con un círculo la letra que, según usted, representa la mejor respuesta y luego valore su elección comparándola con la solución que le proporcionamos y con la subsiguiente Interpretación de las respuestas alternativas.

Cuando haya respondido a todas las preguntas sobre conocimientos, se le pedirá que revise su trabajo y aumente su comprensión de la unidad, o bien que continúe los otros ejercicios. Entonces deberá enfrentarse a algu-

nos ejercicios prácticos. La unidad concluye con unas «Preguntas sin respuesta» y con la recomendación de unas lecturas complementarias.

PREGUNTA 1. ¿Cuál de las respuestas que vienen a continuación refleja lo que esta unidad ha considerado como los fundamentos esenciales que deben ser considerados para determinar el valor de un objeto que se está evaluando?

- a) Aunque otros criterios también son importantes, los esenciales son el mérito, la necesidad y la equidad.
- b) Deben ser considerados todos los valores más apreciados por el personal de la evaluación y sus clientes.
- c) Los evaluadores son los que deben escoger los fundamentos, basándose en su propia filosofía.
- d) Las expectativas de los clientes, el mérito, la necesidad, la viabilidad y la equidad son los criterios que deben determinar esos fundamentos.

Respuesta correcta: Debe usted haber señalado la «d». Las distintas expectativas de los *clientes*, el *mérito* del objeto (o servicio) en cuestión, la *necesidad* que se tenga de él, la *viabilidad* (incluso política) del objeto y la *equidad* o igualdad de oportunidades para que todos puedan tener acceso al objeto, deben ser las normas por que se debe guiar el evaluador.

Respuestas incorrectas

- a) Es incorrecta porque es incompleta.
- b) Es incorrecta porque la unidad identifica específicamente cinco fundamentos.
- c) Es incorrecta porque la subjetividad del evaluador puede dar una inclinación inadecuada al estudio, mientras que si toma como guía los criterios sugeridos no quedará ignorado ningún fundamento principal.

PREGUNTA 2. En esta unidad se afirma que las evaluaciones, en general, deben ser

- a) comparativas, cotejando las alternativas disponibles antes del desarrollo del programa o del servicio.
- b) comparativas tanto antes del comienzo del programa como, periódicamente, después de su desarrollo.
- c) comparativas antes de que el programa empiece, no comparativas durante su desarrollo y ocasionalmente comparativas después.
- d) ni comparativas ni no comparativas.

Respuesta correcta. Debe usted haber señalado la «c». La evaluación comparativa antes del programa permite realizar comparaciones críticas entre distintas alternativas; la evaluación no comparativa durante el desarrollo permite al personal del programa concentrarse en descubrir las mejores maneras de avanzar; y la evaluación periódicamente comparativa después del programa despejará el camino para el logro de mayores progresos.

Respuestas incorrectas

- a) Es incorrecta porque tanto la fase no comparativa como la fase comparativa postoperativa han sido omitidas.
- b) Es incorrecta porque la fase no comparativa ha sido omitida.
- d) Es incorrecta porque cualquier evaluación es comparativa no comparativa o ambas cosas, dependiendo de las necesidades informativas de los clientes y/o la audiencia de la evaluación.

PREGUNTA 3. ¿Cuál de las siguientes afirmaciones es más coherente con lo que se ha dicho en esta unidad respecto de los métodos que pueden ser apropiadamente empleados en los estudios evaluativos?

- a) Evitar el trabajo descuidado, que a menudo es consecuencia de intentar utilizar muchos métodos; un evaluador debe especializarse en la utilización de una sola técnica, como la planificación experimental o el estudio de un caso.
- b) En general, resulta apropiado seguir un enfoque ecléctico para la selección de métodos, pero el evaluador debe evitar la utilización de métodos poco científicos, como el estudio de casos o el análisis político.
- c) El panorama actual de los métodos que pueden emplearse en una evaluación es ya muy avanzado y los evaluadores deben concentrarse no en valorar los métodos existentes y desarrollar otros nuevos, sino en aprender y utilizar la metodología disponible, especialmente la que procede de la investigación y la psicología educativas.
- d) En general, los evaluadores necesitan ser capaces de utilizar una amplia gama de técnicas de investigación, y los que se ocupan de elaborar métodos de evaluación deben desarrollar nuevas técnicas.

