

CAÍTULO 6

LOS ESTÁNDARES PARA LA INVESTIGACIÓN EVALUATIVA

OBJETIVOS

Tras el estudio del presente capítulo el lector debería ser capaz de:

- Comprender la necesidad de disponer de estándares que guíen la realización de evaluaciones de programas.
- Comprender y saber utilizar el vocabulario técnico correspondiente a los estándares.
- Conocer, comprender, analizar y valorar los cuatro grandes criterios en que se organizan los estándares.
- Conocer, comprender, analizar y valorar los estándares integrantes de los diferentes criterios organizativos.
- Analizar críticamente los estándares y su organización así como las ausencias y las propuestas alternativas del autor.

A estas alturas de la exposición habrá quedado suficientemente claro que la evaluación de programas es una actividad compleja, con evidentes connotaciones técnicas, orientada a la aportación de información útil a diferentes partes interesadas, pero difícil por ambiciosa y con evidentes riesgos de ser mal utilizada.

En tales condiciones, no parece difícil aceptar la conveniencia de establecer algún tipo de reglas de buen hacer que orienten a las personas que se adentren en este tipo de actividad. Y así ha sido.

6.1. ESTÁNDARES PARA LA EVALUACIÓN DE PROGRAMAS

Ya en 1974, un Comité nombrado por la APA (Asociación Americana de Psicología), la Asociación Americana de Investigación en Educación y el Consejo Nacional de Medición en Educación, dio a conocer unos estándares para los Manuales y los tests de Psicología, revisando la edición de 1966.

En 1975 el *Joint Committee on Standards for Educational Evaluation* se amplió con la incorporación de doce nuevos miembros, procedentes de sendas organizaciones, dando comienzo a un trabajo que dio a conocer en 1981; se trataba de la publicación *Standards for Evaluations of Educational Programs, Projects and Materials*, en la que se propone un total de 30 normas, agrupadas en cuatro grandes categorías, exigencias o criterios (cuadro 1). En 1988 publicó los estándares para la evaluación del personal, editados por Sage. Posteriormente, en 1989, el proceso seguido por el Joint Committee fue acreditado por el Instituto Americano Nacional de Estándares.

CUADRO 1: *Criterios de calidad de las evaluaciones.*

CRITERIO	NORMAS
UTILIDAD	Se trata de ocho normas destinadas a asegurar que la evaluación proporcione la información que necesite cada una de las audiencias ¹ legitimadas. Habrá de ayudarles a identificar lo que está bien y lo que está mal, lo que es más importante y las formas de abordar las mejoras.
VIABILIDAD	Se trata de tres normas, concebidas para promover una evaluación realista, prudente, diplomática y moderada. Los procedimientos a utilizar deben estar al alcance del evaluador, sin crear problemas entre los afectados, pero controlando los intentos de ciertos grupos de inmiscuirse en la evaluación.
HONRADEZ	Este criterio está integrado por ocho normas, destinadas a asegurar una evaluación realizada con planteamientos legales y éticos, respetando a quienes intervienen en la evaluación y están interesados en las conclusiones. El informe debe ser equitativo.
PRECISIÓN	Este bloque es el más amplio, con un total de once normas, fundamentalmente técnicas, destinadas a asegurar la calidad de la información sobre la realidad estudiada, de forma que sea posible determinar su <i>valor</i> y su <i>mérito</i>

Años más tarde, en 1994, tomando como base los principios de la publicación de 1981, se dio a conocer una revisión de los mismos bajo el título *The Program Evaluation Standards*, que incluye algunas modificaciones.

6.2. CONCEPTOS GENERALES SOBRE LOS ESTÁNDARES

Los autores de los estándares los han concebido como una herramienta útil tanto para los evaluadores –bien desde una con-

¹ El término “audiencias” no parece muy adecuado en nuestro idioma. Personalmente prefiero hablar de “partes interesadas, implicadas o afectadas”.

cepción limitada (sólo evaluadores) bien en una concepción amplia (diseño, asesoramiento y evaluación)– como para los usuarios de los programas, a fin de que sean capaces de valorar las evaluaciones a las que puedan acceder. Sus autores los consideran también como un instrumento de formación en la Universidad, tanto para profesionales como para directivos y profesores.

Los autores, el Joint Committee, definen un estándar como un principio acordado por las personas comprometidas en una práctica profesional destinada a mejorar la calidad de la misma. Consideran que los estándares “proporcionan una guía para evaluar los programas de educación y formación, proyectos y materiales”; en conjunto, y a su juicio, proporcionan una estructura para diseñar y asesorar evaluaciones.

A los efectos de entender el sentido y naturaleza de los estándares, conviene indicar cómo se definen en ellos los conceptos básicos de *evaluación*, *programa* y *proyecto*.

Por *evaluación* se entiende la investigación sistemática del valor o mérito de un objeto, en este caso el programa de educación y formación, aunque pueda aplicarse a proyectos y materiales. *Proyecto* se define como un conjunto sistemático de actividades educativas cuyo desarrollo ocupa un período determinado de tiempo. Cuando un proyecto se institucionaliza se convierte en un *programa*, entendido como un conjunto de actuaciones sistemáticas.

Otros conceptos importantes son los relativos al evaluador, la información, el cliente o la audiencia. Se entiende por *información* el conjunto de datos, cuantitativos o no –desde una narración o un gráfico a una estadística o informe– que ayuda a concretar los resultados, responder a las preguntas que se hace el evaluador o el cliente e incrementar el conocimiento y comprensión disponibles sobre el programa. *Evaluador* es un término amplio que puede aplicarse a cualquier persona que realice una evaluación; *cliente* es la persona, grupo u organización que encarga la evaluación, que la contrata en definitiva; por *audiencia* se entiende las personas o grupos implicados o afectados por una evaluación de programas².

