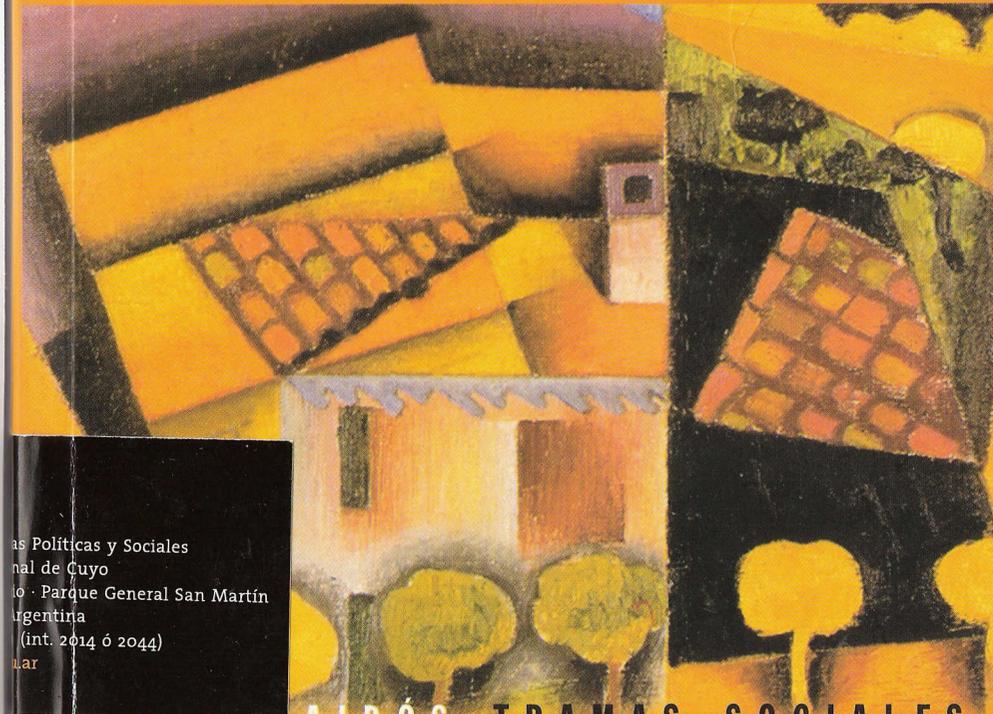


# EVALUAR PARA LA TRANSFORMACIÓN

Innovaciones en la evaluación  
de programas y proyectos sociales

Olga Nirenberg  
Josette Brawerman  
Violeta Ruiz



As Políticas y Sociales  
nal de Cuyo  
lo · Parque General San Martín  
Argentina  
(int. 2014 ó 2044)  
uar

PAIDÓS TRAMAS SOCIALES

**EVALUAR PARA LA TRANSFORMACIÓN**

## TRAMAS SOCIALES

Directora de colección:

*Lic. Irene Gojman*

20. S. Checa (comp.)  
*Género, sexualidad y derechos reproductivos en la adolescencia*
21. T. Puig  
*La comunicación municipal cómplice con los ciudadanos*
22. M. Caracciolo Basco y M. Laxalde  
*Economía solidaria y capital social*
23. M. Montero  
*Introducción a la psicología comunitaria*
24. H. Lejarraga (comp.)  
*Desarrollo del niño en contexto*
25. G. Kessler  
*Delito, exclusión y juventud*
26. V. Ruiz  
*Las organizaciones comunitarias y gestión asociada*
27. M. Mucci  
*Psicoprofilaxis quirúrgica*
28. T. Puig  
*Se acabó la diversión*
29. F. Elías  
*La adopción de niños como cuestión social*

Si desea recibir información mensual de nuestras novedades/publicaciones, y ser incorporado a nuestra lista de correo electrónico, por favor envíenos los siguientes datos a [difusión@editorialpaidos.com.ar](mailto:difusión@editorialpaidos.com.ar)  
Nombre y apellido, profesión y dirección de e-mail.

# EVALUAR PARA LA TRANSFORMACIÓN

## Innovaciones en la evaluación de programas y proyectos sociales

OLGA NIRENBERG  
JOSETTE BRAWERMAN  
VIOLETA RUIZ



PAIDÓS  
Buenos Aires  
Barcelona  
México

Cubierta de Gustavo Macri

Motivo de cubierta: *Fragmento de Paisaje de Cèret (1913)* de Juan Gris

302.3 Nirenberg, Olga  
NIR Evaluar para la transformación : innovaciones en la  
evaluación de programas y programas sociales / Olga  
Nirenberg, Josette Brawerman y Violeta Ruiz.- 1ª ed.  
2ª reimp.- Buenos Aires : Paidós, 2005.  
224 p. ; 21x13 cm.- (Tramas sociales)

ISBN 950-12-4508-X

I. Brawerman, Josette II. Ruiz, Violeta III. Título -  
1. Interacción Social Intergrupar

1ª edición, 2000

1ª reimpresión, 2003

2ª reimpresión, 2005

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del *copyright*, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© 2000 de todas las ediciones en castellano  
Editorial Paidós SAICF  
Defensa 599, Buenos Aires  
e-mail: [literaria@editorialpaidos.com.ar](mailto:literaria@editorialpaidos.com.ar)  
[www.paidosargentina.com.ar](http://www.paidosargentina.com.ar)

Queda hecho el depósito que previene la Ley 11.723  
Impreso en la Argentina - Printed in Argentina

Impreso en Gráfica MPS,  
Santiago del Estero 338, Lanús, en febrero de 2005  
Tirada: 1000 ejemplares

ISBN 950-12-4508-X

### **Olga Nirenberg**

Es socióloga (Universidad Nacional de Buenos Aires -UBA-), con posgrados en Planificación Social y en Salud Pública. Es directora ejecutiva del Centro de Apoyo al Desarrollo Local (CEADEL) y coordinadora del área de evaluación del Programa de Atención a Niños y Adolescentes en Riesgo (ProAme, Secretaría de Desarrollo Social/BID). Trabaja como consultora en evaluación para otros programas nacionales de la Secretaría de Desarrollo Social y en el Programa Materno Infantil y Nutrición (PROMIN, del Ministerio de Salud/BIRF), y para programas internacionales, en especial el Programa Regional de Salud Integral del Adolescente de la OPS-Washington DC y el Programa UNI de la Fundación W. K. Kellogg.

### **Josette Brawerman**

Es socióloga (UBA), con estudios de posgrado en Planificación Social. Ha sido consultora del Centro Interamericano de Desarrollo Social de la OEA y de UNICEF

*Olga Nirenberg, Josette Brawerman, Violeta Ruiz*

Argentina. En la actualidad es miembro del equipo técnico de CEADEL y consultora del área de evaluación del Programa Materno Infantil y Nutrición (PROMIN, Ministerio de Salud/BIRF) y del Programa de Atención a Niños y Adolescentes en Riesgo (ProAme, Secretaría de Desarrollo Social/BID).

### **Violeta A. Ruiz**

Es socióloga (UBA), con estudios en Programación Social. En 1997 recibió el Premio Adebá por el trabajo presentado sobre la evaluación del Programa Materno Infantil del Ministerio de Salud. Se desempeñó como consultora en programas sociales dependientes de distintas áreas de gobierno y Organizaciones de la Sociedad Civil (OSC). Ha coordinado y colaborado en evaluaciones de programas para la Dirección Nacional de Gasto Público Social del Ministerio de Economía, CEADEL, UNICEF Argentina y programas dependientes de la Secretaría de Desarrollo Social de la Nación.

## **ÍNDICE**

<b>Prólogo, Leopoldo Halperín</b> .....	13
<b>Introducción</b> .....	21
<b>Capítulo 1: ¿Para qué la evaluación?</b> .....	27
Introduciéndonos en tema .....	27
La evaluación como momento o función de la gerencia .....	29
Intentando una definición .....	31
Evaluación e investigación: bastantes similitudes y algunas diferencias .....	36
Cambios en los estilos de evaluación .....	38
La evaluación para la transformación .....	42
Aprendiendo con la evaluación y de ella .....	46
El proceso de la evaluación .....	48
<b>Capítulo 2: Los tipos de evaluación</b> .....	51
Antes, durante o después: siempre hay que evaluar .....	53
Conjugando el verbo evaluar: ¿nos evaluamos o nos evalúan? .....	58
Los múltiples propósitos de la evaluación .....	62

Énfasis según temas: ¿qué evaluamos? .....	65
Los distintos objetos posibles de evaluación:	
¿tipos o niveles de análisis? .....	72
Para evaluar las evaluaciones .....	75
La multidimensionalidad de la evaluación .....	77
<b>Capítulo 3: Los abordajes metodológicos</b> .....	79
Evaluación cuantitativa y cualitativa.	
Un debate histórico.....	80
¿Evaluación cuantitativa o cualitativa?	
Ésa no es la cuestión.....	85
Una cuestión central: el rigor metodológico .....	91
Actores y audiencias .....	96
<b>Capítulo 4: Las técnicas de evaluación</b> .....	105
Acercándonos a las técnicas .....	105
Los diferentes tipos de técnicas .....	107
¿Cómo decidir qué técnicas utilizar? .....	129
¿Y sobre los instrumentos? .....	130
<b>Capítulo 5: El diseño de un modelo evaluativo</b> .....	133
¿Por qué un modelo? .....	133
Los atributos deseables de un modelo evaluativo .....	135
Las partes que integran un modelo evaluativo .....	138
Parte por parte .....	138
Las dimensiones del modelo evaluativo.....	144
Abriendo el abanico de las dimensiones .....	148
Seleccionando y definiendo indicadores .....	154
Las técnicas y las fuentes .....	161
Las preguntas orientadoras .....	162
La matriz síntesis del modelo evaluativo .....	164
El instrumento valorativo:	
un modo de cuantificar lo cualitativo.....	165

<b>Capítulo 6: La aplicación del modelo evaluativo y la diseminación de los resultados</b> .....	169
Transferencia del modelo evaluativo y ajustes .....	169
Los énfasis o focos en los distintos momentos.	
Inclusión de ajustes .....	170
La participación de los actores .....	173
La formulación de juicios valorativos fundamentados y de recomendaciones .....	177
La devolución de los hallazgos evaluativos .....	181
El informe de evaluación .....	183
Talleres para el acuerdo con los actores .....	188
El resumen ejecutivo .....	189
La difusión de los hallazgos.....	190
<b>Anexos</b> .....	191
Matriz síntesis de un modelo evaluativo .....	192
Instrumento de valoración.....	205
<b>Bibliografía</b> .....	217

El título del libro *Evaluar para la transformación* adelanta la intención de las autoras, fruto de su aquilatada experiencia en la formulación y evaluación de programas sociales. Y esta intención se resume, a nuestro juicio, en la propuesta de una reflexión profunda del significado de la evaluación como elemento clave en la programación de las políticas sociales.

Es frecuente, en los análisis de esta problemática, confundir la actividad de evaluación –con los supuestos de racionalidad implícitos en ella– con la de programación. Atribuir a la programación –entendida como la asignación de recursos *ex ante*– un carácter racional en todos los casos induce a error. La programación, como metodología de asignación de cualquier tipo de recurso en la sociedad, puede ser racional o irracional tanto en objetivos como en métodos, y es de suma importancia tener en claro este punto al abordar este tema, especialmente en la evaluación de programas.

En este trabajo, Nirenberg, Brawerman y Ruiz hacen un minucioso análisis del significado de la evaluación en

la programación, partiendo del entendimiento claro y simple (pero pocas veces incorporado en los textos) de que la evaluación no es una parte inicial o final del proceso de creación u operación de un proyecto o programa social, sino que se trata de un elemento constitutivo e inseparable del mismo. Es el núcleo motor de la racionalidad de la que debiera estar dotado todo proyecto destinado a modificar una realidad social, desde el comienzo hasta más allá de su finalización, desde las ideas iniciales que guían al programador en su diseño hasta el momento en que, cumplidos los objetivos, es necesario reflexionar tanto sobre los efectos de lo propuesto y actuado como sobre las consecuencias –deseadas y no deseadas– de los procesos que se han puesto en marcha.

En este sentido, en todo el texto hay un explícito rechazo a pensar la evaluación como un momento aislado, parcial y confinado en el tiempo, reafirmando el papel del “aprendizaje” que ésta brinda como punto de partida y de llegada, de prueba abierta y permanente base de nuevas propuestas y gestiones superadoras en la política y en la gestión social.

Un análisis apresurado podría llevar a confundir esta concepción de la evaluación con la racionalidad que Weber<sup>1</sup> define como propia de la modernidad y del capitalismo, racionalidad instrumental u operativa en la que fines, medios y condiciones son analizados por el actor y esto le permite elegir las conductas que lo llevarán, con los medios disponibles, a alcanzar los fines “racionalmente seleccionados”. A su vez, Mayntz,<sup>2</sup> profundizando

1. Weber, M.: *Economía y sociedad*, México, Fondo de Cultura Económica, 1964.

2. Mayntz, R.: *Sociología de la administración pública*, Madrid, Alianza, 1994.

esta caracterización, en su análisis del sistema (modelo) político-administrativo moderno, redefine el problema de la racionalidad instrumental en la gestión del subsistema administrativo, mostrando las variaciones que registra ésta de acuerdo con la ampliación y evolución de sus funciones. En la etapa constitutiva del Estado moderno, tanto en el absolutismo fundante como en el período del Estado gendarme del liberalismo, el predominio de funciones como las de recaudación, orden externo e interno hacen que domine una racionalidad burocrática, cuyo criterio de éxito se basa en el apego a la norma, independientemente de los efectos que ella tenga.

La propia evolución del Estado, en el posterior período democrático de participación, se acompaña de nuevas funciones (obras públicas, desarrollo económico, políticas de bienestar social) y se corresponde con un nuevo modelo de gestión, el de la planificación, cuyos criterios de éxito (y racionalidad) se basan en el cumplimiento de “metas” que sustituyen la racionalidad burocrática. De este modo, para las autoras, coexisten en el seno del Estado dos tipos de racionalidad instrumental, con distintos argumentos.

Pese a esta distinción que surge de los paradigmas normas-metas, la base de racionalidad sigue teniendo características parciales, de tipo instrumental. Así concebida, la racionalidad del planificador, aunque distinta de la del burócrata, se detiene en el momento en que alcanzó el objetivo propuesto en el plan o proyecto. La exigencia –como afirmamos– es llegar a la meta, cumplir acabadamente con los fines del plan. En ambos casos, la racionalidad aparece como un elemento fragmentado, dada su definición instrumental, es decir, parcial.

Esta concepción limitada se contrapone con la racionalidad asociada a la evaluación, que constituye el modelo de la propuesta contenida en este libro y que R. Gómez<sup>3</sup> (utilizando los criterios epistemológicos de K. O. Apel) caracteriza como comprensiva, en tanto discute o elucida la propia racionalidad (y justeza) de los fines que se persiguen.

La mera racionalidad instrumental, restringida y fragmentaria, está irremediablemente ligada a una concepción de la sociedad y su organización, y elimina la reflexión sobre las consecuencias de la acción, con un notorio empobrecimiento y limitación en el análisis, tanto a nivel epistemológico como de la propia concepción de la sociedad, de su organización, de los “modos aceptados como racionales y para la acción social misma”.

La propuesta evaluativa que se aborda en este texto –en tanto elaborada como proceso que abarca y contiene todo el desarrollo del plan o proyecto, sometiendo a crítica sistemática no sólo los fines seleccionados y el ajuste de la operación a éstos sino además, y fundamentalmente, los efectos deseados y no deseados de la implementación, y aun el propio ejercicio de la evaluación misma, mediante la “metaevaluación”–, supera los límites de la razón instrumental y de la crítica expuesta.

Un elemento destacable de esta propuesta es el de la profunda asociación que se establece en el texto entre evaluación y método científico, relación estrechamente ligada a la concepción de racionalidad anteriormente expuesta. Sin embargo, esta asociación está lejos de identificar ambas actividades, e ignora sus diferencias. Recuperando la

3. Gómez, R.: *Neoliberalismo y pseudociencia*, Buenos Aires, Lugar, 1995.

distinción que Bunge<sup>4</sup> hace entre ciencia y tecnología, la programación-evaluación aparece firmemente encuadrada en los cánones del método científico, respondiendo a sus exigencias de prueba y verificación, pero con identidad propia, en tanto acción práctica dirigida a la transformación, a la operación sobre la realidad social.

Un aspecto sobresaliente del trabajo reside en el esfuerzo didáctico con que se ha asumido la tarea a lo largo de los seis capítulos que componen el libro.

Desde el análisis minucioso de la propia definición de evaluación y sus implicancias, sus vinculaciones con el método científico, los criterios de validación y operacionalización, el análisis de las técnicas hasta la matriz de evaluación, la metodología expositiva elegida privilegia en todos los temas tanto la desmitificación de cuestiones aparentemente irreductibles (como es, por ejemplo, la aparente antinomia entre lo cuantitativo vs. lo cualitativo en investigación social), como las definiciones sencillas y accesibles, tanto para quien se asoma por primera vez a estas cuestiones como para los especialistas interesados en otras perspectivas analíticas.

El tratamiento de estos problemas resulta relevante en nuestro país, particularmente en el ámbito de las políticas sociales, dado que éstas muestran –con escasas excepciones– una generalizada orfandad tanto en la evaluación como en la formulación y gestión de programas y proyectos.

Resulta doblemente paradójico que esto ocurra en la Argentina, a la luz de su temprana vocación en la formulación de programas sociales y de salud a partir de la década del cuarenta, ya que al mismo tiempo, y en sentido

4. Bunge, M.: *¿Qué es la ciencia?*, Buenos Aires, Siglo XXI, 1960.

contrario, aun con distintas definiciones, en el resto de América Latina, países como Chile, Venezuela y Cuba, entre otros, han recurrido al modelo de la planificación como método de enfocar su problemática social.

Seguramente son múltiples los factores que explican la renuencia del sector público a incorporar la programación (y la evaluación) con estos criterios de racionalidad en la gestión de los programas sociales. Entre ellos, constituye un elemento de peso la prevalencia en el subsistema administrativo del Estado argentino de una "cultura excesivamente burocrática", que se ha convertido en la orientación cuasi universal de toda gestión de Estado, trascendiendo los periódicos cambios de gobierno e, incluso, la alternancia entre modalidades democráticas y autoritarias que se sucedieron en los últimos cuarenta años.

