

**La evaluación de políticas públicas en
Colombia: una mirada crítica a partir
de las prácticas evaluativas oficiales de
los programas de la "Red de Apoyo
Social"**

André-Noël Roth Deubel

André-Noël Roth Deubel

Profesor Asociado del Departamento de Ciencia Política de la Universidad Nacional de Colombia, Bogotá, desde 2006; y Director del Programa de Doctorado Interfacultad en Estudios Políticos y Relaciones Internacionales (2008-2009). Director del grupo de investigación "Análisis de las políticas públicas y de la gestión pública". Es politólogo, Magíster en Ciencias Políticas y Doctor en Ciencias Económicas y Sociales, mención Ciencias Políticas, por la Universidad de Ginebra (Suiza). Ha enseñado en el área de administración y de políticas públicas en el Departamento de Ciencia Política de la Universidad de Ginebra (Suiza); ha sido profesor y director del Departamento de Ciencia Política de la Universidad del Cauca (Colombia); así como profesor de Políticas Públicas invitado por universidades de Colombia: Universidad del Valle, Universidad Externado de Colombia y Escuela Superior de Administración Pública. Principales libros publicados: *Políticas públicas: formulación, implementación y evaluación*, Bogotá, Ediciones Aurora, 2002 (siete ediciones); y *Discurso sin compromiso: la política pública de derechos humanos en Colombia*, Bogotá, Ediciones Aurora, 2006. Ha publicado artículos relacionados con el tema en varias revistas.

Las comunicaciones con el autor pueden dirigirse a:
Departamento de Ciencia Política
Facultad de Derecho, Ciencias Políticas y Sociales
Universidad Nacional de Colombia
Av. Carrera 30 N° 45-03, Edif. 201
Bogotá, D.C. - Colombia
E-mail: anrothd@unal.edu.co

La evaluación de instituciones y de políticas públicas puede contribuir a la discusión acerca de la capacidad de las ciencias sociales para aportar conocimiento útil a la sociedad.

La evaluación de políticas públicas en Colombia: una mirada crítica a partir de las prácticas evaluativas oficiales de los programas de la "Red de Apoyo Social"*

Introducción

El tema de la evaluación de la acción pública -tanto de las instituciones como de las políticas públicas- está tomando una importancia creciente en toda América Latina. En Colombia, en el marco de la reforma del Estado emprendida desde los años 80, que pasó por la adopción de una nueva Constitución en el año 1991, progresivamente se ha instalado en el paisaje y el lenguaje públicos la idea de la necesidad de evaluar los resultados e impactos de las acciones públicas. Se pretende instaurar una cultura de la evaluación que permitiría, entre otras cosas, conocer mejor el desempeño institucional y los efectos de las políticas y programas, limitar la corrupción, aumentar la transparencia y, ante todo, dar elementos para mejorar la acción pública e informar a la ciudadanía (legitimación) y a las instituciones públicas de la eficacia y eficiencia de ésta.

La preocupación central en cuanto al control del gasto público y a la efectividad de la acción pública generó, en las últimas décadas del siglo XX, una fuerte demanda de herramientas de evaluación. En este contexto, las ciencias sociales y económicas han desarrollado una serie de herramientas y técnicas que pretenden lograr medir objetivamente las realizaciones de los gobiernos, y aportar, a la vez, elementos de juicio para tomar decisiones más racionales o más acertadas hacia el futuro. La evaluación de instituciones y de políticas públicas, por lo tanto, puede contribuir a la discusión acerca de la capacidad de las ciencias sociales para aportar conocimiento útil a la sociedad.

Recibido: 16-05-2009. Aceptado: 07-09-2009.

(*) Una versión preliminar fue presentada en el Seminario Internacional "Evaluación de las Instituciones Públicas en América Latina", organizado los días 12 y 13 de noviembre de 2008 en la Universidad de los Andes, Bogotá (Colombia), por el *Institut de Recherche et de Débat sur la Gouvernance* (IRG) de la Fundación Charles Léopold Mayer (Francia), y por el *Congreso Visible* (Universidad de los Andes).

Si bien la actividad es aún joven en Colombia, existe, en particular en los Estados Unidos, en los países del Norte de Europa y en la Unión Europea, una larga experiencia en materia de evaluación de políticas o programas públicos. De hecho, la evaluación de las políticas públicas se ha prácticamente constituido en una subdisciplina autónoma en el campo científico, con sus asociaciones profesionales, sus centros de formación y sus publicaciones.

A partir de Guba y Lincoln (1989) se presentan, en una primera parte, las características principales de las diversas perspectivas evaluativas desarrolladas en el curso del tiempo por esta subdisciplina y que han sido clasificadas por estos autores en cuatro generaciones. Para ellos, la estrategia de unas ciencias sociales empeñadas en seguir el mismo paradigma científico positivista (en lo epistemológico, ontológico y metodológico) que las ciencias naturales ha conducido a la evaluación, como a las ciencias sociales en general, a un camino sin salida. A pesar de los perfeccionamientos sucesivos ocurridos con las tres primeras generaciones, los autores consideran necesario romper con la perspectiva científica dominante si se desea realmente desarrollar una práctica de evaluación a la vez más efectiva y más democrática. Esta "cuarta generación" de evaluación se fundamenta en un paradigma científico constructivista, interpretativo o hermenéutico que promueve la deliberación razonada entre la diversidad de valores que se expresan en una sociedad o en torno a un problema social específico. En este sentido, la evaluación debe orientar las decisiones públicas no sólo desde una perspectiva técnica (*tecne*), como tendencia predominante, sino en una perspectiva ética (*phronesis*), es decir en valor. La evaluación de la acción pública podría ser, en este sentido, el lugar de intersección entre la perspectiva tradicional que privilegia la lógica causa-efecto o comando-control (las tres primeras generaciones) y la perspectiva práctica y deliberativa en valor (la cuarta generación).

En la segunda parte se argumenta, a partir del análisis de unas evaluaciones realizadas por el Departamento Nacional de Planeación (DNP), principal instancia de planeación y de evaluación del Estado colombiano, que el enfoque elegido no permite generar reales cambios sociales ni profundizar la democracia y la participación por su fuerte amarre a una concepción tecnocrática de la evaluación, que tiende a cumplir con una función de legitimación de las políticas del gobierno, alejándose así de su objetivo central.

La evaluación de primera generación se centra en la actividad de medición. En esta perspectiva, el papel del evaluador es el de un técnico, experto en crear, adaptar y aplicar técnicas e instrumentos de medición donde haya lugar.

