

SE-valoración

Revista de la Sociedad Española de Evaluación

Año 3 | nº 11 | Junio 2010


Andalucía
se mueve con Europa

Evaluación de las Políticas de Empleo



SOCIEDAD
ESPAÑOLA
DE EVALUACIÓN

Exitosa respuesta a la convocatoria del I Premio de Investigación "Carlos Román"



PATROCINA: Instituto de Estudios de Economía, Evaluación y Empleo (I3E)

BASES: www.sociedadevaluacion.org

Ocho trabajos de investigación de distintas Universidades y Entidades, públicas y privadas, de Almería, Sevilla, Granada, Murcia, Pamplona, Madrid, Cataluña. Se han presentado a la 1ª Convocatoria del Premio de Investigación que lleva el nombre del primer Presidente y Fundador de la Sociedad Española de Evaluación de Políticas Públicas, Carlos Román del Rfo.

La exitosa acogida es un motivo de doble satisfacción; en primer lugar, porque pone de manifiesto el interés por la "herramienta" y, en segundo, porque refleja la capacidad de convocatoria que tienen tanto la revista como éste portal de la red; máxime cuando dificultades presupuestarias nos impidieron hacer uso de otros soportes publicitarios.

La Junta Directiva de la SEE quiere hacer público su agradecimiento a los/as investigadores/as que, aceptando las Bases del Concurso, han puesto sus trabajos a disposición del Jurado; éste se reunirá, en primera instancia, durante la primera semana de Julio en Sevilla con objeto de certificar la admisión definitiva a valoración y comunicarla oficialmente a los interesados/as.



¡¡Visítanos!!

www.sociedadevaluacion.org

La Sociedad Española de Evaluación de Políticas Públicas (SEE) se constituye con el objetivo de contribuir al desarrollo e implantación en España de la evaluación de políticas, programas y proyectos públicos, como instrumento fundamental para la mejora de la eficacia, eficiencia y utilidad social de la actividad realizada por la Administración Pública.

Para ello, la SEE fomenta la investigación científica y la difusión de experiencias sobre evaluación entre los agentes e instituciones implicados, tanto en la esfera universitaria como en la función pública, y en el ámbito empresarial y profesional. La presencia de la evaluación en los procesos de decisión y en el debate público se considera como uno de los objetivos primordiales.

La SEE considera imprescindible elevar la calidad de la práctica evaluadora, formando expertos, verificando la validez de los métodos usados y proponiendo mejoras en su aplicación. En este sentido, reúne a especialistas en la materia, tanto de las Administraciones Públicas como de cualquier organismo que aplique la evaluación en su gestión, propiciando la comunicación y el debate entre sus miembros para incrementar su uso, así como para confrontar experiencias sobre sus resultados.

SOCIEDAD ESPAÑOLA DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

Universidad de Sevilla
Facultad de Ciencias Económicas
Avda. Ramón y Cajal, 1. C.P.41018 Sevilla
tlf. 954556636 Horario (9:30 h.-14:30 h.)
secretaria@sociedadevaluacion.org

SE-e-valoración

SUMARIO

4 EVALUACIÓN DE POLÍTICAS PÚBLICAS EN MÉXICO. LA EXPERIENCIA DE PROGRESA Y EL CONSEJO NACIONAL DE EVALUACIÓN DE POLÍTICAS PÚBLICAS
Rodolfo Canto Sáenz.

13 PANORAMA DE LA EVALUACIÓN DE LOS PROGRAMAS SOCIALES EN MÉXICO
Sindy Agustina González Tijerina.

20 ALGUNAS CUESTIONES CONCEPTUALES Y METODOLÓGICAS EN TORNO A LA EVALUACIÓN DE IMPACTO
José María Larrú.

32 EL ESTUDIO DE LA EVALUABILIDAD COMO CONTRIBUCIÓN A EVALUAR CON CALIDAD
Itziar Arispe.

36 ENTREVISTA A D^o. Oneida Álvarez.

40 EVENTOS, FORMACIÓN Y NOTICIAS

EDITORIAL

La buena noticia de la magnífica acogida que ha tenido la Convocatoria del I^o Premio de Investigación Carlos Román, confirma las esperanzas puestas en la fuerza que tiene la herramienta, la Evaluación de Políticas Públicas. En efecto, ocho trabajos de diferentes territorios del Estado – Barcelona, Navarra, Almería, Sevilla, Madrid, Cartagena y Granada – y con orígenes diversos – de universidades, de consultoras, de organismos públicos... – en tiempos de escasez, con un tema tan acotado y sin haber podido, por falta de recursos económicos, realizar mayor difusión que la que proporciona en la red nuestra web es, sin duda, un éxito que queremos compartir con los lectores y lectoras de “e-valoración”.