De este modo, en la formulación y ejecución de la política social, aunque las acciones y los propósitos se etiqueten con frecuencia con el nombre de "programas" y recurran en no pocos casos a la liturgia formal de la propia programación, sus criterios de ejecución, racionalidad y éxito se atienen a los cánones estrictos del procedimiento contable, siendo el principal, y a veces el único criterio de ejecución, su apego al reglamento y a las normas de contabilidad pública.

Esta característica negativa "administrativista" que tiene la política social argentina no puede desvincularse del propio accionar del subsistema político en el marco del Estado. Esto se debe, a nuestro juicio, en parte, a una falencia generalizada de dicho subsistema (el político), el cual, además de carecer de información y condiciones técnicas que le permitan limitar y controlar eficazmente el poder real del subsistema administrativo, se ha volca-

do crecientemente a utilizar las políticas sociales con un sentido fuertemente clientelístico, y para cuyo ejercicio la programación y la evaluación resultan superfluas. Igual consideración merece, a nuestro juicio, la gestión en las políticas sociales de un número importante de ONG, que han pasado a formar parte subsidiaria de la cadena mencionada de acciones clientelares.

El conjunto de estas consideraciones nos permite afirmar que sin duda la lectura de *Evaluar para la transformación* es ineludible, sobre todo para quienes asuman la tarea de la planificación social, para todos los que desean resolver las cuestiones que preocupan a Alicia (y no sólo a ella) acerca de dónde está parada y cómo llegar al fin del camino.

LEOPOLDO HALPERÍN  
agosto de 1999

## INTRODUCCIÓN

---

*La imagen que nosotros construimos de la sociedad, la construimos para los ingenieros hidráulicos, para los jardineros... y para los revolucionarios. Todos ellos son invitados a nuestro teatro y les pedimos que no olviden su interés en la representación cuando estén con nosotros, porque queremos volcar el mundo sobre sus cerebros y sus corazones, para que puedan cambiarlo de acuerdo con sus deseos.*

BERTOLT BRECHT<sup>1</sup>

En este libro nos dirigimos fundamentalmente a aquellos que desarrollan acciones sociales, ya sea que formulen o que ejecuten programas o proyectos, tanto desde organizaciones de la sociedad civil como desde distintas áreas de los diferentes niveles gubernamentales. Son ellos o ellas quienes se habrán visto más de una vez en la urgente necesidad de evaluar sus acciones y las de otros, pero desprovistos de las concepciones y herramientas adecuadas para tal responsabilidad. Creemos que también podrá ser de utilidad para agencias de cooperación o filantrópicas que apoyan proyectos y programas brindando financiamiento, asesoramiento y/o capacitación para temas de desarrollo social.

Uno de nuestros propósitos principales es el de despegar el sentido de la evaluación del puro control que ejercen unos actores sobre otros, para situarlo en el pla-

1. Brecht, B.: *Kleines Organon für das Theater; Schriften zum Theater*, Francfort, Suhrkamp, 1964, pág. 20.

no del aprendizaje compartido entre quienes participan de las gestiones de los programas o proyectos, de modo que sean capaces de provocar las transformaciones necesarias en las situaciones insatisfactorias y en los modos de actuar para su resolución.

Cuando decimos que nos importa la evaluación para la transformación, hacemos referencia al menos a dos niveles en los que las transformaciones deberían producirse. Uno es el de contribuir a modificar las situaciones de adversidad, vulnerabilidad o inequidad en que se encuentran ciertos grupos poblacionales; como aclararemos posteriormente, todo proyecto o programa social se formula para cambiar situaciones de ese tipo, y la evaluación debe mostrar en qué medida se produjo el cambio deseable y debido a qué factores. Otro es mostrar de qué manera las acciones desplegadas para producir esas transformaciones resultan o no las adecuadas, y qué cambios se aconsejan introducir en la forma de hacer las cosas. De modo que principalmente estamos hablando de evaluación para la acción transformadora.

Compartimos el pensar de Krishnamurti (1998) acerca de que no podrá existir una transformación del "orden" internacional -del mundo- si no hay transformaciones en el ser humano, en su propio mundo interior, en sus concepciones, en sus formas de mirar lo que lo rodea y, sobre todo, en las formas de vincularse con los demás humanos y con el ambiente. En tal sentido, puede hablarse también de la existencia de un pensamiento transformador.

Si bien a lo largo del texto reiteramos que la evaluación, considerada como una tecnología social, consiste básicamente en transformar información en nuevo cono-

cimiento, entendemos ese mecanismo como instrumental para la acción orientada a transformar las situaciones de grupos humanos en condiciones desventajosas. Es por eso que, pese a las similitudes, la evaluación se distingue de la investigación social.

No intentaremos aquí discutir otros enfoques ni criticar otras formas de evaluar; más bien tomamos los aportes que cada una de esas modalidades puede brindar, incluyéndolas en un esquema abarcativo. Sí intentaremos aclarar algunas confusiones provenientes tanto de distintos usos lingüísticos para nombrar las mismas cosas como de iguales términos para nombrar cosas distintas, como puede verse sobre todo en los capítulos 2 y 3, donde discutimos los diferentes tipos y abordajes de evaluación y desmitificamos algunas de las distinciones tradicionales.

La mirada evaluativa que aquí proponemos es integradora, comprehensiva, con momentos de detalle y momentos de síntesis gestáltica.

Este libro no sirve como manual o receta para evaluar de forma universal todo tipo de proyecto o programa; en tal sentido no desarrolla instrumentos sino que procura transmitir la concepción de cuáles son los procesos requeridos y las orientaciones acerca de las formas de encararlos, para diseñar evaluaciones distintas para los diferentes proyectos o programas donde los profesionales del campo social solemos hacer cada vez una nueva experiencia.

Para cumplir con nuestro propósito dividimos el texto en seis capítulos y dos anexos.

El primer capítulo pretende dejar aclarada nuestra concepción general acerca de la evaluación y su utilidad

como función de la gerencia, es decir, incluida en la gestión social. El segundo capítulo hace una recorrida por las tipologías usuales de la evaluación en la literatura consultada y, como ya advertimos, se abordan algunas discusiones aclaratorias al respecto. En el tercer capítulo hablamos de los abordajes, de los falsos dilemas entre las cifras y las palabras –lo “cuanti” y lo “cuali”– que envolvieron durante años la cuestión de la evaluación y que ahora parecieran ir superándose, aunque no del todo, concluyendo que es el rigor metodológico lo que importa en las evaluaciones y que es sólo mediante la subjetividad de muchos diferentes que se puede resolver la subjetividad individual, aislada e idéntica a sí misma. El cuarto capítulo presenta una reseña de las técnicas a las cuales puede recurrir la evaluación para generar o recopilar y analizar información, sin pretender con eso tratar la cuestión de forma exhaustiva, pues es tal el caudal de técnicas existentes que se requeriría un libro sólo para ese desarrollo; sin embargo, hemos incluido algunas que nos parecieron novedosas y también discutimos algunas otras que últimamente se levantaron al nivel de “métodos”, pero que continúan siendo herramientas simples y útiles a disposición de los evaluadores (como las relacionadas con “evaluaciones rápidas” o de *rapid assessment*).

Finalmente, los capítulos quinto y sexto forman un solo bloque pero los hemos dividido para dar un respiro a los eventuales lectores. En el quinto se describen las características y los pasos del proceso de diseño de un modelo evaluativo, y en el sexto recorreremos las fases de su aplicación y de la diseminación de los hallazgos, las conclusiones y las recomendaciones producto de la evaluación, cuyo meollo es el informe evaluativo. En los anexos

se incluyen ejemplos que ilustran los pasos descritos en el capítulo quinto.

Las autoras tuvimos la fortuna de trabajar en diferentes escenarios de las políticas sociales durante nuestra vida profesional, que ya lleva muchos años, más de los que nos gustaría numerar. Dedicamos a la evaluación gran parte de los últimos diez o quince años. Por un lado, trabajamos en el marco de programas dependientes del Estado nacional, pero con ejecución en los niveles locales, en algunos casos con financiamiento bilateral o bien del presupuesto nacional exclusivamente; también como miembros de una organización no gubernamental para el desarrollo social; asimismo, como consultoras en el marco de fundaciones o agencias de cooperación internacional para el desarrollo. De modo que estamos muy agradecidas a la gente de las instituciones con las que tuvimos la oportunidad de tan vastos y diferentes aprendizajes.<sup>2</sup> Fue en el marco de esas instituciones que pudimos relacionarnos con innumerables funcionarios, profesionales, técnicos, miembros de organizaciones locales y vecinos y vecinas, quienes nos abrieron y mostraron generosamente

2. Algunas de esas instituciones son: el Centro de Apoyo al Desarrollo Local –CEADEL–, del cual las tres autoras somos miembros desde 1987; el International Development Research Center –IDRC–, de Ottawa, Canadá; el Programa UNI de la Fundación W. K. Kellogg; el Programa Regional de Adolescencia, de la Oficina Panamericana de la Salud –OPS– de Washington DC; UNICEF Argentina; el Centro Interamericano de Desarrollo Social –CIDES/OEA–; el Programa de Atención a Niños y Adolescentes en Riesgo –ProAme– de la Secretaría de Desarrollo Social y el Banco Interamericano de Desarrollo; el Programa Alimentario y Nutricional Infantil –PRANI– de la Secretaría de Desarrollo Social; la Dirección Nacional de Gasto Público Social, del Ministerio de Economía; el Programa Materno Infantil y de Nutrición –PROMIN– del Ministerio de Salud.

*Olga Nirenberg, Josette Brawerman, Violeta Ruiz*

te sus ámbitos de trabajo, para que juntos pudiéramos aprender a partir de la práctica. Es a esos amigos y amigas, compatriotas o de otros países del continente, a quienes va nuestro especial reconocimiento y nuestro deseo de que este libro les resulte útil en la hermosa pero sacrificada tarea que cotidianamente llevan a cabo.

Queremos también expresar nuestro agradecimiento, por su colaboración crítica e inteligente, a dos respetados amigos, que hicieron el esfuerzo de leer nuestros borradores; ellos escribieron con lápiz, de forma sintética y en los márgenes, sus observaciones, las que luego debimos redactar, tipear y colocar en el centro del texto, por su relevancia. Nuestro reconocimiento entonces a Néstor Perrone y a Pablo Vinocur, por sus acertadas e importantes contribuciones. Ojalá hayamos podido interpretar sus comentarios y expresarlos adecuadamente; pero es nuestra absoluta responsabilidad si, aun a pesar de ellos, subsisten los errores u omisiones.

Ahora llegó el momento en que la crítica debe ser puesta en las mentes y las voces de nuestros lectores; deseáramos que con ellos se constituya un nuevo diálogo, que permita hacer continuo el flujo de innovaciones en los modos de realizar y evaluar la acción social.

## CAPÍTULO 1

### ¿Para qué la evaluación?

*¿Me podrías indicar hacia dónde tengo que ir desde aquí?, preguntó Alicia.*

*Eso depende de adónde quieras llegar, contestó el Gato. A mí no me importa demasiado adónde..., empezó a decir Alicia.*

*En ese caso da igual adónde vayas, interrumpió el Gato. ... siempre que llegues a alguna parte, terminó Alicia, a modo de explicación.*

*¡Oh!, siempre llegarás a alguna parte, dijo el Gato, si caminas lo bastante.*

LEWIS CARROLL<sup>1</sup>

### Introduciéndonos en tema

El famoso diálogo entre la ingenua Alicia y el Gato sagaz resulta útil para abordar la cuestión de la evaluación en los programas y proyectos sociales.

Alicia nos habla de la *direccionalidad* cuando pregunta hacia dónde tiene que ir; esa pregunta, tan frecuente en la acción social, adopta diferentes formas: ¿cuál es la direccionalidad deseable para nuestro proyecto?, ¿nos estamos dirigiendo hacia donde nos proponíamos?, ¿en qué medida estamos transitando procesos que nos llevan

1. Carroll, Lewis: *Alicia en el País de las Maravillas*, Madrid, Alianza, 1986. La excelente traducción fue hecha por Jaime de Ojeda, quien al respecto de ese diálogo aclara que "es uno de los pasajes más citados de Alicia, sobre todo en libros científicos, en los que ha servido para ilustrar el contraste entre la realidad informe de la materia y el carácter intencional que tiene toda ordenación lógica".

a obtener los resultados que nos propusimos? Y como suele suceder con la mayoría de las preguntas, éstas parecen tener más de un sentido, y por ende más de una respuesta.

Cuando uno se pregunta por la dirección correcta o deseable –que desde ya son dos cosas diferentes– se plantea frente a varias alternativas, frente a caminos que se bifurcan; el grado de conocimiento o de incertidumbre que tengamos acerca de las dificultades o ventajas que cada camino presenta, tendrá que ver con la decisión acerca de cuál tomar; no es sólo la distancia lo que nos inclinará hacia una u otra de las elecciones.

Además, Alicia no plantea su pregunta en abstracto sino que dice “desde aquí”, o sea “desde este punto donde ahora estoy”; eso permite inferir que de algún modo ella llegó a ese punto donde en ese momento está parada, hablando con el Gato. ¿Fue errónea o correcta su manera de llegar hasta ahí? Pareciera que está perdida, temerosa, y que muchos acontecimientos extraños, que no puede explicar bien, le están pasando mientras no sabe exactamente dónde está. Un signo de la ingenuidad de Alicia es que ella no se cuestiona acerca de cómo llegó a donde está, sino que quiere saber cómo tiene que seguir desde allí. No se ha percatado de que, para seguir, debería analizar mejor cómo fue que llegó a donde llegó, para entonces preguntarse acerca de cuáles caminos alternativos existen y cómo son.

Una relevante afirmación, sugerida entonces por el texto de Alicia, vinculada a la cuestión de la evaluación en los programas sociales es que para hacer algo, para introducir cambios en una situación no deseable o que no nos conforma –para ir a alguna parte– hay que tener ob-

jetivos claros y explícitos y hay que saber de qué manera lograrlos, por cuál camino ir, es decir, hay que saber la *direccionalidad deseable*. Incluso eso es así para quedarse donde uno está, para decidir no hacer cambios.

Para que eso esté claro, previamente debimos planificar, programar nuestras acciones en pro de objetivos expresos, es decir, debimos formular un *plan de trabajo*. También necesitamos saber desde dónde partimos, dónde estamos parados; en otras palabras: debimos realizar un *diagnóstico de la situación problemática* que intentamos resolver mediante nuestra intervención. Si no, ¿de qué forma sabremos luego si cumplimos con lo que nos propusimos, si no podemos comparar confiablemente el estado posterior con el anterior? Obviamente, podemos llegar a cualquier parte, sólo se trata de caminar lo bastante, diría el Gato.

### La evaluación como momento o función de la gerencia

De lo que se trata es de identificar, corregir, poner “bajo control”, la máxima cantidad de eventos que puedan oponerse a nuestra acción o dificultarla, de modo de obtener los cambios deseados en la situación que encontramos inicialmente como desfavorable. No cabe duda de que para eso debemos cumplir con procesos metódicos de reconocimiento de la realidad y de programación de acciones en plazos definidos de tiempo, contando con recursos accesibles.

La *planificación o programación* podría sintetizarse como la “reflexión que precede y preside la acción” (Mathus,

1987). En verdad, puede decirse que esa reflexión continúa, a medida que la acción se desarrolla, es propiamente la evaluación, la que alimenta con sus hallazgos la planificación y gestión de los proyectos.

Visto así, se trata de un proceso que sólo conceptualmente es separable, pues en la práctica el proceso de planificación/ejecución/evaluación es una integridad indivisible, conformada por momentos –funciones– en continua retroalimentación. Los teóricos de las ciencias de la administración engloban ese proceso en el concepto de “gerencia”, y también se usa el término “gestión”.

Queriendo transmitir metafóricamente esa cualidad integral de la gerencia y dentro de la gerencia la función o momento de la evaluación, en una de sus brillantes exposiciones Paulo Motta<sup>2</sup> la comparó con el caso de un ómnibus repleto de pasajeros, donde el conductor debe arreglar defectos del motor sin detener el coche y con todos los pasajeros dentro.

Pero aun sabiendo que resulta complicado, si no reflexionamos sobre la acción que desarrollamos mientras los programas o proyectos se realizan, si no cuestionamos o miramos críticamente cómo llegamos a donde llegamos, difícilmente podamos responder a la pregunta de Alicia: ¿cómo seguir desde aquí?, y a la otra, pre-

2. Paulo Motta es profesor de la Escuela Brasileña de Administración Pública (EBAP), de la Fundación Getulio Vargas, y se desempeña además como consultor internacional en diversas agencias. Esa metáfora fue utilizada durante un seminario organizado por CEADEL, en Buenos Aires, en marzo de 1992, financiado por la Fundación W. K. Kellogg, para los becarios del Programa Internacional de Liderazgo.

via, que aquí agregamos: ¿cómo fue que llegamos hasta aquí?

Una misión relevante de la evaluación en el proceso de gestión es justamente ésta: permitir a los conductores de programas o proyectos, y a los otros actores estratégicos, tomar decisiones acertadas, confiables y fundamentadas acerca de cómo seguir, de cómo dar a sus acciones la direccionalidad deseable, basados en las apreciaciones valorativas sobre lo que se vino haciendo y logrando. No en vano se llama “conductores” a los que dirigen proyectos: porque orientan el camino, imprimen la dirección.

En el sentido apuntado también se dice que el proceso de la evaluación viene a ser como la imagen en un espejo del proceso de la planificación: mientras la planificación mira hacia adelante, plantea previsiones de la acción en los escenarios posibles y deseables, la evaluación enfatiza la mirada hacia atrás –más lejos o más cerca, pero siempre hacia atrás–, buscando aprender de lo hecho, de los errores y los aciertos, valorando lo positivo y lo negativo, para poder recomendar giros o refuerzos en la acción futura. Hay un momento en el cual la planificación y la evaluación tienen su intersección: es justamente en el punto de las recomendaciones evaluativas para la acción futura, las que deberán retomarse en el momento de la reprogramación.