1. Cuatro generaciones de concepciones y prácticas evaluativas

La evaluación de primera generación se centra en la actividad de medición. Su origen se debe a la preocupación inherente a la educación de medir el grado de aprendizaje y de la eficiencia de las técnicas de enseñanza de lo que una sociedad considera como fundamental que se inculque a sus futuros ciudadanos. La medición se hacía principalmente a través de pruebas de memoria. El ejemplo paradigmático de esta preocupación se encuentra probablemente en el desarrollo de las pruebas de medición del "cociente intelectual", desarrolladas inicialmente por el francés Binet para determinar la "edad mental" de los sujetos a fin de su clasificación. Rápidamente estas nuevas técnicas de medición se utilizaron para orientar y seleccionar las personas obligadas a prestar servicio militar durante la primera guerra mundial. Por ejemplo, en los Estados Unidos, una prueba de esa naturaleza, llamada Alpha, fue elaborada y administrada por la Asociación Americana de Psicología (APA, por sus siglas en inglés) a más de 2 millones de personas.

El desarrollo de estas perspectivas sicométricas, sociométricas y econométricas se debe a la existencia, según Guba y Lincoln (1989: 25-26), de dos factores. El primero, al desarrollo durante el siglo XIX de una ciencia *social* basada en los mismos presupuestos que la ciencia física o química (positivismo). El segundo, a la emergencia de un movimiento a favor de la gestión científica de la industria y el comercio considerando necesario aplicar criterios de eficiencia y de efectividad para el principal factor de producción: el trabajo humano.

Esta perspectiva tuvo su mayor desarrollo a partir de los años veinte. A los trabajos de determinación de los mejores métodos de trabajo realizados por Taylor, en EE.UU, y Fayol, en Francia, se sumaron los famosos estudios sobre los modelos de autoridad conocidos bajo el nombre de experimentos de Hawthorne. Rápidamente, esta preocupación relativa a la ciencia del *management* se trasladó al ámbito escolar. Se trataba de concebir a los estudiantes como una materia prima que la "planta" escolar debía "procesar". Este contexto facilitó el desarrollo de múltiples técnicas de medición de los resultados escolares y se empezó a usar como sinónimos los términos de medición y de evaluación. En esta perspectiva, el papel del evaluador es el de un técnico, experto en crear, adaptar y aplicar técnicas e instrumentos de medición donde haya lugar. Aunque esta perspectiva es antigua, es preciso anotar que sigue vigente.

La introducción del juicio en la evaluación marca el nacimiento de la evaluación de tercera generación. De este modo, el evaluador asume no sólo el papel de medidor y descriptor, sino de juez.

En la educación colombiana, los exámenes para la obtención del título de bachiller (llamado ICFES) o que culminan los estudios profesionales (ECAES), entre otros, son muestras características de esta perspectiva de evaluación de primera generación basada en la medición y la clasificación.

Si la primera generación de evaluación se centraba en los individuos como los sujetos centrales de la evaluación, la segunda generación de la evaluación emerge en el ámbito escolar cuando se consideró necesario revisar y adaptar los currículos a las nuevas realidades científicas, sociales y económicas. Para ello, era evidente que los datos existentes no servían para estos objetivos. La preocupación por saber si los nuevos currículos permitían alcanzar los propósitos de formación llevaron a considerar estos propósitos como *objetivos*. ¿Cómo saber si los estudiantes asimilaban lo que los profesores intentaban enseñar? Esta preocupación condujo a la invención por parte de Ralph Tyler, como miembro de la Oficina para la Investigación Educativa del Estado de Ohio (EE.UU.), de lo que se conoce como la *evaluación de programa*. Al mismo tiempo que se implementaba la reforma curricular, Tyler recolectó información relativa a la adquisición por los estudiantes de los objetivos definidos. Esta información fue utilizada para revisar los currículos mientras se desarrollaba la reforma educativa. Un proceso de este tipo hoy se llama evaluación formativa. Este tipo de evaluación es considerada por Guba y Lincoln (1989: 27-28) como descriptiva, debido a que el papel del evaluador se enfocaba en la descripción de fortalezas y debilidades en relación con el logro de ciertos objetivos. Desde ese momento, la medición ya no podía ser asimilada a la evaluación, sino que apareció como *una* de las herramientas disponibles. La publicación en 1949 de su informe hizo a Tyler (1949) famoso y fue reconocido como el padre de la evaluación.

Sin embargo, la evaluación descriptiva no fue una herramienta suficiente para resolver las deficiencias del sistema educativo de los Estados Unidos. En el marco de la guerra fría, el lanzamiento exitoso del Sputnik por la Unión Soviética fue considerado como un fuerte y evidente campanazo de alerta. No bastaba describir las aptitudes y los objetivos alcanzados. Era indispensable que la evaluación ofreciera elementos de juicio comparando los objetivos con estándares externos al programa. La introducción del juicio en la evaluación marca el nacimiento de la evaluación de tercera generación.

De este modo, el evaluador asume no sólo el papel de medidor y descriptor, sino de juez. Eso implica que es necesario considerar también los objetivos como problemáticos y, por lo tanto, susceptibles de ser juzgados en relación con unos estándares externos al mismo programa: una evaluación descriptiva de objetivos puede mostrar que se logran de manera perfecta los objetivos, pero es muda en cuanto a la pertinencia de los objetivos perseguidos. Introducir elementos valorativos en una empresa científica, por naturaleza considerada como libre de valores (*value free*), era inaceptable para evaluadores apegados a una perspectiva científica tradicional. A pesar de la reticencia de los evaluadores, muchos terminaron aceptando este rol, ya que, entre todos los otros jueces posibles, se consideraron como los que serían probablemente los más objetivos. A final de cuentas, la necesidad de juicio no pudo ser ignorada y aceptaron el reto.

En esta perspectiva, a partir de la segunda mitad de los años 60, se desarrollaron nuevos modelos de evaluación que compartían el hecho de que el juicio hacía parte integral del trabajo de evaluación, y que el evaluador era, de forma más o menos explícita, un juez. Saber hasta qué punto el evaluador debía o podía suplantar al decisor político a través de sus juicios, si podía aún considerarse como un analista "objetivo" o "independiente", fue un punto de intenso debate.

Cada una de estas generaciones representa una etapa suplementaria en la sofisticación de la evaluación. La primera generación se centró en recolectar datos relativos a los individuos; la segunda etapa amplió la evaluación a objetos como los programas, los insu- mos, las estrategias pedagógicas, el diseño institucional, etc. La tercera generación introdujo en la evaluación el juicio a los objetivos -a partir de criterios o estándares tanto intrínsecos como extrínsecos.

Por último, una evaluación de cuarta generación pretende suplir a unas debilidades que, siempre según Guba y Lincoln (1989: 31-32), permanecen en las prácticas evaluativas de las tres generaciones desarrolladas anteriormente. En particular, se trata de la fuerte tendencia al "managerialismo", la difícil o imposible tenida en cuenta del pluralismo de los valores y la tendencia a privilegiar el paradigma de investigación científica tradicional.