² Recuérdese mi preferencia por el término “partes interesadas, implicadas o afectadas”.

6.3. ORGANIZACIÓN Y CONTENIDO DE LOS ESTÁNDARES

Como hemos indicado más arriba, los estándares están organizados en torno a cuatro grandes criterios que deben ser tomados en consideración a la hora de encargar evaluaciones, diseñarlas, dirigir las, llevarlas a cabo y darlas a conocer. Se trata de los cuatro siguientes:

- Utilidad
- Viabilidad
- Honradez
- Precisión

Dado que algunos estándares podrían situarse dentro de más de uno de los anteriores criterios, la propuesta del Joint Committee los ubica en aquel que, a su juicio, resulta especialmente adecuado.

Al leer y reflexionar sobre los diversos estándares, el lector podrá comprobar que algunos de ellos, en su acepción más completa, tienen poco que ver con la situación en que se desarrollan muchas evaluaciones en nuestro ámbito educativo; del mismo modo, muchos de ellos están pensados para la actuación de evaluadores profesionales, que formalizan su trabajo profesional mediante acuerdos o contratos con clientes. Desde ese punto de vista, pues, parecerían poco adecuados y hasta se podría preguntar por la conveniencia de su inclusión en esta monografía.

Debemos decir, no obstante, que el fondo de la gran mayoría de los estándares debe ser respetado, puesto que sí es de aplicación a cualquier evaluación de programas, y que los evaluadores, sean profesionales de la evaluación o profesionales que realizan evaluaciones (directivos de centros o entidades educativas, educadores sociales, orientadores) deben tomarlos como referencia y modelo para su actividad evaluadora.

En la medida en que los estándares son principios, su aplicación o no, y el detalle y profundidad con que deben aplicarse es una cuestión de juicio, que deberá considerarse con el debido cuidado el evaluador junto a las personas responsables del programa y las que le encargan –si así fuere– la evaluación.

Analícemos brevemente los diversos estándares, agrupados en torno a los diferentes criterios:

3.1. *Estándares de utilidad*

Este bloque está integrado por un total de siete normas, cuyo sentido es el de asegurar que la evaluación ofrezca la información necesaria y adecuada a cada una de las audiencias o partes interesadas.

A los efectos oportunos, será preciso que los evaluadores informen sobre sí mismos, delimiten sus audiencias o partes interesadas, identifiquen sus necesidades de información, planifiquen correctamente la recogida de información adecuada a tales necesidades y les informen de forma clara y suficiente.

Veamos lo fundamental de cada uno de estos estándares.

U1. Descripción de la audiencia. Las personas implicadas o afectadas por la evaluación deben ser identificadas a fin de poder atender sus necesidades.

Teniendo en cuenta que una evaluación puede afectar o, al menos, interesar a diferentes audiencias –prefiero el término *partes interesadas*– parece necesario que, antes de comenzar una evaluación, se identifique, además de a los clientes³, a tales partes interesadas.

De este modo se podrán conocer sus intereses y necesidades para poder ofrecerles la información necesaria, adecuada y útil. Junto a ello, en la medida precisa y siempre que sea necesario, se podrá facilitar su implicación en la propia evaluación (recepción de sugerencias e iniciativas para una correcta planificación, diseño e implementación).

Debe tenerse siempre en consideración que una evaluación, por buena que sea, si no es conocida y aceptada por las partes afectadas e interesadas, puede quedar, en el mejor de los casos, en algo inútil; en ocasiones puede llegar a ser objeto de críticas que la desacrediten y hasta del rechazo activo por algunos colectivos.

U2. Credibilidad del evaluador. Las personas que llevan a cabo las evaluaciones deben ser dignas de confianza y competentes para realizar su actividad, de forma que sus trabajos y aportaciones alcancen el máximo de credibilidad y aceptación.

³ Quienes contratan los servicios del evaluador.

Por lo general, las evaluaciones, en la medida en que deben concluir en juicios valorativos y en la identificación de los puntos fuertes y débiles de los programas, suscitan resistencias más o menos fuertes entre las personas afectadas.

Para que tales juicios puedan ser aceptados y, por consiguiente, puedan incidir en la mejora de los programas, resulta imprescindible que al evaluador –persona o equipo– se le reconozca competencia profesional, fruto de su experiencia y nivel de formación; junto a ello, el evaluador debe suscitar confianza entre las partes interesadas y afectadas como consecuencia de sus cualidades personales, entre las que cabe destacar la integridad, la ecuanimidad y la profesionalidad.

Dada la importancia de este estándar, sería conveniente asegurarse de que el evaluador –persona o equipo– merece tal credibilidad para su cliente y demás partes interesadas pues, de no ser así, parece preferible no abordar un trabajo que puede no tener consecuencias para la mejora del programa y de su personal.

El evaluador debe ser consciente de que en cada evaluación afianza o reduce su credibilidad como consecuencia de su actuación durante el proceso y del rigor y seriedad de sus informes, aspecto este para el que los estándares le ofrecen una guía de enorme valor y utilidad.

U3. Alcance y selección de la información. La información recogida debe ser tan amplia y adecuadamente seleccionada como para incluir cuantas cuestiones se consideren pertinentes y relevantes en relación con el objeto de evaluación –el programa– y para poder dar la debida respuesta a los intereses y necesidades de las diferentes partes interesadas (audiencias).

Como se ha indicado en otro lugar, la evaluación no puede nunca ser mejor que la información en que se basa; cualquier insuficiencia, deficiencia o sesgo afectará necesariamente a la calidad de la evaluación, de los juicios de valor.