### **Intentando una definición**

Llegados a este punto, es posible adelantar la definición de lo que consideraremos evaluación y que nos

permitirá destacar algunos de sus elementos esenciales, independientemente de la diversidad de propósitos, tipos y abordajes que se puedan encontrar:

Entendemos por evaluación de programas o proyectos sociales una *actividad programada de reflexión sobre la acción*, basada en *procedimientos sistemáticos* de recolección, análisis e interpretación de información, con la finalidad de *emitir juicios valorativos fundamentados y comunicables* sobre las actividades, resultados e impactos de esos proyectos o programas, y formular *recomendaciones* para tomar decisiones que permitan ajustar la acción presente y mejorar la acción futura.

Desbrocemos brevemente los pasos relevantes de esa definición preliminar, aunque en los restantes capítulos volveremos sobre ellos.

- *Actividad programada*: quiere decir que la evaluación –con todas las tareas que ella implica– debe ser prevista tal como las otras acciones del proyecto o programa, puesto que es una actividad más. De ahí que en muchas guías de formulación de proyectos, en los últimos tiempos, se ha incluido como ítem a formular el de la evaluación. Decir que la evaluación es *programable*, significa, entre otras cuestiones, que deben atribuírsele recursos específicos (personas idóneas, momentos y lugares específicos, equipos adecuados, insumos, dinero).

- *Reflexión sobre la acción*: ya se dijo que la evaluación consiste en “detenerse” –no en el sentido absolutamente literal sino en el de distanciarse– para analizar qué es lo que se está haciendo (o se hizo), de qué modo, si ello se orienta según la direccionalidad deseable, cuáles escollos y qué oportunidades se presentan, y cuáles logros se han obtenido.
- *Basada en procedimientos sistemáticos*: supone una metodología y técnicas para la recolección y el análisis de la información relevante que alimentará la reflexión y fundamentará los juicios valorativos que se emitan acerca de las actividades, los resultados y los impactos de los cursos de acción implementados.
- *Emitir juicios valorativos fundamentados y comunicables*: es el núcleo de toda evaluación, e implica atribuir un valor, medir o apreciar si se ejecutan las actividades de acuerdo con lo programado, si los resultados obtenidos corresponden a los objetivos y las metas propuestos, así como en qué medida ha mejorado la situación de los destinatarios de las acciones como producto de nuestra intervención. De lo dicho se desprende que para evaluar siempre es necesario hacer *comparaciones* sobre la base de las cuales se emitirán los juicios; cualquiera que sea el objeto de nuestra evaluación se lo contrastará con un parámetro que constituye lo deseado, lo previsto, lo correcto, en suma, la direccionalidad o las metas que habremos definido previamente. Ésas son comparaciones contra estándares, sea como sea que se los haya definido; pero también hay una comparación contra sí mismo en el tiempo, que es la

que se realiza cuando se comparan situaciones o estados actuales de grupos poblacionales, contra sus situaciones o estados pasados, para saber qué cambió y poder emitir un juicio acerca de si la situación está mejor, peor o igual. Esa comparación contra sí mismo en el tiempo es a veces –sobre todo en los primeros tramos de la ejecución de un programa o proyecto– más relevante que la otra, contra estándares, pues aunque se concluya que se está lejos aún de lo deseable –el estándar–, la situación puede haber mejorado mucho debido a la gravedad de la situación inicial o punto de partida –“línea de base”–, y eso es en sí altamente valorable y muy estimulante para los que ejecutan las acciones, ya que la evaluación les demuestra los logros obtenidos y les permite programar sobre mejores bases. La fundamentación de nuestros juicios valorativos dependerá de la consistencia y confiabilidad de la información –cuantitativa y cualitativa– que recojamos, al estilo de las “evidencias” que corroboran los testimonios en un juicio frente a un tribunal. Los juicios valorativos deben poder comunicarse a las audiencias identificadas, lo que significa que se deben redactar informes comprensibles y programar devoluciones de los hallazgos en modalidades adecuadas para la mejor comprensión por parte de los actores involucrados.

- *Formular recomendaciones para tomar decisiones que permitan ajustar y mejorar las acciones:* la evaluación tiene múltiples aplicaciones y puede perseguir diferentes propósitos, pero de modo general interesa destacar que se trata de pensar sobre el hacer, para identificar errores y problemas que dificultan la acción, para per-

feccionar la gestión de un proyecto o programa, para aprender de lo hecho –haya sido acertado o erróneo–, para introducir correcciones, e incluso para comprobar que nuestro proyecto produce efectos no previstos, algunos de ellos no deseados; y finalmente, ¿por qué no?, para el reconocimiento social del esfuerzo realizado y de los actores participantes que lo llevaron a cabo. Sistematizar las *lecciones aprendidas* y brindar, a partir de ellas, las recomendaciones pertinentes constituye uno de los meollos de la evaluación, según nuestra perspectiva. Mediante los procedimientos de la evaluación reconocemos los logros y las fortalezas de la acción desarrollada, así como los obstáculos y las debilidades; las recomendaciones en general deben sugerir apoyarnos en las fortalezas, a modo de palanca, para superar o saltar las dificultades. Y deben servir a todos aquellos que toman decisiones en los diferentes niveles, no sólo a los “gerentes” o directores de los programas o proyectos, sino también a aquellos que en la tarea cotidiana, en el *escenario donde se hacen las cosas*, deben tomar decisiones operativas. Ya se dijo: el momento de las recomendaciones es la intersección entre la evaluación y la reprogramación; de la profundidad y pertinencia de las recomendaciones que se formulen dependerá la utilidad de la evaluación y, por ende, la viabilidad de su aplicación posterior en la toma de decisiones y en la acción.<sup>3</sup>

3. En el capítulo 3 se verá que la viabilidad de la aplicación posterior de los hallazgos evaluativos también depende del involucramiento de los actores durante el proceso mismo de la evaluación, así como de la completa y oportuna devolución de los resultados por parte de los que tienen la responsabilidad de evaluar.

### **Evaluación e investigación: bastantes similitudes y algunas diferencias**

Una evaluación no es equivalente a una investigación, aun cuando comparte muchos de sus procedimientos. La investigación busca fundamentalmente incrementar el cuerpo de conocimiento, estableciendo relaciones entre hechos, fenómenos, situaciones, etc., mientras que la evaluación se preocupa ante todo por el perfeccionamiento de la acción y compromete esencialmente un juicio de valor. En suma, "la evaluación debe ocuparse fundamentalmente de los posibles cambios en un programa existente" (Hamilton Smith, 1981). La investigación termina con conclusiones acerca de cómo ocurren los hechos, alimenta la teoría y el cuerpo de conocimiento sobre cierto segmento del saber, mientras que una evaluación debe terminar con recomendaciones que lleven a modificar lo que se venía haciendo, o bien a agregar nuevas acciones (lo cual, sin duda, se funda en un "saber" acerca de cómo ocurren las cosas y en ciertas hipótesis acerca de cómo debieran ocurrir).

Tanto lo expuesto en el acápite anterior como lo recién expresado permiten concluir que son las *recomendaciones* para la toma de decisiones las que distinguen la evaluación de la investigación; no pareciera lógico pensar en una evaluación por el solo hecho de describir y emitir un juicio sobre cómo se hacen las cosas, si no se le agregan las recomendaciones emergentes para decidir sobre los cursos de acción y cambiar lo que se detecte como mejorable.

El proceso de explicación o fundamentación de los juicios evaluativos es similar al de la ciencia; el rigor en las formas de llegar a emitir aseveraciones certeras es compartido. No se trata sólo de criticar o de elogiar, hay que fundamentar las aseveraciones —para distinguirlas de las creencias, las opiniones o los juicios emocionales—, para luego, a partir de esas explicaciones o fundamentaciones, poder emitir las recomendaciones que permitan superar deficiencias o profundizar las ventajas ya obtenidas con la acción desplegada.

Por lo tanto, no hay límites demasiado precisos entre la evaluación y la investigación; cuando se pone a prueba una modalidad de intervención considerada innovadora, la evaluación, en tanto incluye un proceso riguroso y sistemático de recolección de información para emitir juicios y recomendaciones, sirve además para probar la eficacia de la intervención innovadora, comparándola con otras formas tradicionales de intervenir en el mismo problema. Pongamos como ejemplo los proyectos que desarrollan modalidades de complementación escolar con un abordaje lúdico e integral, inclusivo del seguimiento del rendimiento del niño en la escuela, en articulación con maestros o maestras, en lugar del anterior modo de apoyo escolar con clases enfocadas exclusivamente a "ayudar a hacer los deberes". La evaluación de proyectos con los dos abordajes y la comparación entre ambos pueden permitir detectar que la nueva modalidad es más *eficaz*, mostrando evidencias de que los niños cubiertos bajo esa nueva forma de intervención tienen mayor rendimiento en la escuela y menor índice de deserción o repitencia que aquellos que se cubren con el método tradicional, y porque además la primera

es *sinérgica*,<sup>4</sup> ya que se logran otros resultados no específicos o buscados primariamente, como puede ser la incorporación de hábitos positivos de autocuidado, de higiene, de relaciones con los pares y en la familia, entre otros.

El ejemplo previo es un caso en el que la evaluación puede ser vista también como investigación, ya que posibilita el avance del conocimiento a partir de la sistematización de la práctica. También permite recomendar mejoras para la acción (aun dentro de la misma modalidad innovadora), pues se puede descubrir algún caso entre los analizados donde, además de las actividades lúdicas y de articulación con las maestras de las escuelas, se agrega la interacción periódica con los padres, para concientizarlos acerca de las ventajas de los nuevos modos de llevar a cabo la complementación escolar, lo que redundará en mejores resultados. Ese hallazgo implicará recomendar también actividades concretas dirigidas a los familiares de los niños, promoviendo así enfoques más integrales. Resumiendo: es muy sutil la distinción entre evaluación e investigación.

### **Cambios en los estilos de evaluación**

Las concepciones y los modos de evaluar fueron cambiando en paralelo con los cambios ocurridos en los pa-

4. Max-Neef (1986) considera sinérgicas las acciones que, además de orientarse a la resolución de un problema puntual, contribuyen a la vez a la resolución o prevención de otros problemas. Un ejemplo muy utilizado es el de la lactancia materna, que a la vez que contribuye a la buena nutrición del bebé, lo preserva inmunológicamente, y consolida un lazo afectivo entre la madre y el bebé que podrá contribuir a su apropiado desarrollo posterior.

radigmas de la ciencia, particularmente durante la segunda mitad de este siglo. El paso de las formas mecanicistas, neopositivistas, racionalistas, objetivistas a las concepciones fenomenológicas, más relativistas e inclusivas de factores contextuales y que incorporan metodologías intersubjetivas, al incidir en los modos de conocer y abordar la realidad, influyeron también en las modalidades de planificación, gestión y evaluación de los proyectos y programas sociales.

Hace unas décadas, la evaluación se planteaba como la última etapa de un proceso lineal de formulación/implementación de proyectos y programas, por la cual se trataba de mirar hacia atrás y establecer –siguiendo métodos y herramientas especializados provenientes de la investigación en ciencias sociales y científicamente válidos según el paradigma dominante– si las acciones ejecutadas habían producido los efectos deseados o buscados.

Esto llevó a la elaboración de diseños de evaluación bastante sofisticados con el fin de reproducir condiciones “experimentales” o “cuasi experimentales” que permitieran determinar si se produjeron cambios en la situación problemática inicial que motivó la intervención (proyecto o programa social) y si esos cambios se debieron a las acciones desplegadas. No desconocemos la necesidad de contar con elementos de juicio acerca de los impactos de estrategias o cursos de acción de uso difundido, pero lo cierto es que estos diseños requieren personal técnicamente capacitado para llevarlos a cabo, por lo que suelen convertirse en evaluaciones externas, con escasa o nula intervención de los integrantes de los programas o proyectos; por otra parte, producen información que puede ser útil en ciertos niveles decisorios (por ejemplo, para

decidir la continuidad o no de determinadas líneas de proyectos), pero tal vez menos para quienes están directamente involucrados en la gestión.

Otro camino difundido, dentro de ese modelo de planificación vertical o “de línea recta”, consistía en comparar, también en la última etapa de ejecución o incluso a posteriori, las metas previamente establecidas con las que se habían efectivamente alcanzado, complementando esta comparación con una explicación (o justificación *ex post*) de las discrepancias encontradas. Este camino suponía que lo “planificado” debía traducirse casi mecánicamente en acciones sin tener en cuenta el contexto organizacional y las racionalidades de los diversos actores en su ejecución.

En la actualidad, con el amplio desarrollo de las metodologías de planificación estratégica y/o participativa (que fueron desplazando lentamente, aún no del todo, a los modelos verticales), se reconoce que el proceso de planificación/programación no responde a una racionalidad lineal, que los actores involucrados en la gestión y ejecución reinterpreten permanentemente los objetivos buscados a través de su accionar, en síntesis, que todo el proceso de la gestión social está signado por grandes dosis de incertidumbre, en gran parte como producto de la naturaleza del contexto en que se desarrolla.

Esa incertidumbre es también producto de la propia naturaleza del desarrollo teórico y de las tecnologías utilizadas en las disciplinas sociales y en el campo de la acción social, en comparación con las ciencias exactas o naturales y sus campos de aplicación, como la ingeniería o la electrónica. Vale decir: en las ciencias sociales no hay una única escuela o un paradigma dominante de forma excluyente, como ocurre en un determinado período his-

tórico con la física o la biología, sino que coexisten diversas corrientes, en algunos casos contradictorias entre sí –paradigmas rivales (Khun, 1970)–, en otros con diferentes enfoques o énfasis (caso de las teorías del aprendizaje y las metodologías para la lectoescritura en psicopedagogía, o las diferentes corrientes en metodologías para la planificación participativa, las diferentes escuelas de la psicología social y la influencia de ellas en la determinación de las técnicas más eficaces para las dinámicas grupales). Por ende, no existen certezas acerca de cuáles son los requerimientos mínimos para la correcta aplicación de las diferentes corrientes teóricas.

La evaluación puede verse entonces como una respuesta a la necesidad de reducir esa incertidumbre en cada uno de los momentos de la acción, para reconocer si los esfuerzos están “correctamente” encaminados. Es decir que se requiere una retroalimentación continua en la implementación para ir corrigiendo sobre la marcha los cursos de acción teniendo en cuenta las cuestiones que preocupan a los involucrados en la ejecución (tanto los ejecutores como los destinatarios).

Es en el sentido recién apuntado que la evaluación social se distingue de otro tipo de evaluaciones, por ejemplo, de aquellas que se aplican a proyectos de construcción o de infraestructura; hay modelos de evaluación que han sido trasladados directamente de los proyectos de ingeniería o de obras a los proyectos o programas sociales, lo cual generó muchas dificultades y resistencias por parte de quienes los ejecutan, puesto que esos modelos, a veces impuestos por algunas agencias de financiamiento, son poco adaptables a los escenarios comunitarios donde los proyectos sociales tienen lugar.

### La evaluación para la transformación

Si bien es cierto que cada vez más frecuentemente se habla de evaluar los resultados de un proyecto o programa, el funcionamiento de sus actividades –los procesos– y hasta el uso de sus recursos, también lo es que sigue siendo una actividad poco frecuente y en general bastante resistida. Aun más, cuando finalmente se decide evaluar un programa, luego muy raramente se utilizan sus resultados y recomendaciones para producir cambios.

Creemos que la resistencia a la evaluación, así como las dificultades para incorporar las recomendaciones evaluativas en la gestión, son cuestiones que tienen que ver con el fenómeno más general de la resistencia al cambio; pues al fin y al cabo, la evaluación implica cambio, tal como la planificación y, en general, tal como la acción social: se formula e implementa un proyecto o programa para transformar una situación que resulta insatisfactoria y se evalúan la gestión y los resultados porque interesa, por un lado, apreciar si se cambió lo que se deseaba cambiar, es decir, para saber en qué medida mejoró la situación que se consideró insatisfactoria al inicio de las acciones; y, por otro lado, para detectar qué fue lo que cooperó a ese logro y qué lo dificultó, es decir, qué se hizo bien, qué se podría hacer mejor y qué se hizo mal. Todo eso para poder aconsejar qué cambios introducir o no en la acción posterior.

Existe una tensión entre la propensión al cambio y los impulsos a mantener la estabilidad de las cosas que nos rodean y la manera en que ellas funcionan; lo primero brinda satisfacción a las necesidades humanas de creativi-

dad y de enfrentar desafíos probando las propias fuerzas, mientras que lo segundo permite la idea de pertenencia, de seguridad y de continuidad. Esa tensión sucede tanto en el plano individual o familiar como en lo microsociales y lo macrosociales (Sztompka, 1993). Como seres humanos, como sociedad, siempre han existido fuerzas o impulsos tendientes a cambiar los modos de relacionarnos, de mejorar nuestras instituciones, nuestras formas organizativas, nuestras formas de expresión.

Claro está, como se dijo, que esas intencionalidades no necesariamente son dominantes, y suelen operar en contra o a pesar de otros que tienen posturas opuestas o diferentes, y que prefieren mantener el statu quo, pues ello les brinda seguridades de alguna índole (afectivas, de prestigio, de influencia, económicas, etc.). En el campo más específico de las políticas sociales, esta lucha permanente del hombre por operar cambios en las relaciones sociales se manifiesta a través de actividades tendientes a mejorar la situación de grupos desprotegidos, con mayores desventajas o que han sufrido algún tipo de catástrofe.

Una forma –no la única– en que pueden operarse cambios sociales de manera no cruenta es a través del diseño de programas y proyectos que partan de realizar un recorte y describir una realidad –*diagnóstico*–, para luego identificar las disconformidades –*problemas*–, y de allí inferir qué queremos modificar –*objetivos*–, lo cual es lo mismo que pensar en las transformaciones que esperamos obtener –*resultados*–, para lo que se requiere el diseño de una *estrategia* para operar el cambio, que se traduce en un *plan de trabajo* –*actividades*– ubicado en el tiempo –*cronograma*– y asociado a los insumos –*recursos*

*humanos, materiales y financieros*— requeridos para llevar a cabo las acciones.