La primera debilidad hace referencia a la relación entre el contratante de una evaluación y el evaluador contratado, donde es el contratante quien dispone del poder último de orientar el contenido y

forma de evaluación y de establecer las condiciones de su conducción gracias a su poder de veto, expresado en su poder de escoger a quién va a contratar. Igualmente, esta perspectiva tiende a favorecer una lógica de ingeniería social que favorece las actitudes tecnocráticas. También, este desequilibrio de poder facilita las actitudes benévolas por parte del evaluador hacia el programa y su responsable, como estrategia para obtener otros contratos. Esta situación conlleva a la segunda debilidad señalada: la casi imposibilidad de tener en cuenta valores distintos a los de la dirección, y que pueden expresarse, por ejemplo, en otros grupos relacionados con el programa a evaluar (beneficiarios y funcionarios de base en particular). De modo que en cualquier evaluación, aunque pretende ser "objetiva", es implícitamente orientada por determinados valores. Y, finalmente, los evaluadores, con el deseo de quedar considerados como auténticos científicos, tienden a usar exclusivamente el paradigma positivista vigente en las ciencias naturales, lo que se refleja en particular en el uso de métodos cuantitativos sofisticados, con la pretensión de encontrar la "verdad" de forma inobjetable.

La evaluación de cuarta generación, propuesta por Guba y Lincoln (1989: 38), y llamada por ellos *responsive constructivist evaluation*, o evaluación constructivista no focalizada (ECNF), se considera una perspectiva alternativa a las tres anteriores, permitiendo superar las debilidades señaladas. En la literatura especializada se la denomina también como "evaluación pluralista" (Monnier, 1992)¹. Para Guba y Lincoln (1989: 43-45), el carácter no focalizado hace referencia a que los parámetros y límites de la evaluación, sus contornos, no son fruto de una simple imposición o, en el mejor de los casos, negociación entre el contratante y el contratado, sino que corresponden a un proceso interactivo y negociado que incluye a todas las partes involucradas. Monnier (1992), en su propuesta de modelo de aplicación, denomina este lugar de negociación como la *instancia de evaluación* que reúne a todas las partes involucradas (ver Roth, 2002: 164). De modo que los contornos de la evaluación no son fijados de manera autoritaria y *a priori*, sino que su definición es parte del proceso mismo de evaluación. El término constructivista se usa para señalar la metodología empleada para llevar a cabo la evaluación. Tiene sus fundamentos en una metodología interpretativa y hermenéutica, en ruptura con el paradigma científico positivista dominante. En esta perspectiva ya no se pretende descubrir la "verdad",

La evaluación no aporta la verdad universal o casi universal, sino una nueva interpretación que consiste en una nueva construcción (o un metadiscurso) a tomar en cuenta para generar cambios políticos o institucionales de manera (más) consensuada.

sino varias interpretaciones resultantes de las construcciones de la realidad que hacen las partes involucradas a partir de sus valores (pluralismo de valores). El trabajo del evaluador ya no corresponde al de un científico que va a descubrir (e imponer) la verdad, sino que se parece más a la labor de un "orquestador" (Guba y Lincoln, 1989: 45) de procesos de negociación entre diferentes visiones, quien intenta alcanzar un consenso con base en mejor información. La evaluación no aporta la verdad universal o casi universal, sino una nueva interpretación que consiste en una nueva construcción (o un metadiscurso) a tomar en cuenta para generar cambios políticos o institucionales de manera (más) consensuada. Eso implica que la evaluación sea una actividad práctica, deliberativa y situada.

Como se puede anotar, esta última perspectiva evaluativa es mucho más acorde con una visión de la acción pública y de la construcción de una experticia pública basada en la participación y la deliberación públicas. En vez de privilegiar una perspectiva científica tradicional, vertical, tendiente a favorecer los valores y los puntos de vista de la autoridad, de las instituciones, y de la tecnocracia, la evaluación pluralista pretende tomar en cuenta los diversos puntos de vista para la construcción de consensos más amplios, de forma más horizontal. En este sentido, tiene una relación estrecha con la idea misma de gobernanza.

2. Las prácticas evaluativas en Colombia

Progresivamente, desde los años 90, la evaluación se ha instalado como una práctica relativamente habitual en Colombia. Tanto las universidades, las ONG, las organizaciones internacionales como las instituciones públicas han llevado a cabo numerosas experiencias de evaluaciones en el país. Aquí nos centraremos en una mirada a las prácticas evaluativas realizadas, o contratadas, por las instituciones públicas oficiales.

Como ente de control, la Contraloría General de la República ha iniciado una serie de *Estudios Sectoriales* para servir de insumos para el ajuste de las políticas públicas. Estos estudios son principalmente de tipo descriptivo: por lo general se realiza sintética pero minuciosamente una descripción del problema, su evolución, y asimismo para lo referente al marco legal e institucional operante. Igualmente, la Defensoría del Pueblo, a través de sus *Informes Defensoriales*, ha producido documentos de tipo evaluativo basados generalmente

Se pretende someter el programa a un proceso experimental a partir del cual se podrá tomar la decisión de expandir, corregir o abandonar un programa. Esta experiencia tiene un carácter muy importante para el país y para el rumbo que puede tomar el desarrollo de la evaluación en Colombia.

en visitas de campo, entrevistas y aplicación de encuestas que corresponden por lo general a descripciones de situaciones problemáticas y de su marco legal e institucional. Prácticamente, en estas evaluaciones adelantadas por los entes de control², no hay una reflexión de tipo metodológico importante, sino en el caso de aplicación de encuestas. Estas evaluaciones corresponden claramente a evaluaciones de segunda generación. Es de resaltar la interesante evolución actual de las contralorías, que tradicionalmente centraban sus actividades en el control de legalidad y presupuestal y que, hoy en día, se dirigen hacia la realización de evaluaciones de programas y de políticas públicas.

A nivel central, la creación de un Sistema Nacional de Evaluación de Resultados de la Gestión Pública -SINERGIA- y de una Dirección de Evaluación de Políticas Públicas en el seno del Departamento Nacional de Planeación es una muestra de la importante institucionalización de este tema (ver Cunill y Ospina, 2003). Mediante el desarrollo de estas iniciativas institucionales, uno de sus componentes pretende realizar evaluaciones estratégicas externas que deben "producir la información necesaria para mejorar la efectividad, el diseño y la ejecución de políticas y programas considerados centrales para los objetivos de desarrollo del país"³. A tal fin se definió una agenda de evaluaciones "estratégicas" para varias políticas gubernamentales, que culminó en agosto de 2009.