En el presente número de “e-valoración” encontramos de una parte, reflexiones sobre la institucionalización de la evaluación de las políticas públicas que abundan en la necesidad de fortalecer la capacitación de las administraciones públicas para gestionar la evaluación y, de otra análisis, relacionados con las “tipologías de evaluación”, como la evaluación de impacto o la importancia del marco normativo en el que se desarrolla la evaluación para comprenderla en su carácter evolutivo.

En el primer bloque situamos los trabajos de Rodolfo Canto Sáenz y Sindy González; en el segundo se encuadran, los de José María Larrú e Itziar Arispe. Necesariamente, “e-valoración” felicita explícitamente a los cuatro por su original reflexión, aportaciones de rigor; a la par que les agradece su inestimable colaboración.

El Prof. Canto, abre de nuevo el debate sobre la institucionalización de la evaluación de políticas públicas desde la aplicación práctica que supone el caso PROGRESA (Programa de Educación, Salud y Alimentación) y la consideración crítica del intento de creación del Consejo Nacional de Evaluación de Políticas Públicas en México. Precisamente, sobre el CONEVAL, la investigadora González hace una detallada descripción e interesante disertación sobre sus características y funciones. Con todo ello, una muestra relevante de que “algo se mueve” en México, en América Latina, con la Evaluación de Políticas Públicas.

Al respecto de las “tipologías”, el Prof. Larrú hace una reflexión sobre los errores que – desde la concepción del diseño de una evaluación de impacto – se pueden cometer si se comienza preguntando por las herramientas y metodologías a aplicar, obviando la importancia del contexto sociopolítico en cada momento para dar una correcta respuesta a la misma. En continuidad con la lectura anterior, Itziar Arispe enfatiza sobre la importancia del marco normativo en la evaluación y su evolución y nos introduce, muy sugerentemente, en el análisis del concepto de evaluabilidad técnica y evaluabilidad política; que, por otra parte, son habituales en la organizaciones internacionales.

En fin, cuatro brillantes y sugerentes análisis que desde “e-valoración”, estamos convencidos que os invitarán al juicio crítico, al debate y retroalimentarán vuestras iniciativas e ideas para desarrollar los artículos de nuestros próximos números.

Edita: SOCIEDAD ESPAÑOLA DE EVALUACIÓN
www.sociedadevaluacion.org

Presidente: José Luís Osuna.
Vicepresidente: Fernando González Laxe.
Vocales: Carmen Vélez, Joaquín Pérez-Hervada,
M^a Angeles Díez, Ester García, Juan Carlos Díaz,
Rafael Monterde.

Coordinadora de contenidos: Ángeles María Sánchez.
Foto de Portada: Juan Vázquez.
Foto de Contraportada: Carlos Román.
Maquetación y diseño: Rocío Jiménez Alba.
Impresión: Gráficas Santa María, S.C.A.
ISSN: 1887-1399
Deposito Legal: SE-4379-06

Los costes de esta revista se adaptan a la Orden 26 de diciembre de 2007, por la que se establecen las normas reguladoras de concesión de ayudas de los Programas de Estudios y de Difusión del Mercado de Trabajo, en desarrollo del Decreto 85/2003, de 1 de abril.



EVALUACIÓN DE POLÍTICAS PÚBLICAS EN MÉXICO LA EXPERIENCIA DE PROGRESA Y EL CONSEJO NACIONAL DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

Rodolfo Canto Sáenz

Profesor Investigador de la Universidad Autónoma de Yucatán, México

INTRODUCCIÓN

El conocimiento y la práctica de la evaluación de políticas públicas tienden a generalizarse entre los países de América Latina. Una revisión de las experiencias latinoamericanas en la materia lleva a concluir que los gobiernos y las sociedades de la región coinciden en que la evaluación de políticas públicas es importante e incluso imprescindible para la buena marcha y el mejoramiento continuo de la función pública.