Por consiguiente, un elemento básico de toda evaluación es la identificación de toda cuanta información se considere relevante *a priori* para satisfacer las necesidades de las diferentes partes interesadas, para responder a las preguntas a las que debe dar respuesta la evaluación.

Ahora bien, una vez identificada toda la información potencialmente interesante y útil, es preciso analizar si con los medios disponibles –de personal, económicos, técnicos– y con el tiempo asignado, los evaluadores deben proceder a establecer prioridades en el caso de que se deba prescindir de alguna información. Para establecer las prioridades, el evaluador debe atender tanto a criterios científicos como a los de carácter profesional y hasta a las demandas de las partes interesadas.

Establecidos estos criterios, es conveniente que evaluador y cliente, junto a las principales partes interesadas –audiencias– logren acuerdos en este importante y básico aspecto.

U4. Interpretación. Los evaluadores deben describir con el debido detalle las perspectivas, procedimientos y bases lógicas utilizadas para interpretar los resultados a fin de que sean patentes las bases de sus juicios de valor.

Lo que convierte a la información en evaluación es su valoración, pero la evaluación es una actividad que, siendo técnica, admite diversidad de planteamientos y resultados en función de los criterios, experiencias, teorías o formación del evaluador.

La subjetividad que afecta a los juicios evaluativos debería ser puesta de manifiesto a fin de que las diversas partes interesadas conozcan con claridad las perspectivas, lógica y procedimientos en que se apoya el evaluador, teniendo en cuenta que los diversos contextos pueden originar valoraciones diferentes para unos mismos hechos, datos o información.

En cualquier caso, lo que parece imprescindible es dejar constancia de quiénes han de emitir las valoraciones y, en su caso, la forma de proceder para integrar tales juicios y para tomar las decisiones pertinentes si procediera.

U5. Claridad del informe. El informe en que toma cuerpo la evaluación debe describir el objeto evaluado –el programa– en su contexto, así como los objetivos, procedimientos y resultados de la evaluación, a fin de proporcionar la información fundamental y facilitar su comprensión.

Como se ha señalado, toda evaluación parte de unas necesidades e intereses, generalmente traducidas a unas preguntas, que guían el proceso evaluador; la evaluación es el proceso sistemático llevado a cabo para dar la respuesta oportuna y se concreta en un *informe* en el que se recoge, como tendremos ocasión de com-

probar, lo fundamental de la evaluación: puntos fuertes y débiles a fin de afianzar los primeros y solucionar los últimos.

Ahora bien, para que tal etapa de la evaluación resulte verdaderamente útil, debe contar con ciertas cualidades, donde destaca la de su claridad a fin de que los destinatarios puedan extraer de él la mayor utilidad posible. A tales efectos, el informe debe ser completo, de forma que recoja el objeto –el programa– en su contexto, además de incorporar los elementos significativos del propio programa, entre los que cabe destacar sus objetivos y procedimientos.

U6. Difusión y oportunidad del informe. Los hallazgos significativos y el informe de la evaluación deben darse a conocer a las partes afectadas e implicadas a fin de que puedan ser debidamente utilizados.

La información que debe aportar la evaluación del programa deja de ser útil si no llega a las partes afectadas e interesadas. Tales partes están legitimadas, además del propio cliente de la evaluación, para conocer los informes.

Dada la diversidad de las audiencias o partes interesadas, es posible que no todas tengan los mismos intereses y que sus diferencias en formación para entenderlos, comprenderlos y extraer información para la mejora aconsejen una formulación diferenciada para los diferentes grupos.

Conviene dejar constancia de que en este punto pueden surgir conflictos entre el evaluador y los clientes que, con alguna frecuencia, pretenden reservarse la información o, al menos, ciertas partes de la misma. Para evitar estos problemas se recomienda tomar acuerdos previos, tal como se indica en el estándar P2. *Acuerdos formales*, recogido en el apartado de Honradez o probidad.

U7. Impacto de la evaluación. Las evaluaciones deben ser planificadas, dirigidas, llevadas a cabo y dadas a conocer de forma que se estime el seguimiento del trabajo por parte de las personas afectadas e interesadas a fin de aumentar las probabilidades de que la información sea utilizada.

Por *impacto* entendemos el influjo o la incidencia que pueda tener la evaluación sobre las decisiones y actuaciones de las partes afectadas e interesadas. A tales efectos, los evaluadores pueden

jugar un importante papel de asesoría y apoyo a las diferentes partes interesadas en el programa y su evaluación.

En principio, la evaluación no produce por sí misma mejora alguna, aunque es preciso dejar constancia de que el mero hecho de evaluar produce efectos –positivos o negativos– en los programas y en las personas a las que afecta.

Para conseguir que la evaluación produzca mejoras es conveniente que los evaluadores asistan y asesoren a las partes interesadas; las orientaciones del evaluador pueden contribuir a que se enfrenten con el programa desde un punto de vista diferente.

6.3.2. Estándares de viabilidad o factibilidad

Se trata de tres estándares concebidos para asegurar que las evaluaciones sean realistas, prudentes y adecuadas. En concreto son los siguientes: a) *Procedimientos prácticos*; b) *Viabilidad política*; c) *Eficacia en los costes*.

F1. Procedimientos prácticos. Los procedimientos utilizados en la evaluación deben ser los adecuados para alcanzar la información necesaria creando el mínimo de problemas posibles.

La información resulta ser el elemento o componente imprescindible de cualquier evaluación. Los procedimientos son diversos e incluyen un amplio abanico, entre los que conviene destacar los relativos a los acuerdos con los clientes, la selección de fuentes e instrumentos, la propia recogida de la información, su análisis y la difusión de los resultados.

Muchos de tales procedimientos pueden originar problemas en los lugares en que deben llevarse a cabo. El estándar reclama la necesidad de que los evaluadores tomen las medidas precisas para minimizarlos; junto a ello, deberán asegurarse de que tales procedimientos sean realistas atendiendo a aspectos como el presupuesto necesario, el tiempo disponible o el personal preciso.