El DNP decidió llevar a cabo esta primera experiencia evaluativa mediante la evaluación de los programas diseñados y ejecutados en el marco de la Red de Apoyo Social (RAS)⁴. Estos programas, de carácter temporal, son: Empleo en Acción, Familias en Acción y, por último, Jóvenes en Acción. Según se señala en el Informe N° 3 de la serie Evaluación de Políticas Públicas (DNP, 2004a: 7), "el diseño y puesta en marcha de una Red de Apoyo Social (RAS), como mecanismo de protección de la población más vulnerable frente a los efectos del desempleo y la reducción en el ingreso de los hogares, fue una de las iniciativas de política más importantes puestas en marcha en el país en muchos años", razón por la cual se consideró necesario realizar una evaluación. Adicionalmente, se arguyó que la reducción de los recursos disponibles para inversión social al final de los años 90, la necesidad de contar con información que diera cuenta de la eficiencia y el impacto de las nuevas políticas, y la necesidad de disponer de mejores esquemas de evaluación de la

política social justificaba plenamente la puesta en marcha de una política de evaluación. Estas evaluaciones se consideran como "la primera iniciativa de envergadura en evaluación de impacto de programas sociales y ha contribuido de manera importante a la institucionalización de la función de evaluación en el Estado colombiano" (DNP, 2004a: 8). Según el mismo documento, además, "se trató de la primera experiencia de diseño de un programa paralelamente con el diseño de su evaluación de impacto". De acuerdo con los autores, esta experiencia de evaluación de los programas de la "RAS constituye un punto de referencia en el análisis de políticas públicas y para futuras evaluaciones de programas sociales" (DNP, 2004a: 9). Con su diseño se pretende someter el programa a un proceso experimental a partir del cual se podrá tomar la decisión de expandir, corregir o abandonar un programa. Por lo señalado anteriormente, es claro que esta experiencia tiene un carácter muy importante para el país y para el rumbo que puede tomar el desarrollo de la evaluación en Colombia.

Hasta marzo de 2009, el DNP ha publicado los resultados de este proceso a través de ocho números de su serie *Evaluación de Políticas Públicas*. Estas evaluaciones se relacionan en el Cuadro 1, señalando la metodología utilizada y los autores.

Cuadro 1
Evaluaciones DNP

Evaluación	Evaluadores	Metodología
1. Programa Familias en Acción: Condiciones iniciales de los beneficiarios e impactos preliminares (marzo 2004).	Unión temporal Institute of Fiscal Studies - Econometría - Sistemas Especializados de Información S.A.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).
2. Programa Empleo en Acción. Condiciones iniciales de los beneficiarios e impactos a corto plazo (octubre 2004).	Unión temporal Institute of Fiscal Studies - Econometría - Sistemas Especializados de Información S.A.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).
3. Red de Apoyo Social: Conceptualización y evaluación de impacto (noviembre 2004).	Grupo de evaluaciones focalizadas de la Dirección de Evaluación de Políticas Públicas del DNP.	Cuasi-experimental.
4. Programa Familias en Acción: Impacto del Programa a un año y medio de su ejecución (mayo 2006).	Unión Temporal Institute of Fiscal Studies - Econometría - Sistemas Especializados de Información S.A.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).

5. Programa Empleo en Acción: Impacto del Programa sobre los beneficiarios y sus familias (enero 2008).	Unión temporal Institute of Fiscal Studies - Econometría - Sistemas Especializados de Información S.A.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).
6. Programas Familias en Acción: Impactos en capital humano y evaluación beneficio-costos del Programa (febrero 2008).	Unión temporal Institute of Fiscal Studies - Econometría - Sistemas Especializados de Información S.A.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).
7. Programa de Vivienda de Interés Social Urbano: Impactos en la calidad de vida y evaluación del proceso de focalización (2007).	Centro de Estudios Regionales Cafeteros y Empresariales.	Cuasi-experimental (comparación entre grupo beneficiario y grupo de control).
8. Evaluación de impacto del Fondo Colombiano de Modernización y Desarrollo Tecnológico de las Micro, Pequeñas y Medianas Empresas - FOMIPYME (octubre 2008).	Econometría S.A.	Aplicación de encuestas (cuantitativas) a una muestra de operadores del programa entre 2001 y 2004.

Fuente: elaboración propia.

Como consta en el Cuadro 1, la mayoría de las evaluaciones contratadas por el DNP fueron con la misma empresa, conformada por el Institute of Fiscal Studies, IFS (un instituto de investigación inglés especializado en el análisis de las políticas sociales y económicas a partir de técnicas microeconómicas), Econometría (una empresa de consultoría privada colombiana especializada en estudios económicos) y otra empresa de consultoría nacional especializada en la recolección y tratamientos de datos. A nivel metodológico, los evaluadores han usado siempre la misma metodología cuasi experimental longitudinal (antes-después), que compara la situación de un grupo de beneficiarios con un grupo de control en distintos momentos. Las dos evaluaciones que fueron realizadas por otros evaluadores usaron una metodología similar. Como se nota, esta concepción de la evaluación se sitúa plenamente en la perspectiva científica tradicional (tercera generación).

Los programas evaluados son principalmente dos: Familias en Acción y Empleo en Acción. Para el programa Jóvenes en Acción, aunque no haya evaluaciones disponibles a la fecha (marzo de 2009), el DNP informa que sus resultados son positivos. Un tercer programa, que no hace parte de la RAS, concierne a la Vivienda de Interés Social Urbana. A continuación se realiza un análisis crítico de las evaluaciones de Familias en Acción y Empleo en Acción, mostrando

Esta concepción de la evaluación se sitúa plenamente en la perspectiva científica tradicional (tercera generación).

sus debilidades o problemas tanto en el plano de la metodología usada como en los resultados presentados y la utilidad de éstos.

Los programas de la Red de Acción Social

La estrategia de la Red de Apoyo Social (RAS), como ya se señaló, incluyó a los programas Familias en Acción, Empleo en Acción y Jóvenes en Acción. Con éstos, el gobierno pretendía entregar subsidios para nutrición y educación a 300.000 familias, generar 300.000 empleos transitorios y capacitar para el empleo a 100.000 jóvenes. La duración de la estrategia se estimó en tres años con un costo total de 520 millones de dólares (DNP, 2004a: 9).

A continuación se reproduce una descripción de los programas de la RAS (DNP, 2004a: 12-13):

"Con la RAS se buscaba, en principio, mitigar los efectos de la crisis económica sobre los grupos poblacionales más vulnerables, de manera rápida, eficaz y con beneficios palpables de forma inmediata para ellos. Sus programas fueron diseñados con el propósito de que su flexibilidad permitiera contraerlos una vez superados los efectos de la crisis económica y mejoradas las condiciones de vida de la población objetivo. Se trataba de proteger el capital humano, entendido como la capacidad futura de generación de ingresos de sus beneficiarios.

Con el programa Familias en Acción se buscaba remover los obstáculos que impedían a los hijos de las familias en estado de pobreza acceder a niveles suficientes de nutrición y cuidado de la salud, así como beneficiarse de la formación que se adquiere a través de la educación primaria y secundaria. El programa otorgaba transferencias monetarias a familias del primer nivel de Sisbén de las zonas urbanas y rurales de municipios de menos de 100.000 habitantes. La transferencia monetaria estaba compuesta de un subsidio nutricional para familias con al menos un hijo entre 0 y 7 años, y de un subsidio escolar por cada hijo de la familia entre 7 y 17 años que asistiera a la escuela. Estos subsidios, como incentivos directos a la demanda de los servicios de educación y salud, condicionaban su pago al cumplimiento del compromiso de asistir de forma sistemática a los centros que los ofrecían. Los municipios en los que se desarrolló el programa debían garantizar, mediante la firma de un convenio, que su capacidad de oferta en los servicios de educación y salud respondería al incremento potencial en la demanda que se generara como efecto del programa.