La idea de cambio está implícita desde el inicio del proyecto o programa que luego será evaluado. El simple hecho de pensar en un proyecto específico está poniendo de manifiesto una necesidad de que algo cambie. Este proceso de transformación está asociado a la creatividad, ya que además de querer modificar una realidad social, institucional, personal, a partir de la generación de un programa, estamos buscando creativamente la alternativa de acción más efectiva para transformar esta realidad: es decir, nos interesa el cambio y cómo se opera ese cambio; ambas cuestiones están relacionadas con el proceso creativo.

En tal sentido, un programa o proyecto social puede ser visto como la puesta a prueba de estrategias o metodologías innovadoras de intervención para la transformación, las que si luego de evaluadas demuestran ser eficaces, podrán ser replicadas en otros contextos, en circunstancias o condiciones similares, o bien con las adaptaciones que convengan. Además, en el mejor de los casos, la modalidad de intervención podrá ser adoptada como estrategia de las políticas públicas, en el nivel local o nacional, para el campo social específico. Eso sería un ejemplo de cambio social relevante y es uno de los impactos más buscados por los que conducen proyectos sociales, pero difícilmente alcanzado; podemos arriesgar que para influir de ese modo en las políticas públicas no sólo deben evaluarse las acciones, sino además deben difundirse los resultados evaluativos, en una estrategia de *advocacy* o *marketing* social más global. Es allí donde se cruzan las acciones evaluativas con las de conceptualización, sistematización y disseminación.

El propio proceso evaluativo implica transformación. Cuando se comienza a pensar en evaluar un programa en funcionamiento, o incluso en evaluar la factibilidad de realización de otro, en realidad estamos pensando que algo cambie en la forma como se vienen haciendo las cosas. De alguna manera esperamos que los resultados de la evaluación nos brinden elementos para operar ciertas modificaciones en la estrategia o metodología de intervención implementada.

Pero en este punto se produce el otro fenómeno: la posibilidad de cambiar la forma en que se venía trabajando produce resistencias. Frecuentemente, el propósito mismo de la evaluación, “evaluamos para cambiar”, genera oposición o al menos renuencia y entonces, siempre que se puede, se aplaza para otro momento la actividad evaluativa.

La posibilidad de cambio evoca con frecuencia el proceso más alto en cuanto a transformación se refiere, que es la revolución, proceso en que difícilmente algo quede en pie; el impulso hacia adelante que se genera es tan arrollador que los viejos cimientos parecen caer debido al impacto. Quizá sea esto lo que genere tanta resistencia, sin tener en cuenta que hay muchas posibilidades de operar transformaciones antes de llegar a una ruptura total. De hecho, la historia nos brinda ejemplos de esto constantemente; hay infinidad de cambios adaptativos, parciales, que van transformando lentamente la sociedad, mientras que, por otra parte, las grandes transformaciones son momentos muy especiales en el devenir histórico.

Si esto sucede en el abarcativo plano de la historia político-social, no debe extrañarnos que una evaluación

—que bien puede terminar en recomendaciones de cortar el desarrollo de un programa o un proyecto, o de introducir ajustes que no convengan o sean traumáticos para algunos individuos o grupos involucrados en él— produzca al menos resistencias y que sus resultados sean escasamente utilizados (muchas veces porque no son “políticamente viables”, aunque estén técnicamente fundamentados).

### Aprendiendo con la evaluación y de ella

El impacto que produce el hacerse cargo de las cuestiones que deben ser transformadas suele interpretarse como “errores cometidos” y esto lleva casi de inmediato a sentir que nada de lo construido sirve y que es necesario empezar de nuevo. Sin embargo, la idea que hay por detrás de la evaluación, en el sentido que le estamos imprimiendo, es integrar dentro de un nuevo esquema lo mejor de las estructuras, las actividades y los recursos existentes en el programa, modificando aquello que ya no resulta útil, correcto o eficaz y aprendiendo de la experiencia realizada.

La afirmación anterior nos lleva a la cuestión del aprendizaje, que también tiene relación con la resistencia a la evaluación. No toda forma de aprendizaje —modalidad pedagógica— resulta “aceptable”, y por ende eficaz, para los que intervienen en el proceso de enseñanza-aprendizaje. Cualquiera sabe que muchas de nuestras resistencias ante ciertas disciplinas tienen más que ver con la forma en que los profesores la transmitían en la escuela que con características intrínsecas de la disciplina misma. El evaluador, tal como lo estamos esbozando aquí, está muy lejos de representar la figura del maestro tradicional, el que viene

a “enseñar”, a transferir su saber a los otros, a los que están haciendo las cosas en terreno, planteando con ellos una relación de tipo asimétrica y “vertical”. Lejos de esa figura se está pensando en el rol del evaluador como facilitador, con un papel importante en cuanto acompaña los procesos, entrando y saliendo de ellos, conservando la posibilidad de una mirada con suficiente distancia aunque involucrada, desencadenando procesos de reflexión, brindando una perspectiva “preguntona” y crítica, involucrando a todos los actores posibles en espacios de reflexión conjunta y de debate a partir de preguntas, muchas veces ingenuas, otras veces rípidas.<sup>5</sup> No debe olvidarse que ese evaluador también está aprendiendo, también tiene su mente abierta a nuevas formas de actuar. No se trata de un juez; se trata de alguien a la par, con el rasgo distintivo del “preguntón”, del cuestionador, que juega de vez en cuando de *outsider*, que se pone el gorro del abogado del diablo, para provocar y estimular la reflexión, a la vez que aprende junto con los demás. No enseña: provoca y estimula el aprendizaje conjunto.

Es relevante la contribución de la evaluación a los procesos de aprendizaje institucional, puesto que resulta prioritario comprometer a quienes trabajan en los proyectos o programas en un “proceso educativo programado que produzca cambios” (Organización Panamericana de la Salud y Fundación W. K. Kellogg, 1998).

5. Sócrates, en el diálogo con Fedro, su alumno, le explicaba que sólo a través de preguntas es posible aprender. Así, la evaluación puede ser vista como un proceso de aprendizaje de estilo “socrático”, pues se centra en formular preguntas evaluativas y responderlas fundamentadamente (Cf. Platón, *Fedro*).

Es posible plantear algunos requisitos para que una evaluación aporte elementos en el citado sentido de aprendizaje institucional:

- debe ser *útil* a las personas comprometidas en el proceso;
- debe ser *viable*, es decir, realizable en un tiempo al alcance de ellas;
- debe *respetar los valores de las personas* involucradas;
- debe estar bien hecha, con procedimientos adecuados, de modo de proveer información *confiable*.

La evaluación es capacitadora e iluminadora pues busca en la teoría, la experiencia y el conocimiento, las argumentaciones que mejor expliquen los resultados y facilita el crecimiento, el desarrollo y el perfeccionamiento de los que intervienen en un proyecto.

### El proceso de la evaluación

Reiteradas veces se dijo que entendemos la evaluación como un proceso, como una función de la gerencia o un conjunto de actividades, siendo, por ende, programable. Es lo mismo que afirmar que requiere para su desarrollo un proceso o secuencia que le es propio, el cual se esquematiza a continuación.

#### *Programación de la evaluación:*

- definición del *sujeto-objeto* de evaluación,
- establecimiento de la *finalidad*,
- identificación de las *audiencias*,
- determinación del nivel de *análisis*,

- clasificación de las *dimensiones* de análisis,
- definición de las *variables*,
- selección de los *indicadores*,
- determinación de las *técnicas e instrumentos* para la recolección de información,
- determinación de la *secuencia de actividades*,
- identificación de *responsables y actores participantes*,
- establecimiento de la *duración, frecuencia y periodicidad*,
- estimación de los *recursos necesarios (humanos y materiales)*,
- *presupuestación*.

#### *Ejecución de la evaluación*

- *relevamiento de información* mediante el desarrollo de las actividades y la aplicación de las técnicas e instrumentos,
- *sistematización, procesamiento y análisis de la información*,
- *elaboración de informes (incluyendo conclusiones y recomendaciones)*,
- *devolución a los actores y diseminación de conclusiones*.

#### *Evaluación de la evaluación (metaevaluación)*

- análisis de la *utilidad* de los hallazgos,
- apreciación sobre la *viabilidad* del proceso evaluativo y del uso de sus resultados,
- juicio acerca de la *pertinencia* metodológica y ética,
- análisis de la *confiabilidad* de los procesos y hallazgos.

Si bien el proceso evaluativo se ha presentado esquemáticamente de forma aislada y lineal, debe interpretarse que forma parte de un proceso mayor que es el proceso administrativo de la gerencia y que como tal retroalimenta

ta la toma de decisiones o se integra como componente evaluativo del diagnóstico. La secuencia presentada sólo procura mostrar la lógica del circuito evaluativo, pero esos pasos no son secuenciales en los escenarios concretos sino que se producirán superposiciones y retrocesos en medio de los avances.

Nuestra concepción es que la evaluación debería ser un *proceso continuo* que se involucre en la gestión desde su inicio, en el planteamiento del problema a resolver o etapa diagnóstica, durante los procesos de la acción y al finalizar períodos determinados, donde el énfasis estará puesto en los resultados o efectos.

Si bien consideramos continua la evaluación, de hecho suele requerir algunos cortes específicos para determinados aspectos, como el estudio de cambios en conductas o actitudes, o situaciones poblacionales (de nutrición, de empleo, de salud, de alfabetización, etc.) que deberán medirse o apreciarse de forma periódica, mediante encuestas, censos o estudios cuasi experimentales, según sean los casos.

Asimismo, como aclararemos luego, a menudo ocurre que los evaluadores somos convocados para realizar nuestra tarea sólo en la etapa final, o después que finalizó el período de ejecución de los programas o proyectos. En esos casos, lejos de rechazar esas demandas, debemos ofrecer nuestras habilidades a los decisores para que en el futuro puedan aprovechar mejor sus propias experiencias, planteando las evaluaciones tempranamente, de modo que les sea posible considerar las alternativas disponibles para tomar decisiones futuras.

## Los tipos de evaluación

*Las ciencias, después de todo, son nuestra propia creación, incluidos todos los severos estándares que parecen imponernos. Es bueno recordar constantemente este hecho [...]. ¿Qué mejor recordatorio hay que el darse cuenta de que la elección entre teorías que son suficientemente generales para proporcionar una concepción del mundo comprensiva y entre las que no hay conexión empírica, puede llegar a ser una cuestión de gusto?*

PAUL K. FEYERABEND (1984)

Existen muchas formas de evaluar y esto ha dado lugar a una profusa y posiblemente confusa terminología, acuñada en la literatura sobre el tema, que remite a distinciones (a veces a oposiciones) entre modos, tipos, enfoques o perspectivas de evaluación. En su mayoría, surgen de la introducción en la práctica de nuevos focos de interés y de los diferentes énfasis que los mismos evaluadores fueron dando al uso de la evaluación. Muchas aparecen incluso para superar restricciones o limitaciones con que se enfrentaron en el desarrollo de las modalidades más tradicionales. Sin embargo, como se verá, responden a diferentes necesidades y no son excluyentes.

Como producto de los intentos de sistematizar la experiencia en propuestas conceptuales y metodológicas, se pueden encontrar innumerables clasificaciones. Algunas han tenido mucha difusión en la literatura sobre evaluación y merecen una breve descripción pues alertan

acerca de la multiplicidad de facetas que pueden y deben considerarse a la hora de concebir y diseñar un proceso de evaluación. En este sentido, adquieren un valor práctico pues sirven para pensar, presentar y provocar nuevas ideas y técnicas y para disponer de listados de cuestiones a considerar y recordar.

Los tipos más difundidos suelen diferenciarse de acuerdo con los siguientes criterios:

- según el momento en que se realiza la evaluación;
- según quiénes realizan, son responsables o intervienen en la evaluación;
- según los propósitos y la índole de las decisiones que se espera tomar en el futuro en función de los resultados de la evaluación;
- según los aspectos a evaluar del programa o proyecto.

Algunos autores distinguen además entre evaluación “informal” y “formal”. La primera se refiere a nuestras conductas cotidianas que nos llevan a emitir juicios sobre utilidad, conveniencia, eficacia, compatibilidad de otras personas, recursos, procesos o productos sobre la base de supuestos subjetivos e individuales, juicios que nos sirven para elegir diferentes cursos de acción. Se la contraponen a la evaluación “formal” o “en sentido estricto”, en tanto ésta es pública y las elecciones se basan en esfuerzos *sistemáticos* de definir criterios *explícitos* y obtener información *precisa* acerca de las alternativas. Nuestro concepto de evaluación se refiere fundamentalmente a este último tipo.

### Antes, durante o después: siempre hay que evaluar

Una distinción clásica proveniente de la tradición económica y trasladada a las ciencias sociales es la que suele efectuarse entre evaluación *ex ante* y evaluación *ex post*.

Como su nombre lo indica, la *evaluación ex ante* (a veces denominada también “de predecisión”, “de factibilidad”<sup>1</sup> o “de pertinencia”) se emprende antes de iniciar un programa o proyecto para tomar una decisión relativa a si debe implementarse o no. En términos generales apunta a “comparar la eficacia o la rentabilidad económica de las diferentes acciones para alcanzar los objetivos deseados” (Pineault y Daveluy, 1987, cap. 6).

Desde el punto de vista del proceso de intervención social se trata, en la fase de formulación y diseño del proyecto, de establecer, además de su *rentabilidad económica*, su *pertinencia* —es decir, la adecuación de las soluciones que propone a la realidad que se pretende modificar—, su *coherencia interna* en términos de adecuación de recursos a objetivos y de éstos a metas y actividades, y su *factibilidad*, o sea, la disponibilidad de los recursos requeridos y la existencia de condiciones para desarrollar las acciones esperadas.

Es el tipo de evaluación que suelen realizar los organismos financiadores a la hora de decidir aprobar un proyecto. Puede servir para mejorar su formulación y también para seleccionar proyectos, por ejemplo, cuando se trata de un concurso.

1. También se denominaba “costo-beneficio”, identificándola con la principal técnica de análisis económico para la evaluación *ex ante* de proyectos.

La *evaluación ex post* es la que se realiza una vez concluida la ejecución del proyecto. Aquí también se utilizan varios términos —aunque no con un significado unívoco— para hacer referencia a diferentes momentos en los que se realiza, sea apenas concluida la fase de ejecución (de fin de proyecto, a posteriori, terminal), sea varios años después que el proyecto o programa ha terminado.<sup>2</sup> De todos modos, en general se concentra en los resultados obtenidos para evaluar en qué medida se alcanzaron los objetivos previstos (situación objetivo), cuáles han sido los efectos, buscados y no buscados, atribuibles al proyecto, sobre la situación inicial que se pretendía modificar; también procura identificar los factores que contribuyeron a producir estos resultados. Merece destacarse que un aspecto importante —a menudo descuidado— de esas evaluaciones es contar con una “línea de base”, o información diagnóstica de la situación inicial, para poder contrastarla con la situación final. Aunque acorde con las metodologías circulantes de planificación/programación, la formulación de un proyecto o programa debe acompañarse de un diagnóstico, no es frecuente que éste se realice con la mirada puesta en el proceso evaluativo, y por ende difícilmente incorpora la información pertinente para poder apreciar los cambios producidos después de su finalización. Lo dicho muestra que, pese al camino ya

2. En algunos casos, la evaluación realizada algunos años después de la conclusión de un proyecto se denomina “de impacto” (véase, por ejemplo, Banco Interamericano de Desarrollo, s/f). Sin embargo, entendemos que la evaluación de impacto puede realizarse también en momentos intermedios si existen hipótesis acerca de la ocurrencia a mediano plazo de los efectos esperados en la población.

recorrido, todavía resulta difícil encontrar traducida en hechos la estrecha relación que se requeriría entre la programación y la evaluación.

En las evaluaciones *ex post*, el foco está colocado en la eficacia, la efectividad y la eficiencia de un proyecto o programa, conceptos que retomaremos y explicaremos en otro punto de este capítulo (véanse las págs. 69 a 71). Pero es necesario destacar que apuntan fundamentalmente a obtener elementos útiles para decidir si conviene seguir implementando esta línea de proyectos y, en caso afirmativo, qué aspectos pueden mejorarse en futuras formulaciones.

A partir del énfasis en la evaluación como instrumento de gestión estrechamente ligado a las distintas fases de la planificación, se introdujo una nueva diferenciación relativa al tipo de *evaluación* que se realiza *durante* la ejecución del proyecto, denominada de muy diversas maneras: *de progreso o proceso, de gestión, concurrente, concomitante, intermedia, de operación*, etc. Su objetivo básico es evaluar en qué medida se va cumpliendo el proyecto o programa de acuerdo con la propuesta inicial, es decir, si se orienta según la direccionalidad deseable. Genera información sobre los procesos, las actividades y los productos, y los compara con las metas formuladas y las que se van reformulando a lo largo del proceso de implementación. Se realiza de forma periódica de modo de detectar oportunamente las debilidades y los obstáculos así como las fortalezas y las oportunidades. De este modo proporciona elementos para la toma de decisiones que permiten reforzar o reorientar acciones y corregir aspectos de la gestión, aumentando así las posibilidades de llegar a resultados favorables.

Aunque se lo distingue, también el *seguimiento* o *monitoreo*<sup>3</sup> se refiere a la evaluación que se realiza durante la ejecución de un proyecto. Ha sido definido como

[...] el examen continuo o periódico que efectúa la administración, en todos sus niveles jerárquicos, de la manera en que se está ejecutando una actividad. Con ello se busca asegurar que la entrega de insumos, los calendarios de trabajo, los productos esperados se conformen a metas establecidas y que otras acciones que son necesarias progresen de acuerdo con el plan trazado (ONU, 1984).