Evaluaciones perfectas pueden llegar a ser imposibles de llevarse a la práctica. La viabilidad reclama un ejercicio de búsqueda de evaluaciones que, sin renunciar a los aspectos esenciales, no fracasen por su nivel de utopía.

F2. Viabilidad política. La planificación y dirección de la evaluación debe procurar la cooperación de los grupos interesados así como evitar cualquier intento de intervenir en su desarrollo o en sus resultados.

En la medida en que las evaluaciones se planifican con la finalidad de tomar decisiones, no debe extrañar que haya personas o grupos con intereses específicos que pueden pretender influir en la evaluación, algo que no debe ocurrir. Pero, por otro lado, sin la activa cooperación de las partes implicadas resulta impensable una evaluación que pueda llegar a ser útil.

Al evaluador le incumbe conseguir tal colaboración y evitar la injerencia si en verdad pretende conseguir credibilidad para su trabajo.

Como se puede comprender, en este tipo de cuestiones son importantes algunas de las cualidades a las que nos hemos referido al hablar, en el marco de los estándares de utilidad, de la *credibilidad* del evaluador.

F3. Eficacia en los costes (Coste – efectividad). La evaluación ha de orientarse al logro de la información necesaria utilizando únicamente los recursos precisos.

Los recursos son, en general, limitados. La evaluación consume recursos, pero realiza aportaciones que pueden resultar de gran valor para la mejora de los programas y para las diversas partes interesadas. Cabe, por tanto, esperar que los resultados de la evaluación compensen los recursos que deben ser utilizados.

El logro de la eficiencia en el uso de los recursos puede plantearse en ocasiones en términos de metodologías para la evaluación: en igualdad de condiciones, se debe elegir la metodología que ofrece las mejores prestaciones a menor costo, no sólo en términos monetarios sino de personal y de tiempo.

Con todo, el establecimiento de la efectividad no es una cuestión fácil; si por un lado, los recursos pueden ser relativamente fáciles de contabilizar, la valoración de las aportaciones de la evaluación resulta francamente difícil de llevar a cabo, sobre todo en la medida en que puede generar una serie de efectos no planeados como el cambio en ciertas actitudes que pueden producir nuevos cambios en el futuro, a veces de gran alcance. Junto a ello es preciso reseñar que algunas de las apor-

taciones de las evaluaciones merecen valoraciones diferentes a las distintas partes interesadas, lo que le confiere un carácter evidentemente subjetivo.

En todo caso, sí parece conveniente dejar como un importante principio la necesidad de utilizar con moderación y prudencia los recursos disponibles.

6.3.3. Estándares de Honradez o Probidad

Este bloque de estándares está integrado por un total de ocho, cuyo sentido es el de asegurar que la evaluación se lleve a cabo de acuerdo con las normas legales y los principios éticos, respetando al máximo el bienestar de quienes intervienen en la evaluación así como de las diversas partes interesadas.

He aquí los ocho estándares:

P1. Orientación al servicio. Las evaluaciones deben diseñarse con la intención de ayudar a las organizaciones a atender y satisfacer las necesidades de los participantes.

Las evaluaciones deben hacer aportaciones relevantes a las organizaciones, a los educadores y a la sociedad, favoreciendo de este modo el aprendizaje. En este sentido, se deben someter a reflexión los objetivos y contribuir a la retirada de los programas no sólo negativos sino ineficaces.

P2. Acuerdos formales. Es preciso plasmar en un documento escrito las obligaciones de las partes implicadas, de forma que cada una de ellas esté obligada a respetarlas o, en su caso, a renegociarlas.

Entendemos por *acuerdo formal* el establecimiento de las condiciones en que se llevará a cabo la evaluación, para lo que es necesario tomar en consideración las necesidades, expectativas y responsabilidades de clientes y evaluadores.

El acuerdo puede ser un sencillo documento o un contrato en toda regla. Sea de una forma o de otra, es preciso acordar aspectos tales como la aprobación del plan general, el sistema de recogida, almacenamiento, organización y análisis de datos; la elaboración, edición y difusión de los informes; la financiación o el sis-

tema de protección de los sujetos que aportan diversas colaboraciones.

El establecimiento de un clima de mutuo respeto y confianza entre cliente y evaluador, fundamental para el desarrollo de la evaluación, puede tener en este inicial acuerdo un marco que evite tensiones posteriores.

P3. Respeto a los derechos humanos. La evaluación debe diseñarse y llevarse a cabo de forma que queden protegidos y asegurados los derechos y el bienestar de las personas a las que pueda afectar la evaluación.

Las evaluaciones deben respetar tanto los preceptos legales como la ética profesional. Entre tales derechos deben citarse expresamente los relativos al libre consentimiento sobre la participación, la libertad de retirarse de la evaluación, la confidencialidad de la información aportada, además, claro está, de la protección de su salud y seguridad.

A fin de asegurarse de tal respeto, es preciso que los evaluadores conozcan tales derechos y estén al tanto de los códigos deontológicos, además de asumir los principios éticos de las investigaciones y evaluaciones.

Los errores en este punto dañan la evaluación, que puede encontrar resistencias futuras, además de contribuir al descrédito tanto de los evaluadores como del propio acto y proceso evaluador.

P4. Relaciones humanas. Los evaluadores deben comprometerse en el respeto de la dignidad e integridad de las personas relacionadas con la evaluación, asegurando que quienes participen no sufrirán daños de ningún tipo.

En toda evaluación se da una red más o menos compleja de interrelaciones humanas que afectan, o pueden hacerlo, tanto a quienes participan como a grupos o personas implicadas o interesadas.