Respuesta correcta. Debe usted haber señalado la «d». Esta unidad recomienda, a la vez, el empleo de un acercamiento ecléctico a la evaluación y la necesidad de nuevas estrategias e instrumentos evaluativos.

Respuestas incorrectas

- a) Es incorrecta porque la unidad recomienda el uso de múltiples técnicas.
- b) Es incorrecta porque la unidad fomenta la utilización de una amplia gama de técnicas, incluyendo el estudio de casos y el análisis político.
- c) Es incorrecto porque esta unidad recomienda el empleo de una amplia gama de las metodologías existentes y, a la vez, el desarrollo de nuevos métodos.

PREGUNTA 4. ¿Cuál de las siguientes afirmaciones caracteriza mejor la época tyleriana?

- a) Se introdujeron los tests estandarizados para conseguir que la educación resultara más eficiente.
- b) El método tyleriano se distinguió por concentrarse en unos objetivos claramente fijados y en su consecución.
- c) Los evaluadores aprendieron a ayudar a quienes elaboraban currículos a clarificar el comportamiento de los estudiantes, que sale a la luz con la realización de un currículo.
- d) Las comparaciones internas entre resultados y objetivos resultan costosos.

Respuesta correcta. Debe usted haber señalado la «b». Es lo más destacable de la filosofía de Tyler y del método práctico de evaluación.

Respuestas incorrectas

- a) Es incorrecta porque los tests estandarizados se desarrollaron en el período pre-tyleriano (a principios de este siglo).
- c) Aunque esto no es incorrecto, se trata de una descripción muy pobre de la alternativa de Tyler a las concepciones más tempranas de la evaluación.
- d) Incorrecta; de hecho, por el contrario, las evaluaciones internas han de mostrado ser menos costosas porque no requieren los servicios de un especialista en evaluación.

PREGUNTA 5. ¿Qué afirmación se corresponde *mejor* con el período de la evolución de la evaluación denominado la época de la «inocencia» (1946-1957)?

- a) Aunque se daba una gran expansión de la educación, la sociedad no tenía un particular interés en formar evaluadores responsables.

- b) Las evaluaciones no parecían ocuparse demasiado de la exactitud de las diferencias existentes entre las calificaciones de cada estudiante.
- c) Las agencias federales empezaron a sentir un gran interés por la evaluación de programas.
- d) Durante este tiempo, se crearon nuevos servicios de tests, algoritmos, planificaciones experimentales y procedimientos estadísticos, como respuesta a un análisis de las necesidades de la educación.

Respuesta correcta. Debe usted haber señalado la «a». Esta afirmación refleja fielmente el desinterés de la evaluación por la responsabilidad o, lógicamente, por otras cosas, incluso aunque las disposiciones educativas estuvieran extendiéndose rápidamente.

Respuestas incorrectas

- b) Incorrecta; la concentración en las encuestas regularizadas condujo a una gran preocupación por la exactitud de las diferencias de calificaciones entre los estudiantes.
- c) Es incorrecta, porque este hecho no se produjo hasta los años sesenta.
- d) Es incorrecta, porque, lamentablemente, no fue éste el caso; procedimientos como los mencionados se ocupaban poco o nada de las necesidades educativas, incluida la de su mejora.

PREGUNTA 6. ¿Cuál de las siguientes afirmaciones se corresponde más con el desarrollo de la evaluación durante la época del realismo (1958-1972)?