Es parte de la función gerencial y responsabilidad del personal a cargo de la ejecución, y en este sentido constituye una típica actividad de evaluación interna dentro del marco de un programa o proyecto. En general, las actividades de monitoreo que hemos conocido se centran en la implementación de un sistema de información para verificar el logro de las metas e identificar desvíos. Este énfasis en la recolección sistemática de informaciones útiles para el funcionamiento de programas ha llevado erróneamente a diferenciar el monitoreo de la evaluación propiamente dicha. En efecto, la existencia de un sistema de información no implica necesariamente que exista evaluación y, viceversa, puede hacerse una evaluación sin la existencia previa de dicho sistema ya que puede recogerse información en el momento del corte evaluativo. ¿Pero de qué sirve un sistema de información si no se lo utiliza para juzgar la adecuación de los procesos de ejecución y proveer elementos de juicio para la introducción de ajustes y correcciones oportunas? También

3. Los dos términos son equivalentes, pues "monitoreo" ha sido traducido literalmente del inglés *monitoring*.

hemos leído o escuchado decir que el punto donde radica la diferencia entre "evaluación" y "monitoreo" es que este último provee información para la gerencia... ¿y acaso la evaluación no lo hace? Quiere darse a entender que el monitoreo prende luces de advertencia y/o riesgo en el gerenciamiento social del día tras día, para generar acciones correctivas inmediatas, del mismo modo que los monitores con el paciente internado en las terapias intensivas e intermedias; pero es necesario advertir que, en los proyectos sociales, la metáfora del paciente monitorizado no es del todo afortunada, ya que es imposible la recolección de información de forma continua o ininterrumpida que permita prender luces de emergencia; sólo es posible recoger alguna información de modo discontinuo, con cierta periodicidad (a lo sumo mensual, por lo general trimestral). Esa información recogida con periodicidad frecuente suele ser del tipo cuantitativo (de ejecución presupuestaria, de cobertura, de cantidad de prestaciones o servicios brindados), pero raramente incluye aspectos de calidad o procedimentales, ya que éstos requieren periodicidades mayores.

Se dice también que el monitoreo adecuadamente conceptualizado y puesto en práctica es un

[...] procedimiento empleado para comprobar la efectividad y eficiencia del proceso de ejecución de un proyecto, mediante la identificación de los aspectos limitantes y/o ventajosos que culmina recomendando medidas correctivas para optimizar los resultados deseados (Banco Interamericano de Desarrollo, s/f).

Se puede apreciar, según esta definición, la ausencia de diferencias entre el concepto de monitoreo y el de evaluación.

### Conjugando el verbo evaluar: ¿nos evaluamos o nos evalúan?

En función de quién realiza la evaluación, se han diferenciado dos grandes categorías, la evaluación externa y la evaluación interna, aunque actualmente también se ha roto esta dicotomía para incluir la evaluación mixta, la autoevaluación y la evaluación participativa.

Se llama *evaluación externa* a la que realizan evaluadores que no pertenecen a la organización ejecutora del programa o proyecto que se evalúa, recurriéndose generalmente a investigadores o consultores independientes, departamentos universitarios y, a veces, incluso a expertos internacionales. Es posible que la evaluación externa se realice por decisión del propio equipo responsable del proyecto, programa u organización, pero también ocurre que se realiza por decisión de la agencia financiadora o del organismo del cual depende jerárquicamente el ejecutor; lo cierto es que esas variantes tendrán influencias en las modalidades y orientaciones evaluativas.

La *evaluación interna* es la que llevan a cabo personas o grupos pertenecientes a la institución gestora del proyecto, no involucrados directamente en su ejecución; en cambio, la *autoevaluación* sería aquella realizada por los propios actores a cargo de la ejecución del proyecto. Esta diferenciación en realidad adquiere sentido en el marco de organizaciones de cierto tamaño, relativamente complejas, en las que puede existir una persona/equipo/unidad específica de evaluación, cuasi externa, es decir, por fuera del área de ejecución del proyecto. En organizaciones más pequeñas, con estructuras indiferen-

ciadas, una evaluación interna sería equivalente a una autoevaluación.

La *evaluación participativa* aparece como una categoría sumamente difusa, pues no es claro a la participación de quiénes se refiere, pero apuntaría –y en este sentido aparece como modelo ideal– a designar aquella evaluación que involucra a los destinatarios y/u otros actores de los proyectos en los distintos momentos del proceso evaluativo (es decir, no sólo como fuente de información sino además con protagonismo en su diseño, en la ejecución, en el análisis de los resultados y en la formulación de conclusiones y recomendaciones).

En la mayoría de los libros sobre evaluación que hemos consultado hay discusiones acerca de las ventajas y desventajas de cada modalidad: “la organización ideal sería aquella que se evalúa a sí misma” (Wildavsky, 1972), y nadie mejor que su personal para conocer y comprender los objetivos, la metodología, los destinatarios y la dinámica cotidiana de las actividades, en suma para captar la racionalidad interna del proyecto. El hecho de que los propios actores realicen la evaluación hace más viable la posterior utilización de los hallazgos evaluativos y la adopción de medidas correctivas recomendadas. Pero su principal obstáculo deriva de una baja garantía de objetividad al ser juez y parte del proceso evaluativo: sus valores, intereses y compromisos personales atentarían contra una apreciación imparcial e independiente del programa o proyecto.

Por el contrario, se supone que el evaluador externo no tiene intereses creados –supuesto que casi nunca es del todo correcto–, suele traer experiencia en el tema, puede volcar recursos y conocimientos inexistentes en la

organización y es mayor garantía de objetividad. Sin embargo, sus debilidades estarían en las tensiones entre evaluador y evaluado, que pueden incluso generar dificultades en el acceso a la información, en las restricciones para captar los factores en juego y la racionalidad interna del proyecto y en las probables resistencias de su personal a aceptar resultados que podrían cuestionar sus objetivos, metas y líneas de acción.

Por otra parte, en nuestra concepción, el evaluador o equipo de evaluadores externos contratados traen, además de sus propias orientaciones y preferencias ideológicas y teóricas, compromisos con quienes los contratan y ello puede incidir sobre la supuesta "objetividad" de su labor.<sup>4</sup>

La solución salomónica propuesta en algunos casos es la *evaluación mixta*, conformando un "equipo" constituido por evaluadores externos e internos con momentos de trabajo conjunto y momentos de trabajo independiente.

Otra forma de preguntar acerca del tema de quién evalúa, sobre todo en el caso de organizaciones o programas de cierta complejidad (con diversos departamentos o subsedes o localizaciones) es ¿en qué nivel evaluamos? En tal sentido, si la evaluación es a nivel central suele hablarse de *control de gestión*; en los niveles intermedios se hace énfasis en la *supervisión* y si se realiza en el nivel operativo, donde se hacen concretas las acciones (prestación de bienes o servicios) en relación con la gente (los beneficiarios o usuarios o clientes), se habla de *evaluación*

4. Para un mayor desarrollo de este tema, véase el capítulo 3 del presente libro.

*operativa*, muy ligada al propio trabajo de los equipos técnicos actuantes (por ejemplo, los grupos operativos de mejoramiento de calidad o de auditoría médica que llevan a cabo los grupos profesionales en los establecimientos de salud).

La cuestión de quién evalúa en realidad lleva implícito el problema de la resistencia a la propia práctica de la evaluación y a las dificultades para utilizar sus resultados. No es inusual que se lleven a cabo evaluaciones rigurosas —sean internas o externas, de procesos o terminales, en uno o en otro nivel de la organización, programa o proyecto— cuya implementación y desarrollo han sido aceptados por la organización pero cuyas recomendaciones son escasamente tenidas en cuenta luego por los propios responsables y ejecutores de proyectos. Su compromiso personal y emocional con las metas, la población destinataria y otros aspectos específicos de la organización se puede erigir en una barrera para la aceptación de cambios en la orientación de su accionar. En síntesis, como ya se vio en el primer capítulo, existen contradicciones entre los propósitos de la evaluación y la lógica de funcionamiento de las organizaciones. Este tema, recurrente en la revisión de las experiencias, nos recuerda que la programación de una evaluación debe incorporar acuerdos y negociaciones entre todos los involucrados en el proceso, sobre todo en torno a sus propósitos y a su diseño. También es necesario contemplar la introducción de mecanismos que permitan compartir sus diferentes etapas, incluyendo el análisis y la interpretación de la información y la formulación de las recomendaciones. Sólo así se podrá hablar genuinamente de evaluación participativa.

## Los múltiples propósitos de la evaluación

Una distinción ya clásica y muy difundida es la que suele hacerse entre *evaluación formativa* y *evaluación sumativa*<sup>5</sup> o *de resumen*. La primera se refiere al tipo de evaluación que se emprende con el propósito de mejorar una intervención en un momento determinado y para un grupo específico de personas, particularmente aquellas involucradas en su ejecución. Busca proporcionar información útil para ayudar al programa a mejorar su gestión y contribuye a la toma de decisiones acerca de su desarrollo e implementación (modificación, revisión, ajustes, etc.). Brinda información a medida que avanza el programa o el proyecto, y se refiere al cumplimiento de metas intermedias y a los problemas de implementación. Sirve fundamentalmente al personal del programa.

Un ejemplo de evaluación formativa es la evaluación curricular en el área educativa, de donde proviene por otra parte esta noción (Scriven, 1967). Utilizando técnicas tales como seguimiento de su aplicación mediante pruebas de aprendizaje y juicio de expertos, se obtiene información que permite evaluar la adecuación de los contenidos curriculares a las condiciones de enseñanza-aprendizaje y así poder realizar los ajustes necesarios.

La evaluación "de resumen" se emprende con el propósito de obtener un juicio global sobre el valor de algu-

5. Esta traducción literal ha sido acuñada del inglés *summative* y nos parece poco afortunada, pues en castellano involucra la connotación de suma o acumulación. Parece mejor la traducción que UNICEF hace del término como una evaluación "de resumen", pues en realidad, y tal como la usó Scriven en su origen, connota una recapitulación de los efectos de un conjunto de acciones (programa, proyecto) sobre sus destinatarios (UNICEF, 1990).

na intervención o acción humana (proyecto, programa, política, servicio). Debe producir información acerca de su efectividad y de las condiciones en las que puede ser eficaz en otras situaciones, momentos o lugares. Se emplea frecuentemente para decidir qué programas o acciones merecen una nueva inversión o la obtención de fondos adicionales. Lleva por lo tanto a decisiones relativas a la continuación, finalización, expansión de los programas o proyectos y responde principalmente a las necesidades de quienes los formulan o los financian. En particular, contribuye a suministrar elementos de juicio sobre la replicabilidad de una determinada intervención en diferentes contextos, ampliando de este modo el abanico de sus potenciales usuarios. En el caso presentado anteriormente acerca de la aplicación del currículo elaborado y mejorado, se emprendería una evaluación "de resumen" con el propósito de determinar si es efectivo para mejorar las capacidades de los alumnos o para introducir cambios en sus actitudes hacia el aprendizaje. El juicio evaluativo en este caso sirve para decidir la continuación del programa y la generalización de su aplicación a otros establecimientos escolares.

Sin embargo, si se tiene en cuenta la íntima relación entre los procesos de gestión y el impacto obtenido, puede decirse que muchas evaluaciones sirven a ambos propósitos. En nuestro ejemplo, la evaluación formativa del currículo llevará de hecho a decisiones acerca de su aplicabilidad, mientras que la de resumen puede contribuir a nuevas revisiones y ajustes para aumentar su eficacia.

Si bien en la práctica suelen usar distintos tipos de información y criterios de evaluación, en nuestra opinión,

el valor de esta distinción reside en un llamado de atención a la necesidad de realizar evaluaciones durante todo el ciclo vital de un proyecto. En realidad, como se ha afirmado en el primer capítulo, el propósito de la evaluación es contribuir a un aprendizaje que impulse el cambio o la aplicación de medidas correctivas. Desde esta perspectiva, toda evaluación cumple una función formativa.

Otra distinción también tradicional en la literatura sobre el tema es la que se establece entre *evaluación descriptiva* y *explicativa* (Suchman, 1967). La primera se refiere a los estudios que tienen por objetivo pronunciarse respecto del éxito o fracaso de un programa a partir de la descripción cuantitativa o cualitativa de sus actividades, sus procesos o sus resultados y de su comparación con algún objetivo, meta o estándar preestablecido.

La segunda se propone identificar los factores que explican situaciones de éxito o fracaso de los objetivos de un programa. Construye esquemas interpretativos de la multiplicidad de causas e interdependencia de efectos a partir de marcos conceptuales y teóricos preexistentes o bien de la formulación de hipótesis a verificar.

Entendemos que se trata de una distinción artificial. Toda evaluación debe *describir* (y ello posiblemente constituya su primera etapa metodológica), pero su diseño debe considerar también la identificación de los factores que intervienen para *explicar* los logros y los fracasos. Por otra parte, como ya dijimos, toda intervención social es una estrategia destinada a producir cambios en un contexto determinado. Implica una serie de hipótesis acerca de los resultados esperados que se traducen en la formulación de objetivos; por ejemplo, si el objetivo de un pro-

yecto es producir una mejora en las condiciones de vida de un grupo o una comunidad mediante la puesta en marcha de un microemprendimiento, existe una hipótesis subyacente de que esta acción será efectiva para producir cambios en la situación de empleo y en el nivel de ingresos del grupo destinatario. La evaluación, en última instancia, busca verificar o desechar tal hipótesis a través de la descripción de los resultados alcanzados y explicar las condiciones en las que se verifica, o en caso contrario identificar los factores –tanto del contexto o mundo exterior significativo como de su gestión– que dificultan u obstaculizan los logros previstos.

### Énfasis según temas: ¿qué evaluamos?

Una distinción relevante se basa en los aspectos o facetas de un programa o proyecto que se quiere estudiar y juzgar. Desde esta perspectiva, existen innumerables tipologías pero, en general, responden a tres grandes criterios clasificatorios.

El primer criterio se basa en las *etapas* del ciclo de un programa o proyecto: en este caso se diferencian los aspectos vinculados a su *diseño y conceptualización* (evaluación de necesidades, evaluación de contexto, evaluación del diseño y programación), aquellos vinculados a su *desarrollo o implementación* (estructura, organización, recursos humanos, materiales y económicos, procedimientos, actividades, etc.) y aquellos vinculados a sus *resultados* (efectos, impactos, rendimiento, etc.). Esta clasificación prácticamente coincide con la que distingue los momentos de la evaluación. Sólo difiere en que pone el acento

en el qué –los contenidos– y no en el cuándo –los momentos– se evalúa.

El segundo criterio alude más bien a los componentes relevantes de los programas. En este eje las clasificaciones más difundidas son: evaluaciones de insumos, procesos y productos; de estructura, procesos y resultados o efectos; de procesos y de impactos.

La *evaluación de insumos* tiene por objeto describir y analizar los recursos humanos, materiales y financieros con que cuenta un programa; la forma en que dichos recursos son procesados de acuerdo con criterios de relevancia, factibilidad y economía en su uso y las estrategias de obtención de insumos destinados a implementar dichas estrategias. La *evaluación de procesos* estudia los medios procesales (actividades, procedimientos, prestaciones) que se utilizan para llevar a cabo un programa, tanto en su diseño como en el modo como éste se implementa. Se trata de emitir un juicio acerca de la capacidad y las estrategias de procesamiento de insumos que tiene el sistema. Sirve para elegir cursos alternativos de acción para la operacionalización, ejecución y el control del proyecto. La *evaluación de productos* estudia el logro de metas propuestas que se expresan en términos de productos esperados. Emite un juicio sobre los resultados obtenidos en términos de productividad, concepto que vincula los productos alcanzados a los insumos utilizados para el desarrollo de las actividades. Esta clasificación proviene del campo de la economía, de las evaluaciones de emprendimientos productivos y de obras físicas.

La clasificación en *estructuras, procesos y resultados (o efectos)* proviene del campo de la salud (Donabedian, 1990) e incorpora una terminología más adecuada a la ín-

dole de los proyectos sociales. La noción de estructura alude aquí a la combinación de los recursos y a su organización para la prestación y el desarrollo de determinados servicios y actividades (procesos) con el objeto de alcanzar ciertos resultados. Por otra parte, esta clasificación vincula la evaluación a la planificación/programación. En la metodología moderna de planificación se plantean objetivos en términos de resultados, procesos y estructuras, en el entendimiento de que los resultados se obtienen si se desarrollan los procesos programados, los que a su vez se podrán realizar en la medida en que se disponga de las estructuras necesarias. La evaluación retoma esos componentes y se evalúan estructuras, procesos y resultados, en una secuencia inversa a la de la planificación.

La clasificación relativa a *evaluación de procesos y evaluación de impactos* es muy utilizada en el campo de los proyectos sociales. En este marco terminológico, la primera apunta a determinar el valor del desempeño actual del proyecto, es decir, el grado en que va cumpliendo con las actividades previstas, alcanzando los resultados esperados dentro de los plazos y con los recursos programados, y procura identificar las causas de los problemas que van surgiendo de modo de corregir y revisar oportunamente la marcha del proyecto. La evaluación de impactos pone énfasis en los resultados logrados en los destinatarios, los compara con los esperados y procura encontrar las causas de las diferencias. Muchas veces esas causas se encontrarán con claridad a partir de la evaluación de los procesos realizados; otras veces tendrán más que ver con factores contextuales que con los proyectos mismos.

La multiplicidad de denominaciones para componentes similares de los proyectos merece aquí un párra-

fo aparte, en particular las que se refieren a sus resultados. “Productos”, “efectos”, “impactos” son términos que pueden englobarse dentro del concepto más general de *resultados* aunque no son utilizados siempre de forma equivalente ni unívoca. Los matices que suelen encontrarse remiten en última instancia a distintos tipos de resultados o distintos momentos del ciclo de un proyecto en que éstos pueden verificarse, así como a sus diferentes ámbitos o espacios de repercusión. Los objetivos del programa o proyecto constituyen la contracara de los resultados y se los suele denominar también con distintos grados de generalidad (finalidad, propósito u objetivo general, objetivos específicos, objetivos intermedios, etc.). Podemos ilustrar estos matices mediante un ejemplo sencillo. Tomemos un proyecto de capacitación técnica de jóvenes no escolarizados de bajos recursos implementado con el objetivo de que obtengan una rápida salida laboral para contribuir a mejorar las condiciones de vida de sus familias. Se suele denominar “productos” a los resultados concretos de las actividades, en este caso, los cursos realizados y los jóvenes capacitados. Los efectos (buscados o programados) se vinculan a los objetivos del proyecto, en este caso: la inserción ocupacional de los jóvenes capacitados. También pueden referirse a otros resultados, previstos o no, como el aumento de la autoestima de los jóvenes capacitados, la mayor integración familiar en sus hogares, etc. El impacto se vincula a los objetivos más generales del proyecto y se refiere a los cambios que se persiguen (a veces no formulados explícitamente) y que podrán verificarse en las condiciones de vida de la población; en el ejemplo podrían ser impactos el mejoramiento del ingreso familiar, la reducción de

la delincuencia juvenil, entre otros. Es necesario señalar que la evaluación de impacto busca identificar aquellos cambios atribuibles a la acción del proyecto, despejando otros factores que podrían haber intervenido.<sup>6</sup>

Finalmente, el tercer criterio de clasificación pone el acento en ciertos atributos de los programas o de sus componentes: la pertinencia, la idoneidad o suficiencia, la eficacia, la efectividad, el rendimiento, la rentabilidad, la productividad, la eficiencia, etc.,<sup>7</sup> conceptos que no siempre son unívocos pero que cuentan con algún consenso, al menos en cuanto a su significado general.