Sin embargo, posiblemente sea ésta la única coincidencia. Al aproximarnos a las experiencias concretas de los distintos países latinoamericanos descubrimos una profunda heterogeneidad entre lo que se entiende en cada uno de ellos por evaluación de políticas públicas y por el concepto mismo de evaluación: sus objetivos, funciones, alcances y métodos, así como su ubicación en el marco institucional de la administración pública.

En cuanto a la naturaleza y objetivos de la función las concepciones varían desde una evaluación básicamente gerencial, útil para generar informes que

apoyen la toma de decisiones, hasta una evaluación integral u holística, que se concibe como ubicua en todas las etapas del proceso de la política pública, desde el surgimiento del problema y su inclusión en la agenda pública hasta el diseño e implementación de la política misma.

Tiende a prevalecer en América Latina una concepción que podríamos ubicar a medio camino entre ambas: una evaluación cuya función esencial es la retroalimentación, pero no en el sentido gerencial de emisión de reportes sino capaz de generar, mediante la investigación y el análisis, nuevo conocimiento que contribuya al mejoramiento de todas las fases del proceso y, especialmente, contribuya al rediseño de la política, a una implementación más eficaz y en general a su mejora continua.

Otro gran tema es el de la ubicación institucional de la función. En algunos países prevalece la idea de que la evaluación es básicamente una función del mismo ente que diseña e implementa la política pública, es decir el poder ejecutivo, ya sea el nacional o los subnacionales; en general se acepta, en este mismo

tenor, la posibilidad de que el ejecutivo en cuestión contrate los servicios de consultores externos que le ayuden a hacer mejores evaluaciones.

Otra gran corriente, en cambio, sostiene que las evaluaciones deben estar a cargo de un poder distinto del que diseña e implementa la política a causa del riesgo, que se presume grande, de que la evaluación no sea imparcial sino adecuada a los intereses del ejecutivo. Usualmente se asume que las evaluaciones deben ser realizadas por el poder legislativo o por una entidad con autonomía frente al ejecutivo a la manera de las universidades públicas o los bancos centrales, por ejemplo. Una gran preocupación, en todo caso, es garantizar la objetividad e imparcialidad de las evaluaciones, algo que ninguna alternativa parece capaz de asegurar por completo.

En general el tema de la evaluación sigue incorporando en América Latina un fuerte componente de polémica. Es, como en todo el mundo, un área sensible, por los costos o beneficios políticos que implica para los actores involucrados; muchos la siguen viendo como el arma arrojada que será lanzada en una u otra dirección según se trate de los responsables de la política en cuestión o de sus adversarios políticos.

Contribuye a la polémica la aún escasa institucionalización de la evaluación; en la mayoría de los países no existen todavía organismos evaluadores formalmente constituidos; suele ocurrir que la evaluación como tal siga siendo una preocupación proveniente del mundo académico que sólo lentamente se va transmitiendo a las esferas públicas; falta mucho para que la evaluación de políticas públicas alcance el grado de institucionalización que tienen las auditorías, por ejemplo. Con todo, nadie pone en tela de juicio la relevancia de la evaluación. Se sabe que es una función imprescindible y de la mayor trascendencia para la función pública; no siempre se sabe, en cambio, cómo, cuándo, dónde y quién debe realizarla.

Una de las áreas más desarrolladas en la región es la evaluación de políticas sociales, que acumula ya experiencias por demás interesantes, como la del programa PROGRESA-Oportunidades desarrollado en México durante la última década. Este mismo país proporciona, con el proyectado Consejo Nacional de Evaluación de Políticas Públicas, un buen ejemplo de lo que suele ocurrir cuando el tema de la evaluación se convierte en un arma arrojada, algo que siempre temen los estudiosos y promotores de la evaluación de políticas públicas.