En la medida en que los evaluadores gocen de las habilidades, capacidades y cualidades ya reseñadas, podrán salvaguardar a tales personas de cualquier daño o perjuicio, evitando herir sus sentimientos, ofenderles o provocarles daños o molestias, lo que, de no darse, contribuiría a provocar hostilidad y rechazo hacia la evaluación.

P5. Informes claros y completos. Los informes en que se reflejan las evaluaciones deben ser claros y completos, incluyendo en sus conclusiones

tanto los aspectos positivos y favorables como los negativos y desfavorables, de forma que los primeros se conviertan en base de nuevas acciones y los segundos en elementos para las actuaciones de mejora.

Contra lo que pudiera parecer al analizar evaluaciones o escritos sobre el tema, las evaluaciones deben servir no sólo para descubrir y poner de relieve los aspectos negativos que, obviamente, deben ser identificados y delimitados de forma clara y precisa; también los aspectos positivos deben identificarse, y ello no sólo por el valor motivador que representan sino porque tales aspectos positivos pueden convertirse en un elemento de base para las decisiones de mejora.

En la misma línea, el evaluador debe poner de relieve los métodos que ha utilizado, con sus pros y sus contras, a fin de que pueda valorarse su incidencia tanto sobre los resultados como sobre las conclusiones del Informe.

P6. Derecho a saber. Evaluador y cliente deben asegurar que los resultados de la evaluación, con sus correspondientes limitaciones, serán accesibles a las personas o grupos afectados, algunas de las cuales tienen derecho a recibir los resultados.

El prestigio, la credibilidad y la propia utilidad de la evaluación dependen en gran medida de la habilidad de los evaluadores para identificar a las personas y grupos afectados e interesados y facilitarles el acceso u ofrecerles la información que les interesa o les afecta.

Ahora bien, dado que los clientes tienen también sus derechos en relación con los resultados de la evaluación, será preciso el establecimiento de un acuerdo formal en este punto.

Las partes interesadas y afectadas deben conocer por qué se realizó la evaluación y cómo se llevó a cabo, así como los métodos utilizados y las valoraciones y recomendaciones realizadas por el evaluador.

El sometimiento de estos aspectos al examen de las partes interesadas deviene a favor de una evaluación de mayor credibilidad y calidad.

P7. Conflicto de intereses. Los intereses en conflicto deben afrontarse de forma abierta y honesta, de modo que no interfieran en el proceso y en los resultados de la evaluación.

De los estándares anteriores se deduce con facilidad que pueden darse situaciones de conflicto entre los intereses de los clientes y de otras partes interesadas o afectadas; el conflicto puede plantearse, del mismo modo, entre los intereses económicos, profesionales o personales del evaluador y la propia evaluación, los clientes y las partes interesadas.

Resulta evidente que los evaluadores pueden estar mediatizados por sus intereses de distinto tipo –conseguir apoyos financieros futuros, mantener buenas relaciones profesionales y personales, afianzar su prestigio profesional...–; del mismo modo, determinadas partes interesadas, o el propio cliente, pueden llegar a sesgar la información, ofrecerla incompleta o insuficiente.

Tomando conciencia de que, en efecto, el conflicto de intereses puede darse, y es normal que se dé, el estándar plantea la necesidad de afrontarlo previamente y llegar a acuerdos que no desvirtúen, cuando no invaliden, las evaluaciones.

P8. Responsabilidad fiscal. Las evaluaciones deben realizarse de forma que se asegure el uso de procedimientos económicos, siendo prudentes y éticamente responsables en lo relativo a los gastos y honorarios, que deben ser adecuados.

El uso de los recursos económicos, tanto públicos como privados, debe responder a criterios de honradez y responsabilidad.

El incumplimiento del estándar, además de las responsabilidades a que pudiera dar lugar, termina por incidir en el descrédito de la evaluación.

6.3.4. Estándares de precisión

Este último criterio de calidad de las evaluaciones es el más amplio: nada menos que doce estándares o normas lo componen.

El núcleo de su contenido se centra en la necesidad de que la evaluación aporte una información técnicamente correcta en relación con aquellas notas o características que deciden sobre el valor y el mérito del programa evaluado.

A continuación se presentan los doce estándares:

A1. Documentación del programa. El programa evaluado debe describirse y documentarse de forma clara y adecuada a fin de que pueda conocerse con toda precisión y claridad.

La credibilidad de la evaluación descansa en gran medida en el hecho de que sus resultados estén respaldados por una completa documentación, que permita al evaluador una comprensión profunda del programa y ello no sólo en su formulación inicial sino en la forma real en que fue aplicado.

A2. Análisis del contexto. El contexto en que el programa se desarrolla debe ser analizado con la debida atención, con el fin de identificar el influjo que pueda representar para su desarrollo y resultados.

Los programas educativos no se desarrollan en el aire; de hecho se ponen en marcha en un contexto en el que hay personas con sus propios programas e intereses, se dan relaciones humanas, se cuenta con más o menos apoyos y recursos... Pensar que estas circunstancias y condicionantes no tienen algún tipo de reflejo e incidencia en el programa, en su puesta en marcha e implementación y hasta en sus resultados, no es realista.

En la medida en que la evaluación implica valoración e interpretación, ni el programa ni sus resultados deberían sustraerse a estos elementos condicionantes y ello no sólo para una valoración más ajustada sino, también, para tomar decisiones que, en ocasiones, no deberían afectar –o, al menos, no hacerlo sólo– al programa sino, precisamente a ese contexto de referencia.

Por otra parte, y de cara a la posible aplicación del programa en otros contextos y situaciones, o de aplicar las conclusiones a otros lugares, parece necesario estar al tanto de estos elementos condicionantes.