La *pertinencia* se refiere a la adecuación de un programa o proyecto (o de alguno de sus componentes) para satisfacer necesidades y demandas o para resolver la situación –problema que le dio origen–. En otros términos, se refiere a la adecuación de la formulación al “estado del arte” en la respectiva rama del saber.

La *idoneidad o suficiencia* se refiere a la capacidad de un programa o proyecto (o, más específicamente, de sus componentes) para contribuir a los objetivos y metas programados. Así, suele hablarse de idoneidad de las actividades, suficiencia de los recursos, etc.

La *eficacia y la efectividad* se refieren a la capacidad de un programa o proyecto para alcanzar los objetivos programados que, como se ha mencionado, no son sino la expresión de los resultados buscados. Si bien los dos términos suelen usarse como sinónimos, es muy útil establecer una

6. Se volverá sobre la cuestión de los distintos tipos de resultados en el capítulo 5 de este libro.

7. Se pueden agregar otros atributos a esta lista, como accesibilidad, equidad, sustentabilidad. Nos detendremos nuevamente en estos conceptos al abordar nuestra propuesta de “modelo evaluativo”.

distinción entre ambos (Pineault y Daveluy, 1987). La eficacia se refiere a la capacidad de un proyecto o de alguno de sus componentes para alcanzar sus objetivos en condiciones ideales, es decir, si se dan las condiciones prefijadas y aplicando las normas tal como se han formulado para la aplicación de las técnicas específicas, mientras que la efectividad se refiere a la comparación de los resultados obtenidos respecto de los esperados, teniendo en cuenta las condiciones concretas de ejecución, en las circunstancias reales del contexto del proyecto o programa y con aplicaciones distantes en mayor o menor medida de las normas formuladas. Los autores citados hablan por un lado de eficacia potencial y por otro de eficacia real (o efectividad) y sus explicaciones se refieren a programas de atención de salud, los cuales suelen estar más normatizados que la mayoría de los programas y proyectos del campo social.

Como puede lograrse una multiplicidad de resultados (expresados en objetivos de distinto nivel de especificidad), la evaluación de la eficacia o la efectividad de los proyectos sociales suele enfatizar sus fines últimos en términos de las soluciones a los problemas sociales o del mejoramiento de la situación del grupo humano al que van dirigidos. Un proyecto puede no ser eficaz, aunque alcance sus metas en términos de los servicios o actividades a realizar, aunque haga un uso adecuado de los recursos, aunque tenga una estructura que favorezca la realización de las tareas programadas, si no logra producir los cambios deseados en las condiciones de vida de la población a la que se dirige.

En este punto es necesario reiterar la importancia que tiene contar con buenos diagnósticos acerca de las características, y agregamos también de la magnitud, del problema específico que se pretende resolver mediante la inter-

vencción social determinada; habrá que responder entonces a las preguntas siguientes: ¿en qué consiste el problema?, ¿qué consecuencias habrá si no se lo resuelve?, ¿cuántas de las personas en la población del territorio considerado padecen el problema a resolver? Al responderlas conoceremos la *relevancia y magnitud del problema*. Y algo muy importante: tendremos los “denominadores” apropiados para poder estimar luego la cobertura del programa o proyecto y a partir de allí su eficacia real o efectividad. Se entenderá mejor con un ejemplo: se trata de un centro donde se brinda complementación escolar para niños con bajo rendimiento que asisten a las escuelas de un determinado barrio. Para poder establecer la eficacia de ese centro hay que conocer la magnitud del problema en las escuelas del barrio; desde un punto de vista restringido, puede decirse que el centro es eficaz pues los chicos que asistieron a él mejoraron su rendimiento y promovieron al grado posterior, pero desde un punto de vista más general, si la cantidad de niños atendidos es muy baja respecto de la cantidad de niños que padecen el problema en las escuelas del barrio, la intervención no está siendo del todo eficaz pues alcanza a resolver sólo una pequeña parte del problema, y por ende una conclusión es que para mejorar la eficacia debería aumentarse la escala de la intervención en ese barrio.

El rendimiento, la rentabilidad, la productividad y la eficiencia constituyen conceptos asociados generalmente a la evaluación económica de un proyecto. El término más general y relevante es el de *eficiencia*. Ésta se define como la relación entre los productos o resultados esperados de un proyecto y los costos de los insumos o procesos de apoyo o funcionamiento que implica. Esta relación puede establecerse de diferentes formas y utilizando diversas técni-

cas procedentes del análisis económico (costo-beneficio, costo-eficacia o efectividad). Las aproximaciones a un cálculo de costo-beneficio en el campo de lo social son muy complejas por la dificultad de valorar y traducir en unidades comparables –monetarias o no– los beneficios o resultados muchas veces “intangibles” de intervenciones sociales, en relación con los costos de los recursos que insumen las posibles alternativas de inversión (*ex ante*). La evaluación de la eficiencia de un proyecto debe brindar elementos de juicio para identificar alternativas de acción que permitan obtener mejores resultados con igual cantidad de recursos o bien los mismos resultados con una mejor combinación o utilización de los recursos existentes. En el campo de la salud y de la educación se han desarrollado instrumentos para la evaluación de la eficiencia de servicios y programas. En el campo más amplio de las intervenciones sociales destinadas a incidir en las condiciones de vida de la población queda aún mucho camino por recorrer. Pero al transitarlo, debería tenerse en cuenta que lo que es necesario mejorar es la *eficiencia de los proyectos eficaces*, lejos de perseguir la eficiencia como un fin en sí mismo o como valor absoluto. Esta última situación es frecuente cuando el campo de la acción social es “invadido” por criterios meramente economicistas.

### **Los distintos objetos posibles de evaluación: ¿tipos o niveles de análisis?**

Existe una multiplicidad de unidades diferentes que han sido y son objeto de evaluación: desde un conjunto de organizaciones (por ejemplo, el sistema educativo, los

sistemas locales de salud), una organización (escuela, hospital, organización comunitaria), una política, un programa o un proyecto hasta productos específicos o servicios brindados por una organización o por un conjunto de organizaciones. Cabe aclarar que se entiende por programa o proyecto social una intervención social planificada destinada a obtener determinados resultados en un segmento de población. Ambos suelen diferenciarse entre sí por su escala y su duración. El programa alude más a un conjunto de actividades integradas con continuidad temporal, mientras que un proyecto constituye una unidad menor desde el punto de vista de su complejidad y la cantidad de componentes, y se diferencia del programa porque tiene una duración limitada y atiende una demanda o una necesidad social puntual. En el lenguaje de la programación tradicional, solía considerarse proyectos a los componentes de un programa.

Cada una de las unidades mencionadas tiene algunas especificidades en términos de criterios de evaluación y dimensiones de análisis, pero muchos aspectos en común. Asimismo, pueden requerir algunas técnicas específicas (por ejemplo, en el caso de organizaciones se utilizarán las que son más propias del análisis organizacional). La experiencia acumulada en la evaluación de algunos de estos objetos ha llevado al desarrollo de subáreas que se manejan con criterios y técnicas propias: por ejemplo, la evaluación de actividades, proyectos o programas de capacitación o educativos, o bien la evaluación de servicios de salud.

Desde el punto de vista conceptual no constituyen de por sí tipos diferentes de evaluación: cualquiera de las clasificaciones anteriormente mencionadas puede aplicarse a estos diferentes objetos de evaluación.

Importa más bien distinguir los niveles de análisis o recortes de la realidad cuyas especificidades deben tenerse en cuenta en el diseño de la evaluación de modo de utilizar encuadres metodológicos, formas de abordaje y técnicas adecuadas. Los riesgos más frecuentes en el abordaje y el análisis son la transposición del nivel –falacia del nivel equivocado– y la utilización inadecuada de técnicas que no se corresponden con el nivel de aplicación para el que fueron concebidas.

Para dar un ejemplo del campo de salud, es frecuente que se trasladen las variables vinculadas a la relación médico-paciente (o equipo de salud-paciente) a la evaluación de programas de salud o del desempeño de establecimientos, o más aún a la evaluación de sistemas locales de salud, cuando cada uno de estos niveles tiene especificidades que le son propias.

La estrategia de evaluación de *clusters*,<sup>8</sup> desarrollada en los últimos años en respuesta a iniciativas de gran escala financiadas por fundaciones privadas y agencias públicas americanas y de varios países europeos (Worthen y Schmitz, 1997), constituye un ejemplo de cómo los diseños de evaluación deben adecuarse al nivel de análisis donde se concretarán. Esta estrategia se aplica a determinadas líneas de programas con ejecución descentralizada dentro de las que se identifican conjuntos de proyectos considerados “agrupaciones” porque tienen en común ya sea su ideario o misión, su modalidad de intervención o el tipo de población destinataria (los niños, los jóvenes, etc.), y adicionalmente, aunque no exclusivamente, el

8. *Cluster* significa, en inglés, racimo, ramillete, haz, enjambre, agrupamiento, constelación.

origen del financiamiento. En esos casos, la evaluación se orienta, mediante un diseño particular, a obtener apreciaciones sobre el desempeño global del conjunto así como sobre el desempeño de cada proyecto integrante del conjunto. Se trata por lo tanto de un abordaje específico para un nivel de análisis particular: un grupo de proyectos que comparten determinadas características. En el capítulo 5 se encontrarán referencias a la evaluación de *clusters* o programas complejos.

Para concluir, cabe aclarar que nuestra conceptualización de la evaluación se aplica a cualquiera de los niveles mencionados; lo que podrá variar es el foco, así como algunas de las técnicas que se seleccionen para recolectar y analizar la información relevante. Sin embargo, pondremos el énfasis y utilizaremos ejemplos de programas o proyectos, términos que usaremos indistintamente.

### Para evaluar las evaluaciones

La *metaevaluación* puede considerarse un caso particular de la clasificación previa, pues se distingue por su objeto, a la vez que constituye un nivel de evaluación; desde otro punto de vista puede ser también catalogada como una evaluación externa. Se realiza con el propósito de juzgar aspectos centrales del propio proceso evaluativo: apreciar su viabilidad y la del uso de sus resultados, analizar la utilidad de los hallazgos, emitir un juicio acerca de la pertinencia metodológica y ética así como de la confiabilidad del proceso evaluativo realizado y de los hallazgos, la adecuación de las conclusiones y las recomendaciones que se formularon.

Si son pocos los programas o proyectos sociales que se evalúan de forma integral y continua, son menos aun aquellos casos en los que se llevan a cabo procesos de metaevaluación propiamente dichos. La literatura refiere algunas experiencias, la mayoría de ellas del campo educativo y en los Estados Unidos.

La metaevaluación se distingue de la evaluación en general por su objeto, pero los pasos metodológicos son los mismos: requiere del diseño previo de un modelo inclusivo de las dimensiones y variables que se considerarán, las técnicas elegidas, los instrumentos y las pautas que uniformen los criterios del equipo metaevaluador.

Se diferencia de esa otra acción de las oficinas de evaluación de las agencias financiadoras o de cooperación, cuando formulan observaciones o comentarios generales sobre los informes y las metodologías de evaluación enviados por los programas o proyectos apoyados.

Conforme se anticipó, la metaevaluación constituye un nivel de la evaluación, cuyo objeto es el propio modelo y proceso evaluativo. Dado que ese nivel está "por encima" de la evaluación propiamente dicha, hay ciertas recomendaciones a tener en cuenta, para que los resultados de esa actividad sean realmente aplicables y útiles.

- En el diseño del modelo para la metaevaluación debe darse intervención a los miembros del equipo de evaluación (ya que será en gran parte el "objeto" a evaluar).
- Los integrantes del equipo de la metaevaluación deben acreditar perfiles de mayor experiencia, idoneidad y prestigio que los del equipo de la evaluación, dado que evaluarán el desempeño, los informes y las conclusiones de esos últimos.

- Es conveniente que la metaevaluación acompañe las diferentes etapas del proceso de evaluación, para permitir realizar ajustes oportunos a la direccionalidad del proceso evaluativo, tal como se aconseja para este último en cuanto a acompañar el proceso de gestión de los programas o proyectos.
- Deberá ponerse especial cuidado en las formas de efectuar las recomendaciones y la "devolución" de los hallazgos, ofreciendo momentos y espacios para que los evaluadores evaluados puedan discutir e introducir modificaciones y ajustes a los informes de la metaevaluación (tal como se aconseja para todas las evaluaciones).

Por lo demás, las metodologías coinciden con las de la evaluación en general.

### La multidimensionalidad de la evaluación

Una conclusión que puede extraerse de las clasificaciones anteriores es que ponen el acento en aspectos similares desde perspectivas diferentes. Así, resulta difícil diferenciar en la práctica la evaluación *ex post* de la sumativa o de la de impacto, o bien la evaluación de procesos de la formativa o del monitoreo. En realidad, abarcan el universo de aplicación de la evaluación, organizándolo para responder alternativamente a diferentes preguntas: para qué se evalúa, quién evalúa, cuándo evaluar y qué se evalúa.

Los desafíos que el desarrollo de la práctica fue planteando frente a las limitaciones de las modalidades más tradicionales de evaluación contribuyeron a su enriquecimiento. Ahora se presta mayor atención a las caracte-

rísticas de los potenciales usuarios de los hallazgos evaluativos. La incorporación de diferentes perspectivas –clientes, financiadores, la misma comunidad– ha ampliado el horizonte de aspectos a considerar; se concibe la evaluación cada vez más como un proceso continuo de autoaprendizaje, considerando tanto los procesos como los resultados, y se procura involucrar de forma periódica a quienes están vinculados con la ejecución.

Los ejes que dieron sentido originalmente a la conceptualización de los distintos tipos de evaluación se van borrando y se tiende a encararla como una actividad más integral.

En los últimos cien años la evaluación ha evolucionado desde un enfoque unidimensional llevado a cabo por un solo evaluador, que generalmente provenía del campo técnico que se estaba evaluando, hasta convertirse en una actividad multidimensional concentrada en múltiples niveles de un proyecto, en la que participa una gran variedad de generadores de productos de evaluación, así como usuarios de los resultados de una evaluación. En la actualidad, en lugar de basarse únicamente en los puntos de vista de un solo evaluador en distintas etapas, un grupo mucho mayor de involucrados ayuda a dar forma a la evaluación y a interpretar sus hallazgos [...] comienza a evolucionar de una actividad en su mayor parte *ex post* para convertirse en una herramienta utilizada cada vez con más frecuencia para fines de monitoreo. La meta de la evaluación ha pasado de la auditoría y culpabilidad a la meta actual del entendimiento y el aprendizaje en función de experiencias adquiridas (Banco Interamericano de Desarrollo, s/f).

El mayor desafío en la actualidad consiste entonces en adoptar abordajes metodológicos y diseños que sean a la vez rigurosos e integrales, compatibles con los requerimientos derivados de la multiplicidad de perspectivas que se ponen en juego durante el proceso evaluativo.

## Los abordajes metodológicos

*Que “saber” signifique algo más que “estar cierto”, supone una situación comunitaria en que varios sujetos, o yo mismo en varios momentos, puedan comprobar lo mismo. En la actualidad instantánea de mi conciencia propiamente nada sé, sólo tengo certezas. Para saber algo, necesito salir de mí y compartir otro punto de vista. El saber requiere la comunidad de un sujeto con otros.*

LUIS VILLORO (1996)

La diversidad está a la orden del día. Así como existen diferentes tipos de evaluación según cuándo, para qué, qué y quién evalúa, ha proliferado en las últimas décadas una multiplicidad de prescripciones, denominadas “modelos evaluativos”<sup>1</sup> acerca de qué es y cómo debe hacerse la evaluación, enfatizando diferentes aspectos según las prioridades y preferencias de los autores. El origen de estos modelos, abordajes o enfoques (como también se los denomina) puede rastrearse directamente a partir de las diferentes visiones de sus proponentes acerca de la naturaleza de la evaluación. Sus orientaciones filosóficas

1. Briones (1991) denomina “modelo de evaluación” a “un esquema o diseño general que caracteriza la forma de la investigación evaluativa que ha de realizarse, las técnicas o procedimientos para la recolección y análisis de la información, el conocimiento final que se desea obtener y los usuarios principales de los resultados del estudio”. Como se desprende del capítulo 5, ese significado no es tan diferente del que nosotras adoptaremos.

e ideológicas, sus predilecciones metodológicas y sus preferencias prácticas los llevan a proponer diseños, métodos de recolección y análisis, y técnicas interpretativas diferentes.

### **Evaluación cuantitativa y cualitativa. Un debate histórico**

Uno de los ejes de los abordajes metodológicos que ha generado mayor discusión, como veremos no del todo cerrada, es el que se refiere al uso de los métodos cuantitativos o cualitativos. Merece que nos detengamos en este debate pues ha atravesado la historia de la metodología de la investigación en ciencias sociales e impregnado fuertemente la de la evaluación, tanto desde el punto de vista conceptual como en la práctica. De hecho, se han llegado a identificar directamente los abordajes con la índole de los métodos de investigación aplicados denominándolos, de modo poco afortunado en nuestra opinión, "evaluación cuantitativa" y "evaluación cualitativa".