- a) Desde 1965, se invirtieron billones de dólares en programas cuyo objetivo era la igualdad y la mejora de las oportunidades en cuestiones de servicios sanitarios sociales y educativos.
- b) El artículo 1 del Acta de la Educación Elemental y Secundaria de 1965 tenía como objetivo proporcionar una educación compensatoria a los niños atrasados. Las adecuadas metodologías evolutivas utilizadas en estos programas tuvieron, afortunadamente, éxito.
- c) El énfasis puesto en las puntuaciones de los tests hizo que la atención se dirigiera hacia los propios programas.
- d) Las duras críticas de CRONBACH (1963) que afirmaron que los métodos evaluativos de entonces carecían de relevancia, y que la evaluación debía convertirse en un proceso de recogida y elaboración de informaciones que sirvieron como guía para la realización de un currículo, fueron muy oportunas

Respuesta correcta. Debe usted haber señalado la «d». Las críticas de Cronbach, contenidas en su artículo de 1963, constituyeron un gran paso adelante en el proceso de descubrir metodologías apropiadas para satisfacer las exigencias evaluativas de los años sesenta.

Respuestas incorrectas

- a) Es incorrecta, porque se trata de una afirmación acerca de un programa político y no estrictamente acerca de la evaluación, incluso aunque los programas posteriores condujeran a un rápido desarrollo de las metodologías evaluativas.
- b) Es incorrecta, porque las metodologías evaluativas empleadas en esta época eran totalmente inadecuadas.
- c) Incorrecta; la verdad está más cerca de todo lo contrario, puesto que las puntuaciones de los tests desviaron la atención de los programas en sí y de su perfeccionamiento.

PREGUNTA 7. ¿Cuál de las afirmaciones siguientes caracteriza mejor, de acuerdo con esta unidad, el estado de la evaluación en 1983?

- a) Los evaluadores han realizado un significativo progreso hacia el establecimiento de la evaluación como práctica formal especializada.
- b) De un modo creciente, la evaluación se fue considerando sólo como una aspirante a la categoría de profesión. Esta área laboral dejó de ser útil en la década de los sesenta: el público la fue dejando a un lado y ya no recibió más ayudas económicas.
- c) En el área de la metodología, los evaluadores han llegado casi a un completo acuerdo mediante el cual las evaluaciones deben emplear métodos naturalistas y evitar la utilización de métodos cuantitativos.
- d) En respuesta al mayor desarrollo metodológico de la pasada década, se han realizado minuciosos cambios y reformas en la práctica evaluativa.

Respuesta correcta. Debe usted haber señalado la «a». Los evidentes síntomas de profesionalización de la evaluación que se dan en 1983 incluyen la aparición de publicaciones periódicas, de numerosos libros y monografías sobre el tema, de organizaciones profesionales que se dedican a la evaluación, de normas profesionales para la práctica, de un programa concreto destinado a la titulación, y un aumento de la investigación y el desarrollo de temas sobre evaluación.

Respuestas incorrectas

- b) Es incorrecta porque, contrariamente a las afirmaciones previas, acerca de que la evaluación formalizada era una moda que pasaría pronto, la especialidad ha empezado a adquirir las características de una profesión.
- c) Es incorrecta porque existe una profunda división entre aquellos que son partidarios de los métodos experimentales/cuantitativos y los que recomiendan métodos naturalistas/cualitativos.
- d) Es incorrecta porque, a pesar de la creciente competencia de los evaluadores, la práctica evaluativa actual, según esta unidad, ha cambiado muy poco en la mayoría de los casos.

INTERPRETACIÓN DEL TEST DE CONOCIMIENTOS. Cuando haya terminado el test de conocimientos le rogamos que examine cómo lo ha hecho. Le recomendamos que revise las partes más importantes del test, prestando más atención a las preguntas en las que se ha equivocado o en las que ha dudado, e incluso aquellas en las que se ha sentido un poco confuso ante nuestras respuestas y explicaciones.

Ejercicio práctico

Este apartado consiste en un ejercicio que pretende ayudarle a valorar si ha comprendido bien cómo pueden aplicarse en la práctica los principales contenidos de la unidad. Le sugerimos que anote su respuesta en una hoja de papel aparte y luego la compare con la respuesta ejemplo que sigue.