El programa Empleo en Acción - Proyectos Comunitarios buscaba atenuar el impacto provocado por la crisis económica y que se evidenció en el desempleo y la posterior caída de los ingresos de la población de los niveles 1 y 2 del Sisbén en zonas urbanas del país. Este programa generaba empleos transitorios, de una duración máxima de cinco meses, en pequeños proyectos de construcción de infraestructura urbana y social intensivos en mano de obra, los cuales se ejecutaban en las localidades y barrios más deprimidos de los municipios. Se trataba así de reducir temporalmente el desempleo y de generar ingresos a los beneficiarios y sus familias. También se buscaba contribuir a solucionar las necesidades de infraestructura de las comunidades más pobres y de esta forma incidir en sus condiciones de vida. Los proyectos se financiaban de forma compartida entre la Nación y los municipios. Además de los empleos, el programa financiaba un porcentaje de los materiales de los proyectos.

Por último, el programa Jóvenes en Acción tenía como propósito mejorar las oportunidades de inserción laboral de los jóvenes (entre 18 y 25 años) pertenecientes a los niveles 1 y 2 de Sisbén, a través de acciones de capacitación en un oficio y con una práctica laboral efectiva en una empresa del sector formal de la economía. El programa estaba dirigido a los jóvenes de las ciudades de Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales y Cartagena, donde se presentaron las tasas de desempleo más altas del país. En el corto plazo, se esperaba mejorar las condiciones de empleabilidad de los jóvenes beneficiarios. En el largo plazo se buscaba mejorar los niveles de calificación del grupo objetivo de manera que alcanzara un mayor grado de competitividad".

La evaluación del programa Empleo en Acción

La evaluación del programa *Empleo en Acción* se presentó en los documentos del DNP de la serie "Evaluación de Políticas Públicas" 2 y 5, publicados respectivamente en octubre de 2004 (DNP, 2004a) y enero de 2008 (DNP, 2008a).

El objetivo del Programa era la de atenuar los impactos de la crisis económica, del ajuste fiscal y de la recesión en el sector de la construcción; es decir, el Programa tenía la intención de atenuar los impactos del ciclo económico sobre los pobres y no la de ser un programa contra la pobreza (DNP, 2008a: 55). Para cumplir estos objetivos, el Programa aportó recursos para el pago de la mano de

obra no calificada y para cubrir parcialmente la compra de materiales. Los proponentes podían ser los departamentos, distritos, municipios, empresas de servicios públicos del Estado, entidades sin ánimo de lucro, empresas asociativas de trabajo, organizaciones no gubernamentales o instituciones de base comunitaria. El proponente debía garantizar los recursos de contrapartida necesarios para complementar los aportes del Programa y cubrir, así, los costos totales del proyecto.

El Programa financió desde fines del año 2000 un total de 3.724 proyectos u obras, de los cuales el 37% (1.384 obras) estaban localizados en un conjunto de 71 de los 78 municipios considerados como prioritarios por el Programa, que incluían las 24 principales áreas metropolitanas o conurbadas del país y otras cabeceras municipales de más de 100.000 habitantes, y el 63% restante (2.340 obras) en un conjunto de 581 municipios no prioritarios de menos de 100.000 habitantes o que no formaban parte de áreas metropolitanas o conurbadas. Hasta fines de 2002 se habían aprobado 3.487 proyectos, cuya ejecución había presentado importantes demoras por problemas de disponibilidad de fondos y por la complejidad de los procedimientos de aprobación, legalización y desembolso. A comienzos de 2003 se simplificaron estos procedimientos y en marzo se aprobaron 358 proyectos adicionales sólo en municipios prioritarios. A partir de esa fecha se canceló la aprobación de nuevos proyectos.

En total, con corte a septiembre de 2004, el Fondo de Inversiones para la Paz (FIP) había ejecutado recursos por valor aproximado de \$250.000 millones, de los cuales \$136.000 millones (54%) fueron asignados a financiación de mano de obra; \$105.000 millones (42%) a financiación de materiales; y \$14.000 millones a pago de Organismos de Gestión. En marzo de 2003, el gobierno decidió suspender la aprobación de nuevas obras, y a fines de 2004 se terminaron las últimas obras del Programa.

Según lo señala el documento de evaluación (DNP, 2004b: 12), la decisión de evaluar este programa se debió a cinco razones. La primera, a lo novedoso del programa en cuanto a la manera de asignar recursos directamente a los beneficiarios del empleo; la segunda, a su operación por fuera de las instituciones que tradicionalmente ejecutan programas de infraestructura social; la tercera, a la participación en la ejecución del programa de actores privados; la cuarta, a la necesidad de proteger en forma rápida y oportuna a las poblaciones vulnerables; y, finalmente, para contar con elementos

Diversas dificultades surgieron para la aplicación de dicha metodología en razón a variaciones en la implementación del programa que obligaron a usar una metodología cuasi experimental en vez de experimental.

para establecer de forma clara las fortalezas y debilidades de este tipo de intervenciones.

Aspectos metodológicos. La metodología experimental inicialmente propuesta para la evaluación de impacto se basó en una comparación entre un grupo de beneficiarios y un grupo de control. Se realizaron encuestas para construir una línea de base que estableciera las condiciones de las poblaciones de estos dos grupos antes de la implementación del programa. Se realizó un primer seguimiento durante el programa (efectos de corto plazo) y otro después de finalizado el programa (efectos de mediano plazo).

Sin embargo, diversas dificultades surgieron para la aplicación de dicha metodología en razón a variaciones en la implementación del programa que obligaron a usar una metodología cuasi experimental en vez de experimental. En primer lugar, no se pudo garantizar que la selección de los beneficiarios del programa fuera totalmente aleatoria (DNP, 2004b: 28). Algunos beneficiarios ya no estaban desempleados en el momento del levantamiento de la línea de base, o incluso ya estaban trabajando en la obra (pp. 29 y 31). Además, el número de personas inscritas en el grupo de control era la mitad en relación con el número de personas beneficiadas (p. 29). Luego, fue necesario excluir casi la tercera parte de los proyectos seleccionados inicialmente (47 de 155) para la evaluación por diversas razones. Al final, la evaluación de impacto se realizó en 108 proyectos de los 3.845 proyectos aprobados, es decir el 2,81%⁵. Por estos motivos, los autores fueron obligados a recurrir a una técnica econométrica conocida como de variables instrumentales, con la cual los errores de estimación son mayores y en donde la corrección del sesgo de selección es parcial (p. 35). Los propios autores de la evaluación señalan que: "se debe ser muy cuidadoso en la interpretación de los resultados en municipios prioritarios y no prioritarios" (p. 31). Es decir, los únicos resultados válidos son los que corresponden a la totalidad del universo evaluado.

Resultados. La evaluación de corto plazo (DNP, 2004b) permitió mostrar los siguientes principales resultados positivos observados en el grupo de beneficiarios en comparación con el grupo de control:

- Un incremento de 8,6 horas de trabajo.
- Un incremento de 39% del ingreso mensual.
- No disminuyó la oferta laboral de otros miembros del hogar y no reemplazó otras actividades productivas.