EL PROGRAMA DE EDUCACIÓN, SALUD Y ALIMENTACIÓN

Este programa, mejor conocido por su acrónimo PROGRESA, fue diseñado y comenzó a implementarse durante el gobierno del presidente Ernesto Zedillo (1994-2000). Fue un programa altamente innovador y exitoso, a tal grado que el siguiente gobierno, encabezado por el presidente Vicente Fox (2000-2006), de signo político distinto, lo mantuvo vigente con el nombre de Oportunidades, en un ejemplo poco común en el país de continuidad en la política pública. El gobierno del presidente Felipe Calderón (2006-2012) lo mantiene en operación hasta la fecha.

En gobierno del presidente Zedillo encargó, hacia el final de su periodo, una evaluación del programa PROGRESA a la firma de consultoría especializada International Food Policy Research Institute (IFPRI) con sede en Washington, D. C. La información relativa al programa que presentamos a continuación proviene del reporte de evaluación elaborado por el IFPRI (Skoufias, 2005).

PROGRESA fue innovador en dos sentidos: por un lado, intervino simultáneamente en salud, educación y nutrición, a partir de la idea de que las tres dimensiones del bienestar son interdependientes y que la

pobreza en salud, educación y nutrición son a la vez causas y consecuencias de la pobreza. Por otro lado, PROGRESA fue diseñado desde el comienzo para ser evaluado y mejorado continuamente, de modo tal que su efectividad en el mejoramiento del bienestar de los más pobres fuera cada vez mayor.

Las evaluaciones realizadas por el IFPRI evidenciaron los aspectos exitosos del programa, a la vez que señalaron áreas susceptibles de mejoría. Por ejemplo, la investigación demostró que PROGRESA había contribuido a que los niños pobres permanecieran más años en la escuela, había mejorado la salud de los infantes y adultos de los hogares beneficiados, aumentado la atención prenatal de las mujeres y mejorado la nutrición de los niños. Por otro lado, la evaluación reveló que PROGRESA habría tenido un mayor impacto en la permanencia escolar de haber focalizado su atención en la asistencia a la escuela secundaria y preparatoria, la etapa en que muchos niños pobres abandonan la escuela.

PROGRESA partió de una cuidadosa focalización de la población objetivo: las comunidades incluidas en el programa fueron seleccionadas con base en un índice de marginación elaborado con información de los censos de población y vivienda. En las localidades elegidas, los hogares beneficiarios fueron seleccionados con base en información socioeconómica que incluía a todos los hogares de la comunidad.

PROGRESA fue también innovador en el mecanismo de transferencia de los apoyos monetarios. En vez de entregarlos a los padres de familia optó por darlos exclusivamente a las madres, al reconocer la mayor capacidad de éstas para usar eficazmente los apoyos en la atención a las necesidades familiares.

En los hogares seleccionados, los niños inscritos en los grados 3 a 9 reciben apoyos monetarios cada dos meses. El monto de los apoyos se calculó, entre otros factores, con base en lo que cada niño podía aportar

a su familia en caso de trabajar. Las becas educativas son algo mayores para las niñas inscritas en secundaria, dada su mayor propensión a abandonar la escuela. Cada dos meses, los profesores y directores de las escuelas informan a PROGRESA de la asistencia de los niños beneficiarios y si ésta ha sido de al menos el 85 por ciento de las sesiones se entrega el apoyo correspondiente.

En el área de la salud y la nutrición, PROGRESA promueve el cuidado de la salud a través de medidas de atención preventiva tales como distribución de suplementos nutricionales, educación en higiene y nutrición y también transferencias monetarias para la compra de alimentos. Las entregas de los apoyos monetarios están sujetas a visitas periódicas obligatorias a las clínicas públicas.

El apoyo monetario mensual promedio entregado a los hogares beneficiarios equivale al 20 por ciento de los gastos mensuales de consumo previos al inicio del programa. A finales de 1999, PROGRESA cubría aproximadamente 2.6 millones de familias, el 40 por ciento de las familias rurales; operaba en cerca de 50 mil localidades, en más de dos mil municipios y en los 31 estados de la República.