EJERCICIO 1. Suponga que el director del Departamento de Bienestar de una gran ciudad le contrata para discutir los planes de evaluación de sus programas que le ha facilitado la agencia. Suponga además que el director de la agencia ha asumido el concepto de «auto-evaluación», de acuerdo con el cual todos los miembros del personal de la agencia son responsables de evaluar su propio trabajo. El jefe advierte también que no se han utilizado los servicios de un departamento de evaluación, ni los de unos especialistas (ya sea durante todo el trabajo o parte de él), ni se ha pagado de una manera especial el trabajo evaluativo. En lugar de eso, ha recordado a los 95 miembros del personal que son responsables de evaluar su propio trabajo y les ha enseñado un método evaluativo con el fin de que sus esfuerzos lleguen a buen puerto. El director cree que este tipo de actividad evaluativa puede ser de utilidad para convencer a la legislatura estatal de que incremente sus ayudas a la agencia.

Las respuestas al ejercicio 1 deben incluir convenientemente los siguientes puntos:

- a) Cada profesional es totalmente responsable de la calidad y, por lo tanto, de la evaluación de sus contribuciones.
- b) Aunque los profesionales necesitan una cierta preparación sobre evaluación, este trabajo es demasiado complejo como para suponer que cada uno de los profesionales, actuando independientemente, puede hacer todo lo que la evaluación necesita para solucionar los problemas que la legislación estatal plantea a la agencia.
- c) Indudablemente, una preparación metodológica apropiada puede ser un positivo paso adelante hacia la profesionalización del trabajo de la agencia, pero entre la planificación de esa preparación y la decisión acerca de cómo distribuir los recursos de la evaluación, el que toma las decisiones en la agencia debe tener en cuenta una amplia gama de funciones evaluativas.
- d) Los papeles que se desempeñan en el trabajo evaluativo de una agencia son, más o menos, los siguientes: cliente de la evaluación, planificador de la evaluación, coordinador de los servicios evaluativos, evaluador especializado, asesor técnico, especialista en información y especialista en comunicación.

Preguntas sin respuestas

Este apartado contiene preguntas destinadas a ayudarle a calibrar y profundizar su comprensión de los asuntos tratados en esta unidad. No se proporcionan respuestas a las preguntas. Le sugerimos que intente dar respuestas a estas preguntas siguiendo ciertos pasos, como revisar la unidad, consultar las fuentes bibliográficas adecuadas, incluyendo las que se consignan al final de la unidad, y discutiendo las preguntas con sus alumnos o colegas.

- 1. La definición de evaluación que hemos dado en esta unidad, ¿es igual mente defendible desde el punto de vista de escuelas filosóficas distintas, como el pragmatismo, el existencialismo o el positivismo lógico?
- 2. Teniendo en cuenta que esta unidad recomienda un enfoque ecléctico a la hora de escoger los métodos que se usarán en un trabajo evaluativo, ¿existe alguna investigación cuyas conclusiones se refieran a la conveniencia o inconveniencia de elegir eclécticamente métodos de especialidades afines —como la investigación educativa, la investigación psicológica, la contabilidad y la planificación ciudadana—, contribuyendo así a recomen-

- dar o refutar el enfoque ecléctico para la elección de métodos investigativos?
3. Algunas personas han sugerido que algunas de las *Normas* del Joint Committee deben ser tenidas en cuenta antes que otras, es decir, que el orden de presentación y aplicación de cada una de las normas debe ser considerado muy importante. ¿Está usted de acuerdo o no? Si lo está, ¿cómo ordenaría usted las *Normas* con el fin de aplicarlas para tomar una decisión acerca de llevar a cabo o no una evaluación que le han propuesto? Si cree que el orden en el que se estudien las *Normas* no es importante, ¿por qué piensa que no lo es?
 4. Examine el repaso histórico de la teoría y la práctica evaluativas contenido en esta unidad. ¿Cómo evaluaría el progreso, la regresión o el estancamiento de la especialidad en lo que se refiere a la responsabilidad social, la erudición o la repercusión?
 5. Teniendo en cuenta el alcance de la unidad y, especialmente, el análisis de los papeles que se desempeñan en un trabajo evaluativo, ¿hasta qué punto cree que la sociedad debe invertir en preparar y emplear a especialistas en lugar de simples evaluadores? ¿Dónde, cómo y hasta qué punto debe proporcionárseles preparación a ambos tipos de evaluadores?