Por último, el estándar nos advierte sobre la conveniencia de aplicar los programas en el contexto más favorable.

A3. Descripción de intenciones y procedimientos. En toda evaluación deben describirse con el detalle necesario para su identificación y valoración tanto las intenciones a las que sirve como los procedimientos que se utilizarán.

La transparencia que debe adornar los procesos de evaluación exige que las intenciones que dan lugar a la evaluación, con-

cretadas en sus objetivos y en el uso que se pretende dar a los resultados, así como los procedimientos que se utilicen, entre los que cabe destacar las formas en que se recogen, organizan, analizan y utilizan los datos, deban ser descritos a fin de que puedan ser identificados y valorados.

Este estándar viene en parte exigido por el desacuerdo manifiesto existente entre la comunidad evaluadora tanto en relación con las intenciones como con los procedimientos. De esta forma, el destinatario de la evaluación conoce las claves de la misma y puede situar sus propias valoraciones en ese contexto de referencia, algo de gran importancia para conferir a la evaluación mayor o menor credibilidad y para utilizar o rechazar la información que le ofrece.

A4. Fuentes de información. A fin de que pueda ser valorada como adecuada o no la información ofrecida, es preciso que se describan con el debido detalle las fuentes de información utilizadas.

Las evaluaciones deben acudir a fuentes diferentes para disponer de la información necesaria para poder valorar los programas. La diversidad de fuentes es necesaria en ocasiones porque cada fuente puede aportar una información específica; pero en otras, esa diversidad es conveniente en la medida en que permite el contraste de puntos de vista de los diversos afectados e interesados en el programa.

Ante la diversidad de fuentes, el evaluador debe considerar las técnicas más adecuadas al caso; cabe pensar en cuestionarios de opinión, en encuestas, en entrevistas estructuradas o en las denominadas “en profundidad”, en situaciones de prueba, en grabaciones en cintas de audio y video, en someter a observación los comportamientos o en utilizar la documentación disponible, por ejemplo.

A los efectos oportunos, es necesario informar sobre las fuentes seleccionadas, los procesos de identificación y selección de informantes –evitando grupos interesados o sesgados en sus puntos de vista–, las técnicas utilizadas y la forma de analizar la información. Con tal información como base, resulta posible a las partes interesadas disponer de la información necesaria para valorar la propia evaluación y tomar las decisiones pertinentes.

A5. Información válida. *La información a recoger y los procedimientos a utilizar deben seleccionarse y aplicarse de forma que se asegure la validez de las interpretaciones para el uso al que se destine la evaluación.*

La clave de la calidad de las evaluaciones descansa en su validez. Las dificultades que debe superar el evaluador en este punto son muy elevadas, sobre todo cuando la evaluación hace referencia a información o datos difícilmente accesibles.

Al evaluador le incumbe recoger las evidencias precisas que respalden sus datos, instrumentos, procedimientos e interpretaciones.

Alcanzar la validez de la información implica en ocasiones, y como ya se ha señalado en el capítulo correspondiente, probar hipótesis que relacionan determinados hechos observables con ciertos constructos inobservables. Por tanto, debe haber un respaldo teórico para estas suposiciones.

Pero, junto a ello, es necesario disponer de evidencias en el sentido de que tanto las técnicas utilizadas para recoger la información así definida, como los instrumentos concretos en que toman cuerpo, ponen de manifiesto tales constructos en lugar de realidades diferentes⁴.

A6. Información fiable. *Se deben elegir y aplicar procedimientos para la recogida de la información capaces de asegurar que es suficientemente fiable para el uso al que se destina.*

En general, se entiende por fiabilidad el grado de consistencia de la información recogida. A tales efectos, es necesario distinguir entre la variabilidad de las puntuaciones explicable como consecuencia de las diferencias reales existentes entre las personas a las que se aplican los instrumentos de recogida de datos y aquella otra que se atribuye a los denominados errores de medida, explicable por causas fortuitas o por azar.

A los efectos de la fiabilidad, conviene tomar las medidas oportunas para minimizar los errores de medida a la vez que se potencia la variabilidad –varianza– verdadera.

⁴ El tema de la validez es clave para todos los instrumentos de recogida de datos. Determinados tests que pretenden medir inteligencia en realidad miden conocimientos e incluso memoria.

La fiabilidad puede tomar formas diferentes; en ocasiones lo que se pretende es conocer el grado de *estabilidad* de las puntuaciones de una a otra ocasión; en otras, se trata del grado de *equivalencia* de dos series de puntuaciones correspondientes a instrumentos que pretenden medir la misma realidad; en otras, lo importante es el grado de *consistencia interna*. Al evaluador le incumbe reducir tanto como pueda las limitaciones a la fiabilidad de su información, e informar sobre las mismas.

A7. Información sistemática. *La información recogida, organizada y analizada debe ser revisada de forma sistemática, de forma que los posibles errores puedan ser detectados y corregidos.*

Se debe ser consciente de que los riesgos de error son variados y elevados, afectando a momentos y etapas de la evaluación que van desde la recogida de datos a su análisis, pasando por los procesos de puntuación, codificación, o registro.

Se deben realizar los esfuerzos precisos para que las instrucciones que se den a las personas que intervengan en estos procesos sean claras; junto a ello, se deben tomar las medidas necesarias para que reciban la formación *ad hoc* a fin de que sean capaces de reconocer la información pertinente. Por último, se debe diseñar y aplicar un plan sistemático de controles y comprobaciones capaces de eliminar o, cuando menos, reducir al máximo los errores.

A8. Análisis de la información cuantitativa. *La información de tipo cuantitativo debe ser analizada de forma sistemática y apropiada a fin de que puedan darse las respuestas adecuadas a las preguntas formuladas.*

La información de naturaleza cuantitativa –edad, nivel socioeconómico, resultados, medidas de actitudes...– permite su organización, resumen y presentación adecuadas.