Esquemáticamente, suele asociarse la evaluación o investigación cuantitativa al paradigma positivista de la ciencia y se la identifica con el enfoque hipotético-deductivo. Generalmente se centra en verificar hipótesis específicas (parte de una teoría mayor) a través de pruebas empíricas y privilegia la explicación de los fenómenos desde el punto de vista de los investigadores/evaluadores. Utiliza diseños experimentales o cuasi experimentales. Enfatiza la estandarización, la precisión, la

objetividad y la confiabilidad de la medición así como la replicabilidad y la generalización de los resultados. Predomina el uso de los métodos de análisis estadístico. Se preocupa no sólo por producir datos numéricos sino por generar números adecuados a pruebas estadísticas.

Por su parte, los enrolados en la evaluación o investigación cualitativa intentan recuperar el contexto y las dimensiones humanas del fenómeno en estudio. Generalmente esa evaluación se realiza en medios "naturales" (investigación naturalista) como escuelas o vecindarios; el investigador/evaluador es el principal "instrumento" de recolección y el que analiza los datos; enfatiza la descripción "gruesa", aquella que ilumina los patrones cotidianos de conducta y su significado desde la perspectiva de quiénes se estudian; se orienta a los procesos sociales y no sólo se concentra en los resultados; emplea múltiples técnicas de recolección de datos, algunas más otras menos estructuradas; y fundamentalmente utiliza el lenguaje de las palabras y un enfoque inductivo en el análisis de datos.

Se plantean a continuación, en forma esquemática y a modo de contrapunto, las principales diferencias que suelen destacarse entre estos abordajes evaluativos.

EVALUACIÓN CUANTITATIVA	EVALUACIÓN CUALITATIVA
Paradigma positivista/empírico	Paradigma fenomenológico/constructivista
Predominio de deducción	Predominio de inducción, comprensión, interpretación
Diseños experimentales o cuasi experimentales	Investigación naturalista
Perspectiva desde afuera (punto de vista del observador)	Perspectiva desde adentro (punto de vista de los actores)
Lenguaje numérico	Lenguaje verbal
Énfasis en los aspectos "objetivos", observables y cuantificables	Énfasis en los aspectos subjetivos: captación y comprensión de actitudes, conductas, valores y motivaciones internas
Análisis estadístico	Análisis de contenido, estudio de casos
Resultados generalizables	Resultados limitados en su generalización

Los desacuerdos entre los defensores de la "evaluación cuantitativa" y los de la "evaluación cualitativa" se han centrado en cuestionarse mutuamente el uso de determinados métodos y técnicas y la capacidad de los datos relevados por una u otra para explorar, explicar, comprender y predecir los fenómenos sociales.

De todos modos, merece destacarse que estos desacuerdos son en parte producto de la evolución de la metodología de investigación social en el último medio siglo, dentro del marco de los paradigmas científicos pre-valetientes en cada momento.

Los *paradigmas* constituyen perspectivas generales o teorías explicativas de la realidad que determinan cómo

percibir e intentar comprender la verdad. Incluyen y combinan creencias ontológicas (¿qué tipo de sujeto es el ser humano?, ¿cuál es la naturaleza de la realidad?, ¿existe en los hechos o en el pensamiento?), epistemológicas (¿cuál es la relación entre el investigador y el sujeto/objeto estudiado?) y metodológicas (¿cómo aprehendemos el mundo y avanzamos en su conocimiento?).

Un paradigma científico puede definirse como un principio de distinciones-relaciones-oposiciones fundamentales entre algunas nociones matrices que generan y controlan el pensamiento, es decir la constitución de teorías y la producción de los discursos de los miembros de una comunidad científica determinada [...]. Se convierte así en un principio rector del conocimiento y de la existencia humana. Hay que considerarlo como un mapa útil, como una aproximación adecuada y como un modelo para la organización de la información conocida (Martínez Miguelez, 1993, págs. 62-3).

Reiteramos entonces que la práctica de la evaluación estuvo fuertemente influenciada por el paradigma o los paradigmas científicos predominantes que establecen qué es lo importante, lo legítimo y lo razonable en cada época.

Hasta fines de los sesenta, la formación en materia de investigación de quienes se desempeñaban como evaluadores provocó un claro predominio del enfoque experimental considerado como el único que podía responder a los requerimientos del método científico, entendido según el enfoque clásico como acumulación del conocimiento por investigaciones sucesivas, formas "objetivas" de conocer, y con énfasis en la medición como sinónimo de precisión y en la comprobación de hipótesis mediante pruebas empíricas. El rigor en los procedimientos es

aquí condición para encontrar la “verdad”. Ello llevó a la aplicación de diseños cuasi experimentales para captar la relación causal entre programa y resultado, es decir, para evaluar si las metas y los impactos logrados por los programas y proyectos sociales son imputables a ellos o son consecuencia de otros factores. Estos diseños requieren como mínimo una medición de la situación en el inicio –línea de base– y otra al final del programa –*ex post*– o del uso de grupos de control. Involucran una serie de supuestos de homogeneidad, estabilidad y racionalidad en la implementación, así como de confiabilidad de los registros, difíciles de verificarse o mantenerse durante períodos prolongados en el contexto de los programas y proyectos sociales. También pueden plantearse problemas éticos al requerir la creación de condiciones experimentales para contar con un grupo de control, lo que puede significar la exclusión de individuos de los beneficios de un programa o proyecto social.

Sin embargo, en los años setenta comenzó a tener mayor aceptación una variedad de métodos y técnicas cualitativas que en los sesenta habían sido considerados inaceptablemente “blandos” y no científicos. Hubo avances en la multidisciplinariedad, ya que se multiplicaron los métodos y las estrategias de investigación alimentados por diversas disciplinas, como la historia, la etnografía, la semiótica, la antropología, la sociología. Los evaluadores adoptaron cada vez más las formas subjetivas de conocer, incorporando valores pluralísticos y utilizando métodos cualitativos. Por un lado, se produjo una especie de deslumbramiento sin reconocer cabalmente la complejidad de esas modalidades y las competencias requeridas por parte de quienes las utilizaban, con lo cual también hubo

una “malversación” de tales métodos y técnicas. Y, por otro lado, sus detractores cuestionaron el abandono de lo que ellos suponían objetividad y rigor científico.

Es a fines de la década del setenta cuando comenzaron a percibirse los beneficios de la integración o la complementación de ambos enfoques dentro de un mismo estudio evaluativo aunque, como veremos enseguida, fueron diversas las formas en que se dio ese enlace. En la última mitad de la década del noventa el debate parece haber perdido intensidad, al menos entre los expertos. La mayoría de los evaluadores acepta hoy ambos enfoques (cuantitativo y cualitativo) como válidos. El énfasis que se ponga en uno u otro depende de la situación o de los momentos de evolución de un proyecto o programa y pueden complementarse o integrarse en un mismo diseño. Es así que se admiten múltiples combinaciones posibles, que se potencian y pueden aumentar la confiabilidad y credibilidad de los resultados.

En realidad, pareciera que cada momento histórico en la evolución de la investigación social y la evaluación aún operara en el presente, básicamente como un conjunto de prácticas que los evaluadores toman o no en función de los diferentes intereses o propósitos y de las particulares situaciones o circunstancias.

### ¿Evaluación cuantitativa o cualitativa? Ésa no es la cuestión

Una de las críticas más severas a la utilización exclusiva de enfoques cuantitativos es que sólo miran aquellas variables que se pueden medir y así pierden de vista el

“bosque” al que pertenecen y que les da sentido.<sup>2</sup> Si bien, como dijimos anteriormente, la discusión entre evaluadores parece estar superada y la corriente actual se orienta hacia la legitimación de la diversidad, el problema se vuelve a suscitar frente a quienes demandan las evaluaciones: los gerentes de programas o proyectos, los decisores políticos, los financiadores; ellos siguen solicitando a los evaluadores que usen diseños cuasi experimentales, técnicas cuantitativas “que muestren cifras concretas y gráficos claros”, mientras que el análisis de la estructura del programa o proyecto, sus mecanismos de funcionamiento, las opiniones y demandas de los beneficiarios son tomadas con recelo, o incluso no tenidas en cuenta como hallazgos de una evaluación. Desearían —y es lícito— tener en muy pocas páginas (dos o tres) y unos pocos cuadros con cifras muy resumidas, el reflejo de los resultados de un programa o proyecto, la visión de los logros, el éxito de la gestión y también, por qué no, los fracasos u obstáculos que signaron el desempeño, pero raramente les interesan los procesos y las explicaciones. Debemos reconocer que muy pocas veces los evaluadores cumplimos en satisfacer las necesidades de nuestros demandantes. Lejos de desestimar sus pretensiones, ellos nos dan un motivo importante para incluir en el diseño de las evaluaciones los enfoques cuantitativos y cualitativos, considerando tanto los

2. Resulta muy ilustrativa al respecto una nota sin firma aparecida en el diario *Clarín*, del 30 de mayo de 1999, en relación con la evaluación de la calidad educativa y los consecuentes *rankings* de las escuelas. Dice un subtítulo de la nota: “Medir o evaluar. [...] Cuando se mide se asigna un puntaje o valor en una escala; evaluar supone elaborar un juicio acerca del porqué de ese puntaje [...] en general, las pruebas que toma la Nación ponen más el acento en medir que en evaluar”.

procesos como los resultados, brindando indicadores sintéticos y numéricos sobre resultados, así como explicaciones que reflejen los procesos.

La cuestión sobre qué es mejor y cuándo usar métodos cualitativos y cuándo cuantitativos, entendiendo por tales el conjunto de procedimientos y técnicas para la recolección de la información y el procesamiento y análisis de los datos, no pasa en realidad por la defensa de uno en desmedro del otro: cada uno tiene sus potencialidades y sus limitaciones y éstas se abordarán brevemente en el capítulo sobre técnicas. Se trata más bien de reflexionar acerca de su utilidad y de los beneficios y costos de combinarlos e integrarlos en el diseño evaluativo que se plantee. Sin pretensión de agotar los argumentos, podemos señalar al menos que esta combinación permite la consecución de objetivos múltiples atendiendo tanto a los procesos como a los resultados del programa, enriquece los hallazgos evaluativos profundizando en los motivos de las asociaciones que pueden encontrarse entre variables cuantitativas y sugiriendo hipótesis para explicar la variabilidad entre individuos o grupos, permite abordar la cuestión de las creencias, motivaciones o actitudes de la población, difíciles de ser reflejadas cuantitativamente, dando vida a los datos. También facilita la triangulación (véase la pág. 95) y acrecienta la comunicabilidad de los resultados. Por otro lado, no puede desconocerse que esta combinación tiene requerimientos en términos de tiempos y costos y, sobre todo, de habilidades para el análisis que no siempre son fáciles de alcanzar. Sin embargo, como veremos en los últimos capítulos de este libro, mediante un proceso de construcción participativa del modelo de evaluación, es posible instalar las capacidades necesarias a tal efecto en la organización ejecutora

del proyecto. De todos modos es tarea de los evaluadores presentar diseños integradores y demostrar la pertinencia de la combinación de técnicas cuantitativas y cualitativas en un proceso evaluativo.

Al decir de Michael Patton (1987):

El arte de la evaluación implica crear un diseño y relevar la información que es apropiada en una determinada situación y en el marco de una política determinada.

O sea que no hay una “receta” que en sí misma determine qué abordaje o qué método usar. Ante cada evaluación se deberán tener en cuenta las demandas de quienes contratan la evaluación, las necesidades de los decisores, las de quienes financian el programa o proyecto, las de los técnicos que lo llevan adelante y las de los beneficiarios, así como la propia visión del evaluador acerca del problema y su experiencia –preferencia– en el uso de determinadas técnicas. Así será posible construir un diseño evaluativo que dé respuesta apropiada a ese determinado contexto. Dicho diseño, por otra parte, seguramente será diferente en la próxima evaluación que asuma ese mismo evaluador.

Por lo tanto, el abordaje metodológico no se limita a una cuestión de métodos.

Las metodologías de evaluación constituyen marcos de referencia coordinados de supuestos filosóficos (acerca del mundo, la naturaleza humana, el conocimiento, la ética), integrados a perspectivas ideológicas sobre el rol y propósito de la investigación social en los procesos decisorios de programas y políticas sociales, acompañados por posturas valorativas acerca de los fines deseados de los programas y de la investigación y finalmente –y es lo menos– por preferencias acerca de métodos complementarios (Greene, 1994:531).

De este modo sería preferible reservar los atributos de cuantitativo y cualitativo para referirnos a diferencias entre técnicas de recolección, métodos de análisis e índole de los datos<sup>3</sup> que se obtienen, más que a distintas formas de abordajes evaluativos. También sería conveniente utilizar, como Patton, el término “estrategia” para aludir a la naturaleza de los métodos y las técnicas a utilizar. En su forma ideal una *estrategia cuantitativa* sería aquella que apunte a obtener datos numéricos susceptibles de ser sometidos a análisis estadístico y una *estrategia cualitativa* sería aquella que recurra fundamentalmente a otro tipo de técnicas de recolección<sup>4</sup> (como la observación en terreno, el análisis de documentación, las entrevistas en profundidad y, en general, todo tipo de técnicas participativas), dando lugar a información susceptible de análisis de contenido o histórico, o de casos). Las estrategias mixtas serían aquellas que utilizan combinaciones de ambos grupos de técnicas y métodos en una misma evaluación.

En el sentido amplio que se ha dado al término “abordaje evaluativo”, nuestro énfasis, como se verá más adelante, es acorde con la concepción de evaluación que hemos presentado en el primer capítulo: apunta a acompañar el proceso de gestión en todas sus etapas, contribu-

3. Si consideramos, como Galtung, que un dato es la forma de registrar una respuesta, el dato será cuantitativo o cualitativo según cómo se registre. Así, una opinión (respuesta) se convierte en un dato cuantitativo si se la registra sobre la base de categorías preestablecidas para luego establecer frecuencias, y es un dato cualitativo si se la registra en su expresión textual. Así, un mismo instrumento (por ejemplo, un cuestionario con preguntas abiertas y cerradas) puede recoger datos cuantitativos y cualitativos.

4. Véase el capítulo 4 de este libro.

yendo al aprendizaje institucional y a una retroalimentación permanente entre evaluación y programación. Incluye el uso de una multiplicidad de técnicas, tanto cualitativas como cuantitativas, predominando aquellas que permiten incorporar la perspectiva de los actores involucrados en los procesos de programación y de ejecución, incluyendo a los beneficiarios, usuarios o destinatarios de las acciones.

Veremos en el capítulo 5 que un paso crucial en la construcción del modelo evaluativo, en el sentido de diseño que le damos en este libro, es la elección de las dimensiones evaluativas, mientras que las variables, los indicadores y las técnicas y los instrumentos a utilizar son elementos subordinados, dependientes de cuáles sean esas dimensiones. La determinación de las dimensiones del modelo es una parte muy creativa del proceso evaluativo, pues es lo que divide conceptualmente el todo que se pretende evaluar, en una forma diferente, para encontrar un nuevo sentido que posibilite otra comprensión y permita la transformación de la realidad considerada; eso no puede lograrse eligiendo variables e indicadores sino con otro nivel de abstracción, más global, que abarque el todo mediante las dimensiones que lo integran. Ya no se trata sólo de acercarnos a los fenómenos con variables medibles o apreciables mediante indicadores; se trata de cuál es la concepción del todo en sus partes, para entonces interesarnos por cuáles variables lo representan y explican. Diciendo lo mismo en otras palabras: para afrontar una evaluación no bastará con “elegir los indicadores apropiados”, aunque ése constituya un paso necesario; se requerirá tener una concepción “teórica” acerca del “objeto de la evaluación” y sobre la realidad donde éste se ubica.

### **Una cuestión central: el rigor metodológico**

Es importante recordar aquí que el resultado de toda evaluación debería acrecentar el conocimiento que se tiene sobre el programa o proyecto, sobre el problema que dio origen a la existencia del mismo y sobre las modalidades adecuadas para la resolución de este problema. Como hemos dicho anteriormente, la evaluación debe ser una instancia de aprendizaje más que de control. En ese sentido podríamos enfocar entonces el tema del abordaje evaluativo alrededor de cómo logramos conocer más y de qué modo incorporamos el saber.<sup>5</sup>

Ello remite a la cuestión del rigor metodológico, es decir, a los parámetros vigentes relativos al proceso de producción de conocimiento para que el que se obtenga sea considerado científico.

La evaluación fundamenta sus juicios en el análisis de un conjunto sistemático de datos, los cuales no existen ni surgen de forma natural o espontánea. Primero hay que poner en claro cuáles serán las dimensiones o los grandes aspectos a considerar; luego, definir las variables y los indicadores correspondientes, y sólo entonces deben obtenerse los datos pertinentes usando determinadas técnicas e instrumentos. Para recoger evidencia válida y confiable, es muy importante la decisión de cuáles técnicas e instru-

5. “La ciencia no tiene que estar limitada forzosamente a mediciones y análisis cuantitativos. Debemos considerar científica cualquier manera de abordar el conocimiento siempre que satisfaga dos condiciones: que el conocimiento se base completamente en observaciones sistemáticas y que se exprese desde el punto de vista de modelos coherentes, aunque sean limitados y aproximativos” (Martínez Miguelez, 1993:178).

mentos se utilizarán en función de las respectivas variables e indicadores seleccionados. Se requiere utilizar técnicas e instrumentos adecuados y a la vez poner sumo cuidado en los procesos de aplicación, de modo que la información tenga *validez interna y externa*,<sup>6</sup> o sea que se refiera a las dimensiones y variables que nos interesan y no a otras, que midan aquello que procuran medir en el caso de cuantificaciones, y que permitan apreciar y emitir juicios fundamentados sobre aquello que se procura comprender o explicar en el caso de cualificaciones. Se requiere *precisión* en los instrumentos, es decir que deben estar diseñados de forma tal que pueda asegurarse que se refieran a lo que importa que se refieran, dejando de lado lo no relevante. Se requiere asimismo que el instrumento sea suficientemente *confiable*, es decir que tenga la capacidad de ofrecer resultados similares o comparables al ser aplicado reiteradas veces y por diferentes evaluadores. Sin duda que para asegurar todo eso deben pautarse procedimientos ajustados en los diseños así como en las aplicaciones. Surge entonces la necesidad de capacitar y adiestrar a los actores que intervendrán en las etapas de diseño y que aplicarán los procedimientos e instrumentos evaluativos.