Referencias bibliográficas

- AMERICAN PSYCHOLOGICAL ASSOCIATION. 1974. *Standards for educational and psychological tests*, Washington, D.C., Autor.
- ANDERSON, S.; BALL, S.; y MURPHY, R. y colaboradores. 1974. *Encyclopedia of educational evaluation*, San Francisco, Jossey-Bass.
- BLOOM, B. S.; ENGLEHART, M. D.; FURST, E. J.; HILL, W. H.; y KRATHWOHL, D. R. 1956. *Taxonomy of educational objectives. Handbook 1: The cognitive domain*, Nueva York, David McKay (trad. cast.: *Clasificación de las metas educativas. Tomo I: Taxonomía de los objetos*, Alcoy, Marfil, 1979).
- BRINKERHOFF R.; BRETHOWER, D.; HLUCHYJ, T.; y NOWAKOWSKI, J. 1983. *Program evaluation: A practitioners' guide for trainers and educators*, Boston, Kluwer-Nijhoff.
- CAMPBELL D. T. y STANLEY, J. C. 1963. Experimental and quasi-experimental designs for research on teaching, en N. L. Gage (comp.), *Handbook of research on teaching*, Chicago, Rand McNally.
- COLEMAN J. S.; CAMPBELL, E. Q.; HOBSON, C. J.; MCPARTLAND J.; MODD, A. M.; WEINFELD, F. D.; y YORK, R. L. 1966. *Equality of educational opportunity*. Washington, D.C. Oficina de Educación, Departamento de Salud, Educación y Bienestar de los EE.UU.
- COOK, D. L. 1966. *Program evaluation and review technique: Applications in Education*, Washington, D.C. Oficina de Imprenta del Gobierno.
- . 1971. *Educational project management*, Columbus, Ohio, Charles E. Merrill.

- CRONBACH, L. J. 1963. Course improvement through evaluation, *Teachers College Record*, 64, págs. 672-683.
- . 1980. *Toward reform of program evaluation*, San Francisco, Jossey-Bass.
- EISNER, E. W. «Educational objectives: Help or hindrance?» *The School Review*, 75 (1967), págs. 250-260.
- . 1975. *The perceptive eye: Toward the reformation of educational evaluation*, Stanford, California, Stanford Evaluation Consortium, diciembre.
- GLASER, R. «Instructional technology and the measurement of learning outcomes: Some questions», *American Psychologist*, 18 (1963), págs. 519-521.
- GLASS, G. 1976. Primary secondary and meta analysis of research, *Educational Researcher*, 5 (10), págs. 3-8.
- GUBA, E. G. «A Study of Title III activities: Report on evaluation» (National institute for the Study of Educational Change, Universidad de Indiana, octubre de 1966) (ciclostilado).
- GUBA, E. G. y LINCOLN, Y. S. 1981. *Effective evaluation*, San Francisco, Washington, Londres, Jossey-Bass.
- HAMMOND, R. L. 1967. «Evaluation at the local level». Conferencia para el Miller Committee for the National Study of ESEA Title III.
- JOHNSON, R. 1980. Directory of evaluators and evaluation agencies, Nueva York, EXXON Corporation.
- Joint Committee on Standards for Educational Evaluation. 1981. *Standards for evaluations of educational programs, projects, and materials*, Nueva York, Mc Graw-Hill.
- KAPLAN, A. 1964. The conduct of inquiry, San Francisco, Chandler.
- KELLAGHAN, T. 1982. «Los sistemas escolares como objeto de evaluación». En Daniel STUFFLEBEAM; Thomas KELLAGHAN; y Benjamín ÁLVAREZ (comp.), *La Evaluación Educativa*, Bogotá, Pontificia Universidad Javeriana.
- KELLAGHAN, T.; MADAUS, G.; y AIRASIAN, P. 1982. *The effects of standardized testing*, Hingham, Mass., Kluwer-Nijhoff Publishing.
- KRATHWOHL, D. R.; BLOOM, B. S.; y MASIA, B. B. 1964. *Taxonomy of educational objectives: The classification of educational goals. Handbook II: Affective domain*, Nueva York, David McKay (trad. cast.: *Taxonomía de los objetivos de la educación. Clasificación de las metas educativas*, Alcoy, Marfil, 1979).
- METFESSEL, N. S. y MICHAEL, W. B. «A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs». *Educational and Psychological Measurement*, 27 (1967), págs. 931-943.
- PECK, H. «Report on the certification of evaluators in Louisiana». Texto presentado en el congreso de la Southern Educational Research Association, Lexington, Kentucky, Fall 1981.
- POPHAM, W. J. 1971. *Criterion-referenced measurement*, Englewood Cliffs, N. J.; Educational Technology Publications (trad. cast.: *Evaluación basada en criterios*, Magisterio Español, 1983).
- PROVUS, M. 1969. *Discrepancy evaluation model*, Pittsburgh, Pennsylvania, Pittsburgh Public Schools.
- REINHARD, D. 1972. *Methodology developments for input evaluation using advocate and design teams*. Tesis doctoral inédita, Universidad del Estado de Ohio.
- RICE, J. M. 1897. The futility of the spelling grind, *The Forum*, 23, págs. 163-172.