Las diversas preguntas planteadas al inicio de la evaluación deben llevar aparejada la selección de los métodos y procedimientos de análisis más adecuados; por lo general, conviene comenzar con un análisis exploratorio, capaz de identificar errores, de apreciar su corrección y exactitud y de permitir una primera aproximación a la realidad evaluada. Las representaciones gráficas permiten alcanzar una visión intuitiva de los datos, facilitando de este modo una mejor comprensión de los mismos.

Tras el análisis exploratorio se puede proceder a análisis más complejos, capaces de encontrar respuestas claras a las preguntas y de facilitar su interpretación. Entre tales análisis pueden destacarse los relativos a comparaciones entre grupos o a inferencias en relación con las diferencias encontradas.

Vale la pena destacar, de una parte, la necesidad de que los evaluadores comprueben si sus análisis son adecuados a los datos recogidos, esto es, si cumplen con los requisitos correspondientes; de otra, deben estar en condiciones de defender su metodología: los análisis, los resultados y las interpretaciones.

A9. Análisis de la información cualitativa. La información cualitativa de un programa debe ser analizada de forma sistemática y apropiada a fin de que puedan darse las respuestas adecuadas a las preguntas formuladas.

Se considera cualitativa aquel tipo de información consistente en narraciones y descripciones realizadas por diferentes personas a través de técnicas muy distintas: desde la entrevista –estructurada o no– a los documentos, pasando por las grabaciones o la observación.

Los análisis cualitativos consisten en el tratamiento adecuado de la información, con frecuencia mediante su organización en *categorías* que, por lo general, no suelen estar predeterminadas. En efecto, en este tipo de análisis es habitual el planteamiento de un proceso inductivo e interactivo mediante el cual el evaluador va y viene de los datos a su organización e interpretación, contrastándola con las diferentes fuentes productoras y con las partes interesadas en el programa.

Aunque no es fácil establecer cuándo se ha alcanzado un suficiente análisis de la información de esta naturaleza, se puede considerar que se ha llegado a ese punto cuando se han elaborado unas categorías que permiten poner de relieve la información necesaria para dar respuesta a las preguntas de la evaluación. Las categorías, por su parte, han debido ser analizadas en cuanto a su validez y fiabilidad.

A10. Conclusiones justificadas. Las conclusiones a las que llega la evaluación deben justificarse explícitamente a fin de que las diferentes partes implicadas y afectadas puedan someterlas a valoración.

En general, la evaluación, plasmada en el Informe final, acaba en una serie de conclusiones en las que se concretan los juicios, valoraciones y recomendaciones del evaluador.

A fin de lograr por parte de los destinatarios de la información la credibilidad y aceptación necesarias, tales conclusiones deben estar debidamente argumentadas. A tales efectos, es importante que se aprecie que derivan lógicamente de la información recogida; en suma, las conclusiones, como ocurre en los informes de investigación, deben ir acompañadas de la consiguiente discusión.

A11. Objetividad del informe. Los informes deben reflejar con claridad los resultados de la evaluación, evitándose toda distorsión debida a las posiciones personales de cualesquiera de las partes interesadas (audiencias).

Los informes pueden presentar distorsiones tanto conscientes e intencionadas, lo que atenta contra los principios éticos y las normas legales, como inconscientes y no intencionadas, fruto de posiciones y sesgos personales del evaluador.

Los evaluadores deben estar atentos a ambos tipos de sesgos; en particular, deben evitarse aquellos que pueden ser fruto de la falta de cuidado o de la presión de las partes interesadas, incluidos los clientes.

A12. Metaevaluación. La evaluación en sí misma debe ser evaluada, tanto formativa como sumativamente, poniéndola en relación con los anteriores estándares. De esta forma, el proceso evaluativo podrá ser llevado adelante de forma adecuada y las diferentes audiencias podrán examinar su realización.

Si la evaluación tiene un potencial de mejora se debe fundamentalmente a sus aportaciones, pero estas lo son en la medida en que tanto su diseño, como su realización y el propio informe se lleven a cabo de forma correcta y adecuada, algo muy difícil como se desprende de las exigencias de los estándares anteriores. La consecuencia parece lógica: será preciso evaluar la propia evaluación y, para ello, los estándares se convierten en elemento de referencia.

La evaluación de la evaluación puede y debe llevarse a cabo por parte de los propios evaluadores, de evaluadores externos, de los clientes y de las demás partes interesadas. La realización siste-

mática de este tipo de actividad metaevaluativa redundante en la credibilidad de las evaluaciones, de los evaluadores y de la evaluación como profesión.

La evaluación formativa de la evaluación se convierte en una guía del propio plan de evaluación y de su realización; al llevarse a cabo durante el proceso de evaluación es posible corregir sobre la marcha las disfunciones detectadas. Cuando la evaluación la realiza un equipo de personas, alguna de ellas puede asumir esta tarea.

Por su parte, la evaluación sumativa de la evaluación se da al final del proceso evaluador sobre el programa completo. En general, es preferible que este tipo de metaevaluaciones pueda ser llevado a cabo por parte de evaluadores externos independientes.

Cuestiones importantes de esta evaluación sumativa pueden ser las siguientes:

- ¿Fue adecuado el diseño de evaluación?
- ¿Se llevó a cabo de forma correcta la aplicación del diseño evaluativo?
- ¿Están claramente definidos los resultados de la evaluación?
- Las conclusiones, ¿están claramente respaldadas por la información y los datos?
- ¿Se difundió adecuadamente la información alcanzada?