Las dos evaluaciones se publicaron cuando ya se había cancelado el programa. Sin embargo, tanto la primera evaluación (a corto plazo) como la final muestran que el programa tuvo algunos efectos positivos.

- Se observó un incremento de 9% del consumo total del hogar (frente al 40% esperado).

La evaluación final, ex post, del programa (DNP, 2008a) atribuye a éste los siguientes principales resultados:

- El programa logró una buena focalización, es decir no desplazó empleos de tiempo completo en el sector formal.

- Los beneficiarios recibieron importantes beneficios en el corto plazo, durante la ejecución de las obras y, adicionalmente, mejoraron su empleabilidad nueve meses después de terminadas, aunque estos beneficios no son sostenibles a mediano o largo plazo (p. 62).

- El programa funcionó efectivamente como un programa de igualdad de oportunidades (*workfare program*) para personas con menores oportunidades de trabajo en otras actividades (p. 65).

- El programa, nueve meses después de terminadas las obras, no había tenido todavía efectos muy perceptibles sobre las comunidades en áreas de tratamiento. Las comunidades beneficiarias percibieron sólo dos impactos positivos menores: i) un aumento en la recreación y en las prácticas deportivas de los niños y jóvenes en los barrios intervenidos; y ii) un aumento de la seguridad en la zona en vías y espacios públicos (p. 70).

Las dos evaluaciones se publicaron cuando ya se había cancelado el programa: se suspendió en marzo de 2003 y todos los proyectos se terminaron durante el 2004. Sin embargo, tanto la primera evaluación (a corto plazo) como la final, publicadas respectivamente en el 2004 (DNP, 2004b) y en el 2008 (DNP, 2008a), muestran que el programa tuvo algunos efectos positivos. A pesar de esto, como se indica en el documento N° 8, los evaluadores consideraron, *a posteriori*, que la decisión del gobierno colombiano de suspender el programa fue correcta. Pero señalan que el gobierno tomó esta decisión "*por diferentes razones*" (DNP, 2008a: 81). En particular, se alude allí, de un lado, *problemas operativos* en la ejecución del programa y, del otro lado, la visión de que *existían otras alternativas*, en particular Familias en Acción y un programa nuevo de subsidios salariales (PADE) para promover mayor empleo en el sector privado (DNP, 2008a: 81) del cual -como señalan los evaluadores- "no se tenía ninguna evaluación piloto de sus posibles impactos, ni se tuvo en cuenta la visión negativa que se tiene de este tipo de programas en la literatura internacional" (DNP, 2008a: 84). Como se deduce de lo anterior, en este caso la evaluación no tuvo impacto en la decisión de dar por terminado el programa.

Se percibe a través del texto de la evaluación, de un lado, un cierto sentimiento de frustración de los evaluadores debido a que el trabajo que adelantaban no fue tenido en cuenta para la toma de la decisión, y del otro, la búsqueda de una justificación *a posteriori* para avalar o legitimar esta decisión, a pesar de que muestran ciertos resultados positivos, y mantienen la esperanza de que su evaluación sea útil para decisiones futuras.

Para concluir, cabe considerar que, a nuestro juicio, no es muy extraño que se obtuvieran algunos resultados positivos. Parece, en efecto, obvio que entregando un sueldo a personas en situación de pobreza se obtenga un mejoramiento en algunos aspectos de su vida. Por lo tanto, no es sorprendente que el Programa Empleo en Acción tenga efectos positivos sobre ingreso, horas de trabajo y consumo de bienes de primera necesidad. En este sentido, consideramos que la comparación hubiera debido hacerse más entre *programas diferentes* para conocer su eficacia respectiva, que entre beneficiarios y no beneficiarios del programa. Adicionalmente, puede preguntarse si la focalización de estos programas no constituye una forma sofisticada de generar un ambiente clientelar, haciendo beneficiar a cuenta gotas algunos beneficios pasajeros a relativamente pocas personas, pero diseminadas en muchas localidades.

La evaluación del programa Familias en Acción

La evaluación del programa *Familias en Acción* se presentó en los documentos del DNP de la serie "Evaluación de Políticas Públicas" Nos. 1, 4 y 6, publicados respectivamente en marzo de 2004 (DNP, 2004b), mayo de 2006 (DNP, 2006) y febrero de 2008 (DNP, 2008b).

El programa. El Programa Familias en Acción consiste en la entrega de "subsidios de nutrición y de educación a familias del nivel 1 del SISBEN, localizados en municipios con menos de 100 mil habitantes. Se considera que la población de estos municipios es netamente rural aunque cada uno tenga su cabecera municipal. El subsidio consiste en otorgar a las familias un apoyo monetario directo de \$46.500 por familia con 1 o más niños de 0 a 6 años, que asistan regularmente a controles de crecimiento y desarrollo en las instituciones de salud, y de \$14.000 y \$28.000 por niño de 7 a 11 y de 12 a 17 años, respectivamente, asistiendo regularmente a una institución educativa de primaria o secundaria. El programa asegura el subsidio nutricional por 12 meses y el subsidio escolar por 10 meses, los cuales se entregan cada dos meses" (DNP, 2006: 11).

En contraposición al caso del Programa Empleo en Acción, el programa Familias en Acción opera para una población para la cual ya existe desde 1986 otro programa: los Hogares Comunitarios de Bienestar. Sin embargo, existe una diferencia fundamental entre los dos programas: Familias en Acción (FA) es un subsidio en dinero condicionado, mientras Hogares Comunitarios de Bienestar (HCB) es un programa de subsidio en especie (alimentación y atención infantil). FA es un programa para el hogar con un subsidio nutricional independiente del número de niños, mientras que HCB provee alimentación y atención infantil por cada niño participante. La participación de una familia en uno de estos programas es excluyente del otro. En esta ocasión se realizó también una evaluación comparativa de los dos programas.

Aspectos metodológicos. La metodología general usada para esta evaluación es la misma que para Empleo en Acción y, por lo tanto, con los mismos problemas ya señalados. Como elemento adicional, se realizó una comparación del programa FA con el programa preexistente HCB.

Resultados. La evaluación de corto plazo considera que el programa dio resultados positivos. En sus conclusiones, los autores señalan que el programa tuvo un "impacto importante en la ingesta de algunos alimentos" y que "el estado nutricional de niños de 0 a 6 años ha mejorado" (DNP, 2004b: 162). En la segunda evaluación (DNP, 2006) se concluye que:

- "Familias en Acción está alcanzando impactos importantes en su corto período de ejecución en la población objetivo del Programa. En particular, la participación en el Programa ha incrementado el consumo de los hogares en un 15% con respecto a lo informado en el 2002. Cabe resaltar que este incremento en los gastos se ha concentrado en el consumo extra de proteínas (carne, huevos y leche), compra de ropa y calzado para niños y en cierto aumento de gasto en educación. No se observó incremento de gasto en artículos para adultos tales como alcohol, cigarrillos o ropa.