El IFPRI recuerda que hasta la implementación de PROGRESA en México, y en la mayoría de los países en desarrollo, los programas de combate a la pobreza no eran habitualmente sujetos a una evaluación rigurosa. En los raros casos en que los programas eran evaluados la decisión de hacerlo era usualmente tomada años después de su implementación, cuando ya era demasiado tarde porque el programa había alcanzado una cobertura tan amplia que ya era imposible elaborar comparaciones entre grupos de beneficiarios y no beneficiarios.

PROGRESA se distinguió porque los elementos esenciales para una rigurosa evaluación de su impacto fueron contemplados desde las primeras fases de

la implementación del programa. La estrategia adoptada para la evaluación incluyó dos elementos fundamentales:

1. La adopción de un diseño experimental, que permitió la medición del impacto del programa mediante la comparación de indicadores claves entre hogares beneficiarios (grupo en tratamiento) con hogares similares que aún no eran incluidos en el programa (grupo de comparación y control) El diseño experimental de PROGRESA contemplaba la remoción sistemática de la influencia de otros factores que pudieran haber contribuido a los cambios observados.
2. La recopilación de información de los dos grupos (grupo en tratamiento y grupos de comparación) antes y después de la implementación del programa.

La muestra total utilizada en la evaluación de PROGRESA incluyó 24 mil hogares en 506 localidades de siete estados de la República Mexicana. De las 506 localidades, 320 fueron beneficiarias del programa y 186 de control. Los 24 mil hogares fueron entrevistados a lo largo de dos años; también se llevaron a cabo talleres con grupos de beneficiarios, líderes locales, oficiales de PROGRESA, trabajadores de clínicas de salud y profesores de escuela. Algunos de los resultados de la evaluación de impacto en los hogares beneficiarios son los siguientes:

- ▶ PROGRESA ha generado un incremento de la matrícula en la escuela secundaria de más de 20 por ciento en el caso de las niñas y de 10 por ciento en los niños.
- ▶ El mejoramiento en la nutrición y el cuidado preventivo de la salud ha hecho a los niños pequeños más robustos frente a las enfermedades. Concretamente, ProgresA ha reducido en 12 por ciento la tasa de morbilidad entre los niños de uno a cinco años.
- ▶ Los beneficiarios adultos de PROGRESA han re-

ducido en 17 por ciento los días de incapacidad por enfermedad y en 22 por ciento los días en cama.

- ▶ En 1999 los gastos medios en alimentos fueron 13 por ciento más altos en los hogares PROGRESA que en los de control; aumentó el consumo de frutas, vegetales, carnes y productos animales. El consumo medio de calorías se elevó en 10.6 por ciento. Los beneficiarios sienten que desde el inicio de PROGRESA en los hogares rurales pobres se come mejor.
- ▶ PROGRESA ha tenido un impacto significativo en la talla de los niños, con un incremento de 16 por ciento anual (equivalente a un centímetro) para los niños entre uno y tres años de edad.
- ▶ Los costos administrativos de PROGRESA son relativamente pequeños en comparación con los costos de programas anteriores y en programas similares en otros países. El análisis de costos del programa ubicó en 8.6 pesos los costos administrativos por cada 100 pesos transferidos.

Como se ha indicado, la evaluación del programa a cargo del IFPRI facilitó su permanencia en el gobierno del presidente Vicente Fox y también su mejoría: la evaluación reveló que el mayor impacto de PROGRESA en la matrícula escolar se daba en el nivel de preparatoria, por lo tanto los apoyos se reorientaron y reforzaron en ese nivel. También, el programa original otorgaba los apoyos a cambio de la asistencia a la escuela, sin tomar en cuenta el desempeño de los estudiantes. Oportunidades (el nuevo nombre del programa) mejoró el diseño de este componente al vincular los beneficios al desempeño escolar, incluyendo bonos adicionales por el cumplimiento exitoso de un grado, o ligando los beneficios a la participación en otros programas, como el de Jóvenes con Oportunidades, que contempla el acceso a microcréditos, mejoras en el hogar, educación para adultos y seguro social.

El IFPRI no duda en afirmar que la evaluación de PROGRESA ha instaurado un conjunto de estándares más elevados para el diseño y ejecución de la política social en México y en América Latina y ha incrementado el conocimiento sobre medidas de política pública que los gobiernos pueden tomar para mejorar las capacidades de la gente pobre. Señala que programas similares a PROGRESA están siendo probados o han sido implementados en otros países latinoamericanos como Brasil, Honduras y Nicaragua y es promisoría su posible adaptación a países de África y Asia.