Atendiendo a lo anterior, se comprenderá el importante papel que juega para realizar la metaevaluación el hecho de haber ido documentando adecuadamente las diversas decisiones y actuaciones que ha ido llevando a cabo el evaluador.

6.4. ALGUNAS REFLEXIONES SOBRE LOS ESTÁNDARES

Un análisis de los estándares nos permite comprobar su concentración en la *evaluación*, sin referencias normativas a los *programas*, además de la ya indicada referencia a un contexto diferente del nuestro, lo que lleva a que algunos de los estándares sean de difícil aplicación en la situación ordinaria de nuestro país. Junto a ello, señalaré lo que considero ausencias y haré mención de ciertas discrepancias en cuanto a la organización de algunos estándares. Veamos.

6.4.1. Separación entre programa y su evaluación

Es comprensible que se planteen unos estándares sobre evaluación y no sobre los programas objeto de atención, por varias razones: estamos ante una actividad técnica, en la que puede darse más fácilmente el acuerdo; la evaluación es concebida como una actividad profesional y, por tanto, ajena a las personas responsables de los programas. Pero no es menos verdad que eso tiene dos importantes consecuencias: un cierto formalismo en las propuestas, por un lado, y posibles disfunciones por otro⁵.

Entiendo que debería plantearse un bloque de estándares sobre *programas* y, en nuestro caso, sobre *programas educativos*. A modo de propuesta, podrían ser, al menos, los siguientes:

- *Calidad intrínseca del programa:*

El documento en que tome cuerpo el programa debe concretar sus metas y objetivos, dando cuenta o razones de su elección en función de la relevancia y pertinencia de las mismas; en la medida de lo posible, ofrecerá información sobre la base teórica que lo sustenta.

- *Calidad técnica del programa:*

El programa debe reunir las condiciones técnicas para ser considerado tal. Entre ellas, la de especificar los objetivos; concretar la naturaleza de la intervención, debidamente fundamentada; plantear los medios y recursos e incluir un sistema de comprobación y mejora de sus resultados.

- *Calidad metodológica del programa:*

El programa debe formularse de forma que contenga información relevante para su correcta implantación e implementación, y, en

⁵ Un caso patente es el relativo a la *evaluación de la calidad de las universidades* sin previamente plantear algún tipo de orientación, pauta o propuesta sobre lo que es *calidad*, que queda al arbitrio de la *autonomía universitaria*. Una de dos, o ambos conceptos van unidos o no será posible establecer pautas comunes comparables... a no ser que consideremos que la calidad se encuentra subsumida de modo más o menos subrepticio en los instrumentos elaborados para la evaluación.

el marco de nuestras actuales reflexiones, para que resulte evaluable y, consecuentemente, mejorable.

6.4.2. Definición de estándares en un contexto diferente

Algunos de los estándares –en concreto F3, P2, P6, P8– están pensados especialmente para evaluadores profesionales, contratados como expertos, situación poco común en nuestro país, donde la mayoría de las evaluaciones es cubierta por universitarios con formación metodológica.

6.4.3. Ausencias y propuesta de modificaciones

Algún estándar puede considerarse innecesario si se aceptan los planteamientos sobre la conveniencia, y hasta la exigencia, de *complementariedad metodológica*; en concreto, los denominados A8 –Análisis de la información cuantitativa– y A9 –Análisis de la información cualitativa– bien podrían quedar refundidos en uno, con la denominación o rótulo de *Análisis de la información*, sin adjetivaciones.

El último de los estándares, el A12. *Metaevaluación*, no parecería ser propio de cualquier evaluación de programas, sino una exigencia técnica para la organización, sistematización y mejora de nuevas evaluaciones y de los propios programas educativos. La *metaevaluación* podría ser entendida como una exigencia técnica que permite a los diseñadores de programas y a los evaluadores, realizar mejor sus tareas de diseño y de planificación de la evaluación. En último caso, y cuando se refiere al propio programa, no es otra cosa que un análisis autocrítico que debería realizarse en paralelo al propio proceso completo de la evaluación para evitar situaciones como las de llegar al final habiendo cometido errores que invalidan todo el esfuerzo.

Por otra parte, se echa de menos una referencia genérica a la *metodología* ya que los estándares que aparecen sobre este ámbito se limitan a los *procedimientos* –A3– o se centran en exclusiva sobre la *información* (A4: Fuentes; A5, 6 y 7: información válida, fiable y sistemática; A8 y A9, como ya se ha indicado, sobre su análisis).

El estándar sobre *metodología*, a mi juicio, debería referirse a la necesidad de que fuera respetuosa con el *principio de complementariedad*, al que hemos aludido reiteradamente, siempre que se trate de evaluaciones integrales de programas.

BIBLIOGRAFÍA

Se incluye a continuación una selección de obras sobre la temática de esta primera parte, quedando la bibliografía general para el final de la monografía.

- Comité Conjunto de Estándares para la Evaluación Educativa (1998). *Estándares para la evaluación de programas*. Bilbao: Ediciones Mensajero.
- Fernández Ballesteros, R. (Coord.) (1996). *Evaluación de Programas*. Madrid: Síntesis.
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*. Beverly Hills: Sage Pub.
- Joint Committee on Standards for Educational Evaluation (1981). *Standards for Evaluations of Educational Programs, Projects and Materials*, New York: McGraw-Hill.
- Pérez Juste, R. (Coord.) (2000). Evaluación de programas educativos. En *Revista de Investigación Educativa*. Número monográfico. Vol. 18, 2.
- Rossi, P.H. y Freeman, H.E. (1989). *Evaluación: Un enfoque sistemático para programas sociales*, México, Trillas.
- Stenhouse, L. (1984). *Investigación y desarrollo del currículum*, Madrid: Morata.
- Stufflebeam, D.L. y Shinkfield, J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Buenos Aires: Paidós.