- El incremento del consumo de los alimentos mencionados se vio reflejado en una mejoría del estado nutricional de los niños. Se observaron efectos positivos en la talla de los niños rurales de 0 a 2 años y en el peso de los niños urbanos. Sin embargo, no se ha identificado aún un impacto del programa en niños de 2 a 4 años viviendo en la zona urbana.

- Adicional al impacto nutricional, se comprobó un incremento en la asistencia a controles de crecimiento y desarrollo de niños

menores de 6 años inscritos en FA, así como una disminución de la proporción de niños rurales menores de 4 años afectados por enfermedades diarreicas en los 15 días previos a la encuesta.

- También se encontró un impacto positivo en la asistencia escolar de niños entre 12 y 17 años (y más específicamente entre 14 y 17). Este incremento fue mayor en la zona rural, donde la tasa de asistencia inicial era menor a la urbana. Por otro lado, se observó un impacto limitado del Programa en la asistencia de niños menores de 12 años; cabe recordar que la tasa de asistencia de este grupo de población es superior al 90%.

- A pesar de identificarse un saldo migratorio negativo de personas al interior de los hogares, éstos tienden a movilizarse en menor proporción cuando son beneficiarios del programa, cuando poseen vivienda propia, cuando viven en zonas en el que el conflicto armado es de baja intensidad.

- En cuanto al programa de HCB, se destaca que tiene una adecuada focalización en estos municipios, en la población realmente más pobre entre los elegibles, y en el grupo de 2 a 4 años de edad.

- Los niños que asisten a los HCB tienen en promedio 2 cms más de estatura que los que no asisten, y después de controlar por un conjunto de variables independientes de orden individual y municipal.

- Se evidencia un efecto positivo en la vinculación al mercado laboral de las madres que envían sus hijos a los HCB.

- Resulta destacable el mejor desempeño escolar de jóvenes de 12 a 17 años, que tuvieron la oportunidad de asistir a algún HCB antes de cumplir 7 años.

- El hecho especial de que los subgrupos de niños más impactados por cada uno de los Programas fueran diferentes, sugiere la conveniencia de explorar algún tipo de integración o combinación" (DNP, 2006: 61-62).

A partir de estas conclusiones, los autores someten a consideración unas recomendaciones. La primera consiste en su continuación debido a sus efectos positivos en el mediano plazo, aunque recomiendan realizar otra evaluación para conocer su efecto y sostenibilidad a largo plazo, y también efectuar evaluaciones relacionadas con los mecanismos de operación de los programas. Luego, consideran que FA y HCB pueden ser programas complementarios, y dejan entender que existe el proyecto de integrarlos. Sugieren también que se podría

La función central de las evaluaciones, y en particular las evaluaciones de tipo experimental, es aportar insumos para la toma de decisión. Sin embargo, en los dos casos estudiados, se nota que la toma de decisión no se realizó con base en las evaluaciones presentadas.

reducir los subsidios para los niños que asisten a primaria para incrementar el de bachillerato. Terminan diciendo que estos resultados son válidos sólo para las poblaciones rurales (poblaciones de menos de 100.000 habitantes), y que la ampliación del programa en las ciudades debe ser objeto de otra evaluación (DNP, 2006: 62).

Conclusión

La función central de las evaluaciones, y en particular las evaluaciones de tipo experimental, es aportar insumos para la toma de decisión. Sin embargo, en los dos casos estudiados se nota que la toma de decisión no se realizó con base en las evaluaciones presentadas. Al contrario, ¡parece ser que son los evaluadores quienes integran y legitiman en sus comentarios y recomendaciones las decisiones ya tomadas!

En el caso de EA, para tomar la decisión de suspender el programa se aludió, de un lado, a problemas operativos en la ejecución del programa y, del otro lado, a la visión de que existían otras alternativas, en particular Familias en Acción y un programa nuevo de subsidios salariales (PADE) para promover mayor empleo en el sector privado (DNP, 2008a: 81) del cual -como señalan los evaluadores- "no se tenía ninguna evaluación piloto de sus posibles impactos, ni se tuvo en cuenta la visión negativa que se tiene de este tipo de programas en la literatura internacional" (DNP, 2008a: 84). Como se deduce de lo anterior, en este caso la evaluación no tuvo impacto en la decisión de dar por terminado el programa.

Sin embargo, los autores consideran, casi diez años después de su lanzamiento y cinco años después de su suspensión, que este tipo de programas puede ser positivo, mediante algunos ajustes, como "programa de apoyo a la construcción y mantenimiento de obras públicas municipales en barrios de estratos socioeconómicos bajos, para garantizar igualdad de oportunidades de empleo a la población más pobre y vulnerable" (DNP, 2008a: 82).

En el caso de FA, los autores señalan que "la visión estratégica se plantea entonces sobre un escenario de *decisiones políticas ya tomadas* y sobre un operativo de expansión del Programa *ya iniciado*" (DNP, 2008b: 117). Además, se señala luego que "es de recordar que los impactos en las cabeceras municipales identificados en la evaluación fueron pequeños" (DNP, 2008b: 118).

Estos dos ejemplos parecen ilustrar a la perfección lo que se conocía desde ya un tiempo: el problema mayor de este tipo de evaluación

es que los resultados se dan a conocer cuando ya no tienen mayor pertinencia política, y que, para tomar decisiones acerca de un programa, la importancia de los impactos observados no son necesariamente los elementos más centrales para fundamentar una decisión.

Además, parece pertinente cuestionar la metodología utilizada -de comparación con un grupo de control-, en la medida que, como ya se señaló, es bastante obvio, o por lo menos no sorprendente, que la entrega de subsidios a población pobre con necesidades básicas insatisfechas (en particular, alimentación, salud y educación) va a generar un comportamiento positivo en relación con estas variables. Aun con mayor razón, cuando los subsidios se entregan de forma condicionada, como en el caso de Familias en Acción, al cumplimiento de ciertas obligaciones como la participación en algunas reuniones de información en salud, por ejemplo, o la asistencia escolar, so pena de perder el subsidio.

Finalmente, estas observaciones confirman la sensatez de las críticas a las evaluaciones que se basan en un intento, siempre difícil y limitado en su validez, de evaluar las políticas y programas mediante el paradigma científico de la medición positivista, y en este caso con ninguna reflexión acerca de los valores, tal como lo plantea la evaluación de la tercera generación. Se construye una imagen positiva de los programas únicamente con base en la *ausencia de política alternativa*, en un contexto de extrema pobreza en donde cualquier esfuerzo generará probablemente algún efecto positivo. Por lo tanto, en este marco, las políticas y programas siempre resultan legitimados por la "ciencia" evaluativa.

La visión que se desprende de esta serie de evaluaciones es la de una estrategia evaluativa sin participación de los beneficiarios ni de los ejecutores. Corresponde a una visión tecnocrática, de arriba hacia abajo. Y, como se observa, ni siquiera los actores políticos toman decisiones apoyándose en las apreciaciones derivadas de las evaluaciones.