Sin embargo, el rigor no tiene que ver exclusivamente con la pertinencia del uso de determinada técnica para apreciar o medir determinado fenómeno ni con el grado de confiabilidad y precisión de los datos obtenidos.

6. "La diferencia principal entre validez interna y externa es que la primera expresa una relación lógica, mientras que la segunda expresa una relación empírica; la validez externa consiste en poner a prueba una hipótesis" (Zetterberg, 1965). En la primera se trata de coherencia y en la segunda se trata de la posibilidad de extrapolar o generalizar.

Por un lado, el rigor metodológico en la evaluación de un programa o proyecto va más allá de la decisión acerca de las técnicas y los instrumentos a usar: importa desde la propia identificación del problema que se abordará, de la adecuada determinación de los principales ejes o dimensiones que lo componen, de la selección de las variables relevantes en el marco de cada dimensión y de los indicadores apropiados para reflejar las variables elegidas; ese rigor en la etapa de la formulación deberá luego ser reforzado en los pasos de la aplicación, con procedimientos apropiados y lo más estandarizados que sea posible, lo cual supone actividades concretas de nivelación y homogeneización de los equipos evaluativos, sobre todo de los que intervienen en el trabajo de campo. Además, deberá mantenerse el rigor en el momento de procesar, sistematizar y analizar la información obtenida, así como en la preparación de los informes para la devolución de los resultados.

Por otro lado, uno de los aspectos positivos que introdujo la discusión entre métodos cuantitativos y cualitativos ha sido la resignificación de las cuestiones de la validez y confiabilidad, asociando éstas a los instrumentos utilizados y agregando la consideración de otros aspectos vinculados a los valores de los investigadores/evaluadores y de los "investigados", incluyendo así el contexto de producción del conocimiento, y la ética. En síntesis, se han sometido a revisión los criterios tradicionales de rigor y se están incorporando otros.

Es necesario comprender que los métodos cualitativos pueden asegurar el rigor aunque de modo diferente del de los métodos cuantitativos. La explicitación de los criterios de selección de participantes e informantes (más

que la búsqueda de representatividad estadística), la explicitación de los valores e intereses del evaluador y del modo en que el contexto incide en ellos, la socialización del proceso de análisis e interpretación, constituyen formas de asegurar el rigor. El *consentimiento informado* (la aceptación de la gente a responder o participar de una investigación o evaluación y además la transparencia en cuanto a los procedimientos y los derechos y deberes de todos los que se involucran en el proceso), el *anonimato* (más que la confidencialidad) y el cuidado en *minimizar las asimetrías* en la relación entre evaluador y otros actores, son formas de introducir la ética en los procesos de producción de información.

Cualesquiera que sean los métodos que se utilicen, los evaluadores son actores privilegiados, y por lo tanto la *validez*, la *precisión* y la *confiabilidad* de los datos dependen en gran medida de su capacidad, sensibilidad e integridad, lo cual nos lleva a la cuestión de la *subjetividad* en el proceso evaluativo.

Más adelante proponemos que la corrección de la subjetividad individual —que produce distintas apreciaciones sobre los mismos eventos— se encuentra en la “*intersubjetividad*”,<sup>7</sup> lograda mediante la incorporación de los diferentes actores en distintos momentos del proceso evaluativo —*multiactoralidad*—, lo cual genera espacios de

7. Habermas (1982) menciona procedimientos de “creación de campos epistemológicos de intersubjetividad”, contraponiendo este tipo de generación de conocimiento al que se realiza de forma aislada o individual y que corre el riesgo de la subjetividad individual, no sometida a crítica. Él, como otros muchos autores, sostiene que la ciencia, aun en el enfoque positivista, carece de objetividad absoluta, lo cual no impide procurar el rigor.

confrontación y reflexión conjuntas, para permitir acuerdos o consensos sobre distintos aspectos y, en particular, sobre los juicios valorativos fundamentados.

Por otra parte, si consideramos la evaluación como la aplicación de ciencia y conocimiento para valorar la calidad de determinado programa o proyecto, los aportes de diferentes disciplinas en el proceso evaluativo incrementan su poder explicativo y resolutorio, por lo cual es relevante la conformación de equipos evaluativos *multidisciplinarios*.

Una estrategia actualmente muy difundida para asegurar el rigor y corregir la subjetividad es la *triangulación*.<sup>8</sup> Ello supone la utilización de diferentes técnicas y fuentes para evaluar los mismos fenómenos o aspectos de la realidad a través de operaciones convergentes, en el sentido de síntesis y complementación metodológica. También se habla de triangulación cuando la evaluación es realizada por un grupo en el que “se cruzan” los criterios y puntos de vista de cada uno de los evaluadores. Con ello se procura mayor *confiabilidad* de la información obtenida, mayor *entendimiento* de los fenómenos bajo estudio, *reducción de sesgos propios de cada técnica, fuente y profesional*, y la *validación de las apreciaciones evaluativas*.

En síntesis, entendemos que el rigor no se limita a la elección de las técnicas o a determinar los pesos relativos

8. La metáfora de la triangulación se deriva del uso —con fines militares o de navegación— de múltiples puntos de referencia para identificar con mayor precisión la posición de un objeto alejado. En las ciencias sociales se intenta algo similar: mejorar la exactitud o precisión de los juicios y aumentar la confiabilidad de los resultados a través de la recolección de datos sobre el mismo fenómeno, desde diferentes perspectivas o técnicas.

de los métodos cualitativos y cuantitativos, sino que depende más de la calidad de las decisiones y conductas que los investigadores/evaluadores adoptan a lo largo de los diferentes momentos por los que se atraviesa en el proceso de investigar o de evaluar. Es así que las decisiones y el rigor que se aplique en los momentos de:

- definición del problema a evaluar,
- elección y explicitación del marco conceptual,
- generación de hipótesis,
- selección de técnicas,
- trabajo de campo,
- procesamiento, análisis e interpretación de los datos y resultados obtenidos,

influirán en la validez de los juicios apreciativos para corroborar, precisar o rechazar las hipótesis explicativas/evaluativas, en la comprensión del objeto de estudio y, en definitiva, en las posibilidades de intervenciones adecuadas que permitan transformar la situación evaluada.

### **Actores y audiencias**

Si aceptamos que los programas y proyectos sociales son el resultado de un conjunto de decisiones, podemos también aceptar que la evaluación de programas está atravesada por decisiones políticas en términos de definición de prioridades sociales, distribución de recursos y, en última instancia, de poder. Es en esta línea de pensamiento que las preguntas evaluativas sobre la importancia de un programa o proyecto, su significado, el de

sus logros y de la efectividad alcanzada, reflejan un proceso de toma de decisiones y establecimiento de prioridades. También los resultados evaluativos forman parte de definiciones técnico-políticas y no son simples deducciones o conclusiones abstractas, descontextualizadas de la realidad en la que el programa o proyecto se desarrolla.

El equipo evaluador describe, infiere y transmite un conocimiento práctico adaptado a cada situación. Pero, ante todo, su tarea es valorar y, hasta la selección de estándares o parámetros que se usarán para contrastar, estará determinada por un contexto. De ese contexto participa una variedad de actores con diferentes intereses y perspectivas acerca del programa.

Lo que diferencia un abordaje evaluativo de otro no son los métodos sino más bien las preguntas, quiénes plantean esas preguntas y qué valores se promueven (Greene, 1994:533).

Nosotros agregamos a eso la relevancia que tiene la determinación de las dimensiones de la evaluación, es decir, cuáles ejes conceptuales serán los que estructurarán el modelo a utilizar, pues son tales dimensiones las que definen la “teoría” subyacente del programa o proyecto a evaluar.

Los actores participantes en el diseño de una evaluación y el rol que desempeñan en la determinación de tales dimensiones y la formulación de las principales preguntas evaluativas es lo que influirá en las distintas formas de abordar la evaluación de un proyecto.

El primer paso de cualquier evaluación estará dado por una “negociación inicial” en la que el evaluador acordará sus límites y alcances con los decisores políticos

o responsables máximos del programa o proyecto y con los representantes de los entes financiadores (en el caso de que puedan distinguirse) teniendo en cuenta diversos factores, en particular los fondos que se le asignen. Son estos *actores* los que determinarán el contexto específico en el que comienza el proceso evaluativo.

A partir de allí, el carácter del abordaje metodológico le da en primera instancia el punto de vista del evaluador,<sup>9</sup> cuyo rol como actor es relevante. Son importantes en ese sentido su ubicación, su concepción, su origen y trayectoria disciplinaria, sus vinculaciones con otros actores, etc. Cada evaluador, de acuerdo con su particular cosmovisión, ubicación, actitudes, experiencia e intereses, releva aquellos aspectos que considera significativos y frente a los cuales está sensibilizado. Como vimos, eso introduce la cuestión de la subjetividad en el proceso evaluativo, sobre todo en el caso de que se trate de un evaluador individual y no de un equipo. Esta consideración de la carga ideológica que el evaluador trae consigo está bastante alejada del paradigma clásico que apela al supuesto de la objetividad y la neutralidad valorativa del investigador, evaluador u observador, como alguien situado afuera del fenómeno o realidad en estudio, no “contaminado” por esa realidad. Ya tales cuestiones no son materia de discusión seria en la filosofía de las ciencias, puesto que en las propias ciencias físicas y naturales

9. Cuando se menciona al “evaluador” deberá tenerse en cuenta que éste puede ser un individuo o bien tratarse de un equipo, en cuyo caso interesarán los factores mencionados para cada uno de sus miembros, así como las características multiprofesionales y multiactorales del grupo y sus modalidades de funcionamiento.

se reconoce la influencia del observador en el campo de experimentación y viceversa.

Sin embargo, aún subsisten formas de encarar los procesos evaluativos vinculadas a las modalidades tradicionales de evaluación, que ponen el énfasis en la supuesta objetividad científica, sobre cuya base el evaluador “llega” al proyecto con un modelo predeterminado, releva la información que le interesa, “se va” para analizar los resultados y presentar sus conclusiones, manteniéndose en todo momento “asépticamente” fuera del programa. Sin ir tan lejos, está el evaluador que, si bien se involucra activamente en el funcionamiento del proyecto o programa, intentando entender su dinámica, elabora aisladamente tanto su modelo como sus instrumentos de recolección y sus conclusiones, es decir, se reserva el rol de la evaluación exclusivamente para sí, ya que es el “experto”.

Otra posibilidad, a la que adherimos y que estaría vinculada a la aceptación de la subjetividad intrínseca de los actores participantes, se encuentra en la construcción de espacios y momentos de “intersubjetividad” mediante la participación de los distintos actores comprometidos con el programa o proyecto objeto de análisis, en todas las etapas de la evaluación. Es en ese sentido que se hablaba anteriormente de la evaluación multiactoral y multidisciplinar.

Consideramos que la forma de compensar el peso del equipo evaluador y sus puntos de vista (por más multidisciplinar que sea), y a la vez hacer viable la aplicación de los procesos evaluativos para luego lograr la aceptación y la apropiación de los resultados y recomendaciones por parte de los miembros del programa o proyecto, es justamente a través del involucramiento en el proceso eva-

luativo de todos los actores intervinientes y significativos. Todos ellos deben tener posibilidades reales de intervenir en las decisiones relativas a qué dimensiones y variables tener en cuenta, cuáles indicadores son los más relevantes, qué técnicas e instrumentos de recolección utilizar, así como en la implementación de la evaluación, en sus correcciones y cambios, en el análisis y seguimiento de sus resultados, en las conclusiones extraídas de ellos y en formular (o corregir) las recomendaciones.

Es importante, entonces, realizar un exhaustivo relevamiento de quiénes son los actores clave en la toma de decisiones y en la ejecución del programa/proyecto, para incluirlos desde el inicio en el proceso de la evaluación.

En las temáticas sociales, los usuarios, destinatarios o beneficiarios, así como la comunidad en general, son actores de ineludible consideración en el proceso evaluativo.

En términos generales podemos distinguir al menos cuatro tipos de actores, que en principio deberían ser tenidos en cuenta, aunque en cada caso se deberá proceder a identificar y caracterizar a los que sean relevantes para cada proyecto o programa.

- Los funcionarios de las agencias financiadoras.
- Los responsables de la conducción político-técnica.
- El personal profesional y administrativo que lleva adelante la ejecución.
- Los miembros de las organizaciones intervinientes.
- La población objetivo, beneficiaria o usuaria.

Si bien el rol del evaluador sigue siendo clave, sólo a partir del trabajo mancomunado con los distintos actores nos parece posible llevar adelante una evaluación que lo-

gre superar los temores, las desconfianzas y las barreras que normalmente surgen en la organización o en el equipo del programa/proyecto que se pretende evaluar ante el hecho concreto de “sentirse evaluados”. Además, como ya se ha visto en el capítulo anterior, los miembros de esos equipos están involucrados en distintos aspectos de la marcha y es natural que se resistan a aceptar cambios aun cuando surjan de las recomendaciones de evaluaciones rigurosas.

La evaluación y la organización pueden ser términos contradictorios. La estructura de la organización implica estabilidad, en tanto que el proceso de evaluación inculca el escepticismo. La evaluación se refiere a la relación entre la acción y los objetivos, en tanto que la organización relaciona sus actividades con su programa y clientela (Wildavsky, 1972).

Por lo tanto, en la fase inicial será imprescindible el esfuerzo del equipo evaluador por conocer a los actores con quienes interactuará, así como su historia en tanto individuos y como miembros de un grupo u organización. También será fundamental el desarrollo de capacidades y habilidades específicas en los participantes, a través de la capacitación, el entrenamiento, el asesoramiento y la asistencia técnica, para facilitar la mirada evaluativa y el conocimiento del proyecto o programa desde esta perspectiva.

Podemos decir, en el sentido que venimos desplegando, que el rol del evaluador no es solamente generar capacidades en los actores del programa o proyecto para facilitar la instalación de una “cultura organizacional evaluativa” sino que, simultáneamente, irá aprendiendo e incorporando la información necesaria para la construc-

ción del modelo evaluativo. Para que refleje lo más acabadamente posible las necesidades y los requerimientos de los actores y del programa o proyecto, el evaluador desarrollará las grandes líneas, conceptualizará las dimensiones de la evaluación y sus principales aspectos, pero dará intervención a los actores en el diseño específico de dichas dimensiones. Pondrá sus habilidades metodológicas al servicio del conocimiento que los actores tienen del lugar y la realidad donde se desempeñan. En un proceso así encarado, los distintos actores expresarán los diferentes objetivos y metas que cada uno persigue con la evaluación, quedando explicitados sus propios puntos de vista.

Por ejemplo, en un proyecto que evalúa la calidad de la atención en los servicios de salud, los profesionales asistenciales suelen poner el énfasis en los aspectos científico-técnicos de la atención, y consideran el acto médico en relación con cada paciente; paralelamente suelen dejar de lado los aspectos de eficiencia (o de costos de la atención). Asimismo, serán heterogéneos dichos énfasis según las disciplinas y especializaciones profesionales.

En cambio, los gerentes o funcionarios de los niveles de conducción de los sistemas de salud enfatizan los aspectos de resultados a niveles poblacionales y la eficiencia, o sea, la relación de esos resultados con los costos (lamentablemente, a veces sólo consideran importantes los costos, o tan sólo los gastos).

Por su parte, los usuarios de los servicios hacen hincapié en los aspectos de trato interpersonal, de accesibilidad y de resultados concretos en la mejoría del estado de salud y bienestar a corto plazo (lo cual no es estrictamente lo mismo que enfatizar en aspectos científico-téc-

nicos), pero casi no consideran los aspectos de costos o eficiencia (aunque sí los de gastos en el caso de que tengan que realizarlos de su bolsillo). También debe tenerse en cuenta que habrá heterogeneidades en las expectativas (y, por ende, en la satisfacción/aceptabilidad), de acuerdo con las características culturales y socioeconómicas –los estratos de pertenencia– de los usuarios.

Todos deben ser incluidos en el proceso evaluativo. Cada uno de esos actores procurará orientar la evaluación hacia los aspectos que considera relevantes; la síntesis a la que pueda arribarse a partir de las instancias de confrontaciones y de consenso<sup>10</sup> otorgará mayor riqueza a la evaluación y colaborará en el uso posterior de sus hallazgos para la mejora del desempeño del proyecto/programa.

En resumen: la especificidad y la parte “noble” de la tarea que recae en el evaluador es la de aportar su experiencia metodológica a los demás actores, a la vez que realizar una síntesis superadora y creativa a partir de lo que ellos brindan.

Por otra parte, es necesario diferenciar a los *actores* participantes en la evaluación de las *audiencias* interesadas en sus resultados. Algunos coincidirán, pero ello no es así necesariamente. El modo de preparar los informes, organizar las presentaciones de los resultados y las formas de difusión de las conclusiones e incluso de especificar recomendaciones, seguramente deberá realizarse diferenciando y enfatizando distintos aspectos, puesto que

10. En el capítulo sobre técnicas se sintetizan algunas formas para efectuar confrontaciones y arribar a consensos, pero deberá profundizarse con técnicas específicas de dinámicas grupales y de negociación, entre otras.

*Olga Nirenberg, Josette Brawerman, Violeta Ruiz*

las inquietudes y preocupaciones de cada tipo de audiencia son distintas. Es relevante reconocer esa diversidad de intereses por parte de las audiencias, pues ello incidirá también en su aceptación de los resultados evaluativos y las recomendaciones.

Sin duda, los actores que se involucraron en los procesos evaluativos serán audiencias ineludibles, a las que habrá que efectuar "devolución" de los resultados para que puedan tener "derecho a réplica", es decir, a modificar las conclusiones y recomendaciones así como las fundamentaciones. Pero las audiencias son más amplias, por ejemplo:

- los pobladores o actores de la comunidad no directamente afectados por la implementación del proyecto o programa;
- las autoridades políticas inmediatas y mediatas;
- los responsables de programas similares en el mismo o en otros contextos;
- los funcionarios de la agencia que financia el programa (en los casos de que se trate de un programa subsidiado u apoyado por una agencia);
- la comunidad académica.

No deberán descuidarse los procesos de socialización de los hallazgos de la evaluación, particularizando según las diferentes expectativas e intereses de las audiencias.