- . 1914. *Scientific management in education*, Nueva York, Hinds Noble and Eldredge.
- ROTH, J. 1977. Needs and the needs assessment process, *Evaluation News*, 5, páginas 15-17.
- SCRIVEN, M. S. 1967. The methodology of evaluation En *Perspectives of curriculum evaluation* (AERA Monograph Series on Curriculum Evaluation n. 1), Chicago, Rand McNally.
- . 1974. Prose and cons about goal-free evaluation, *Evaluation Comment*, 3, páginas 1-4.
- , 1975. *Evaluation bias and its control. Occasional Paper Series*, n. 4, Universidad de Western Michigan. Centro de Evaluación.
- SMITH, E. R. Y TYLER, R. W. 1942. *Appraising and recording student progress*, Nueva York, Harper.
- SMITH, N. L. 1981a. *Metaphors for evaluation: Sources of new methods*, Beverly Hills, Sage.
- . 1981b. *New techniques for evaluation*, Beverly Hills, Sage.
- STAKE, R. E. 1967. The countenance of educational evaluation, *Teachers Collage Record*, 68, págs. 523-540.
- . 1975. *Program evaluation, Particularly responsive evaluation. Occasional Paper Series* n. 5, Universidad de Western Michigan, Centro de Evaluación.
- . 1978. The case-study method in social inquiry, *Educational Researcher* 7, páginas 5-8.
- . «Setting standards for educational evaluators» *Evaluation News* n. 2, 2 (1981), págs. 148-152.
- STUFFLEBEAM, D. L. «The use and abuse of evaluation in Title III», *Theory Into Practice*, 6 (1967), págs. 126-133.
- . 1978. Metaevaluation: An overview, *Evaluation and the Health Professions*, 1 (2), págs. 146-163.
- STUFFLEBEAM, D. L., y otros. 1971. *Educational evaluation and decision making*, Itasca, Ill., Peacock.
- SUÁREZ, T. 1980. *Needs assessments for technical assistance: A conceptual overview and comparison of three strategies*. Tesis doctoral inédita. Universidad de Western Michigan.
- TYLER, R. W. 1967. Changing concepts of educational evaluation. En R. E. Stake (comp.), *Perspectives of curriculum evaluation* (vol. 1), Nueva York, Rand McNally.
- WEBSTER, W. J. 1975. *The organization and functions of research and evaluation units in a large urban school district*, The Dallas Independent School District, Dallas.