El IFPRI concluye que se requiere investigación adicional para responder a preguntas derivadas de la evaluación de PROGRESA, por ejemplo si las transferencias con menores condicionamientos (lo que reduciría los costos administrativos) pueden tener impactos similares en la formación de capital humano de las familias pobres, si el monto de las transferencias a cada familia es el adecuado o si con menores montos se podrían lograr resultados similares (lo que permitiría extender la cobertura del programa); si es posible que similares o incluso mejores resultados en la asistencia escolar pudieran lograrse con programas alternativos, por ejemplo la construcción de nuevas escuelas o la mejora de los servicios educativos; por último, sostiene que la cuestión fundamental de dónde romper el círculo vicioso de la pobreza y su transmisión intergeneracional sólo podrá responderse con el seguimiento de las cohortes de niños beneficiarios.

EL CONSEJO NACIONAL DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

Como parte de una propuesta de reforma fiscal integral, en el mes de junio de 2007 la Secretaría de Hacienda del gobierno mexicano sometió a la consideración del Congreso de la Unión la propuesta de crear el Consejo Nacional de Evaluación de Políticas

Públicas (CNEPP) que sustituiría, con funciones mucho más amplias, al Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), organismo creado en 2004 en el marco de la Ley General de Desarrollo Social. El nuevo CNEPP tendría las siguientes características:

- ▶ Organismo público descentralizado con personalidad jurídica, patrimonio propio, autonomía técnica y de gestión.
- ▶ Un Comité Directivo como órgano de gobierno de la entidad; una Comisión Ejecutiva Evaluadora, a cargo de las funciones sustantivas de evaluación, y un presidente a cargo de la dirección del propio Consejo.
- ▶ El Comité Directivo estaría integrado por el presidente del Consejo, un representante de la dependencia coordinadora del sector y seis consejeros, expertos en materia de evaluación de políticas públicas o, en su caso, investigadores académicos que fueran o hubieran sido miembros del Sistema Nacional de Investigadores, con amplia experiencia en la evaluación de políticas y programas públicos y que colaborasen en instituciones de educación superior y de investigación, en programas inscritos en el Padrón Nacional de Postgrado del Consejo Nacional de Ciencia y Tecnología. Los consejeros no deberían haber ocupado puestos públicos, cargos de elección popular ni cargos directivos en los partidos políticos durante los tres años anteriores a su designación.
- ▶ Los consejeros serían designados por el Presidente de la República de forma escalonada, por medio de una convocatoria y con base en las listas de candidatos que propondrían tres instituciones: la Comisión Nacional de Desarrollo Social, para seleccionar a cuatro consejeros; el Consejo Nacional para Competitividad de la Micro, Pequeña y Mediana Empresa, para seleccionar a un consejero, y el Consejo Nacional de Seguridad Pública, para seleccionar al sexto consejero.

- ▶ La Comisión Ejecutiva Evaluadora estaría integrada por los seis consejeros y el presidente del organismo.
- ▶ El Consejo, por medio de la Comisión Ejecutiva Evaluadora, realizaría las evaluaciones por sí mismo o mediante la contratación de personas físicas y morales especializadas y con experiencia probada en la materia que fuera a evaluarse, independientes del ejecutor de las políticas, los programas y los recursos públicos. Estas contrataciones también podrían realizarse por la dependencia o entidad en cuestión, por sí mismas o por medio de los mecanismos formalmente establecidos, siempre y cuando se garantizara la independencia del evaluador.

El CNPP, como se ha señalado, estaba inspirado en el CONEVAL, dedicado a la evaluación exclusiva de las políticas sociales. La positiva experiencia de la evaluación de PROGRESA demostró la utilidad de la evaluación, al poner en relieve los logros de ese programa, por demás exitoso, así como sus áreas susceptibles de mejoría. El CONEVAL representa en sí mismo un resuelto compromiso del gobierno mexicano con la evaluación de políticas públicas, específicamente de políticas sociales. La Ley de General de Desarrollo Social establece que la evaluación de la política de desarrollo social estará a cargo del CONEVAL, que podrá realizarla por sí mismo o a través de uno o varios organismos independientes del ejecutor del programa, y tiene por objeto revisar periódicamente el cumplimiento del objetivo social de los programas, metas y acciones de la política de desarrollo social para corregirlos, modificarlos, adicionarlos, reorientarlos o suspenderlos total o parcialmente.