Claramente, las evaluaciones realizadas por el DNP, basadas en una ilusoria capacidad de medición objetiva con fines de alcanzar una óptima eficiencia técnico administrativa, no constituyen herramientas e instrumentos para profundizar la democracia y favorecer una revisión de las políticas y programas desarrollados por el gobierno y las instituciones públicas. Tampoco, aparecen capaces de articular una diversidad de fuentes y voces presentes en la sociedad.

En este sentido, es sólo con el desarrollo de evaluaciones de "cuarta generación" que se podría introducir elementos de deliberación, o sea de gobernanza. La práctica evaluativa colombiana "oficial" presente sigue centrándose, a la vez que contribuye a su reforzamiento, en una alianza del saber científico exclusivo en manos de expertos con un poder político fundamentalmente concebido desde una perspectiva autoritaria y vertical.

La evaluación de cuarta generación, o pluralista, pretende remediar los defectos claramente evidenciados en el análisis realizado: actitud benévola por parte del evaluador hacia el programa y su responsable, casi imposibilidad de tener en cuenta valores distintos a los de la dirección, y "positivismo" excesivo. ¿Cómo podría desarrollarse una evaluación pluralista en el tipo de programas aquí examinados?

La actitud benévola o legitimadora desarrollada por los evaluadores puede ser limitada mediante la constitución de una instancia de evaluación que integra expertos, beneficiarios y ejecutores de proyectos desarrollados a nivel local en el marco del programa. Esta instancia de evaluación -en donde se negocia y acuerda entre expertos (los evaluadores contratados), un grupo de beneficiarios (y, si es del caso, también de afectados negativamente por el proyecto), y otro de ejecutores, los asuntos relacionados con los tópicos a evaluar, la metodología y el desarrollo de la evaluación- es la responsable del diseño, de la conducción y del seguimiento de la evaluación. Para el diseño, Guba y Lincoln (1989: 142 y ss.) sugieren el uso de un proceso dialéctico hermenéutico. De este modo, el diseño de la evaluación será, por un lado, menos sometida a los deseos expresados explícita o implícitamente por el contratante -ya que el diseño es parte del proceso de evaluación- y el proceso evaluativo encontrará, por otro lado, una mayor legitimidad en el grupo de beneficiarios y ejecutores locales. El carácter participativo, deliberativo y plural del proceso de diseño, conducción y seguimiento de la evaluación debería permitir, por lo tanto, abrir espacio a la diversidad de valores portados por los diferentes actores del programa y no sólo a los valores de los expertos y de la dirección. Legitimidad y participación generando una mayor probabilidad de aceptación de las conclusiones de un proceso de evaluación.

En el caso concreto de un programa del tipo de Familias en Acción, se podría por lo tanto imaginar la constitución de una instancia de

En términos metodológicos, más que comparar entre un grupo de beneficiarios y un grupo de control, se recomendaría establecer una comparación con proyectos de otros programas existentes, incluyendo estos proyectos en la selección a evaluar.

evaluación conformada por evaluadores contratados, quienes en un primer momento seleccionan con base en un proceso aleatorio un cierto número de beneficiarios y de ejecutores locales de proyectos para conformar dicha instancia. Luego, como etapa de validación, se hace partícipe del resultado -el diseño inicial- a los beneficiarios y ejecutores de los proyectos seleccionados para que éstos hagan sus aportes y que, luego, la instancia de evaluación integre o revise su diseño inicial. A partir de allí, se conforma el equipo de evaluación que realizará el trabajo de evaluación en los proyectos seleccionados. En la medida de las posibilidades, sería también pertinente integrar ejecutores locales al equipo de evaluación.

En términos metodológicos, más que comparar entre un grupo de beneficiarios y un grupo de control, se recomendaría establecer una comparación con proyectos de otros programas existentes, incluyendo estos proyectos en la selección a evaluar. De esta manera, en la fase de diseño de la evaluación, la instancia de evaluación se vería obligada a ampliar sus preguntas evaluativas a más elementos, evitando el riesgo del "efecto túnel" (centrarse únicamente en observar los objetivos explícitos de un programa). Asimismo, sería pertinente no limitarse al uso de indicadores cuantitativos, realizando una serie de entrevistas a beneficiarios y ejecutores que permita disponer de elementos de juicio cualitativos que posibiliten indagar sobre la pertinencia o adecuación de los valores incorporados a los programas. Finalmente, se considera que una evaluación no puede tener un tiempo de ejecución superior a dos años si se pretende producir información útil para orientar la toma de decisión.

Notas

(1) Un resumen de la estrategia propuesta por Monnier (1992) para desarrollar una evaluación pluralista se encuentra en Roth (2002).

(2) Se pueden consultar las evaluaciones realizadas por estos entes de control en sus respectivas páginas web: <http://www.defensoria.org.co> y <http://www.contraloriagen.gov.co>.

(3) Ver el <http://www.dnp.gov.co/PortalWeb/Programas/Sinergia/PublicacionesSinergia>.

(4) Es de precisar que el DNP es de cierta forma juez y parte, ya que la misma institución está encargada de recomendar las principales políticas públicas del país mediante documentos de orientación (documentos CONPES),

de establecer el Plan Nacional de Desarrollo y, a su vez, de realizar la evaluación de éste.

(5) Cabe observar que el documento no señala de qué manera fueron seleccionados estos proyectos.

Bibliografía

- Cunill Grau, N. y Ospina Bozzi, S. (eds.) (2003), *Evaluación de resultados para una gestión pública moderna y democrática: experiencias latinoamericanas*, Caracas, CLAD; AECI/MAP/ FIIAPP.
- DNP (2004a), *Red de Apoyo Social: conceptualización y evaluación de impacto*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 3).
- _____ (2004b), *Programa Empleo en Acción: condiciones iniciales de los beneficiarios e impactos de corto plazo*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 2).
- _____ (2004c), *Programa Familias en Acción: condiciones iniciales de los beneficiarios e impactos preliminares*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 1).
- _____ (2006), *Programa Familias en Acción: impacto del programa a un año y medio de su ejecución*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 4).
- _____ (2008a), *Programa Empleo en Acción: impacto del programa sobre los beneficiarios y sus familias*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 5).
- _____ (2008b), *Programa Familias en Acción: impactos en capital humano y evaluación beneficio-coste del programa*, Bogotá, Departamento Nacional de Planeación (Serie Evaluación de Políticas Públicas; N° 6).
- Guba, E. y Lincoln, Y. (1989), *Fourth Generation Evaluation*, Newbury Park, Sage Publications.
- Monnier, Eric (1992), *Evaluations de l'action des pouvoirs publics*, Paris, Economica.
- Roth Deubel, A.-N. (2002), *Políticas públicas: formulación*,

implementación y evaluación, *of Curriculum and Instruction,*
Bogotá, Ediciones Aurora. Chicago, The University of
Tyler, R. W. (1949), *Basic Principles* Chicago Press.