Entre sus funciones incluye alcanzar la mayor calidad en el desarrollo de lineamientos, metodologías, normas, modelos e instrumentos de evaluación, atendiendo a criterios de confiabilidad, validez, relevancia, transparencia y representatividad, y procurar que los resultados de las evaluaciones sean comparables en el tiempo. Le corresponde incluso revisar y aprobar los indicadores de resultados, gestión y servicios que

de manera invariable deben incluir los programas sociales para su evaluación.

Con la propuesta de creación del CNEPP el Ejecutivo mexicano buscaba extender la experiencia del CONEVAL más allá de la política social, creando un organismo responsable de la evaluación de todo el universo de la política pública. Este propósito, sin embargo, se frustró por las suspicacias que entre los partidos de oposición generó la posible creación de un organismo evaluador con poderes tan amplios.

Los opositores Partido de la Revolución Democrática (de izquierda) y Partido Revolucionario Institucional (de centro izquierda) rechazaron la creación del CNPP desde su análisis en la Comisión de Presupuesto y Cuenta Pública del Congreso de la Unión. Al argumentar su rechazo, los diputados del PRD sostuvieron que el nuevo organismo otorgaría mayores atribuciones a la Secretaría de Hacienda para evaluar y fiscalizar el gasto y programas gubernamentales y duplicaría funciones con la Auditoría Superior de la Federación, órgano de la Cámara de Diputados que, constitucionalmente, es la única instancia legal para fiscalizar la administración pública.

De modo particular, los diputados de oposición rechazaron la propuesta de modificación del artículo 134 de la Constitución Mexicana, que estipulaba que los recursos federales de los gobiernos federal, estatales y municipales, así como los órganos político-administrativos de sus demarcaciones territoriales y sus respectivas administraciones públicas paraestatales serían evaluados por "instancias técnicas". Según los opositores, semejante reforma constitucional dejaría abierta la posibilidad de crear cualquier instancia evaluadora que le pareciera al Ejecutivo, cuando la instancia constitucional para evaluar las políticas públicas es la Auditoría Superior de la Federación que quedaría reducida, en los términos de la propuesta, a una mera instancia de opinión sobre los resultados de las políticas públicas.

Además, según los diputados de oposición, se le quitaría a los congresos estatales la facultad de evaluar las políticas públicas en sus entidades y el empleo de mecanismos para obligar a los gobernadores (ejecutivos locales) a transparentar su gasto. En su opinión, el CNEPP daría superpoderes a la Secretaría de Hacienda, que inventaría indicadores y haría autoevaluaciones a su antojo, capricho y criterio. Con la creación del CNEPP, argumentaban, el gobierno federal buscaba un mayor centralismo al quitar atribuciones a las entidades y sus congresos y atentaba contra el equilibrio de poderes.

Sin embargo, los opositores dejaron abierta una posibilidad: reconocieron que en distintos países hay organismos evaluadores de resultados pero son autónomos o cuentan con la participación de los legisladores. Si el Ejecutivo Federal quería una instancia evaluadora de este tipo tendría que tomar en cuenta a los legisladores, a los gobernadores y a los presidentes municipales. De otro modo sólo se estaría dando a sí mismo poderes absolutos al designar a un grupo de burócratas que lo evaluarían “a modo”, sin que siquiera fueran ratificados por el Congreso.

Otras reservas expresadas por los opositores y algunos analistas iban en el sentido de que con la creación del CNPP se autorizaba a las dependencias del gobierno a contratar sus evaluadores, lo que de entrada ponía en tela de juicio la independencia de éstos, por aquello de que “el que paga manda”. También sostenían que se abría la puerta para que no sólo las universidades, centros de investigación y organizaciones no lucrativas evaluaran los programas gubernamentales, desde los de combate a la pobreza hasta la compra de armamento, sino que también lo hicieran las personas físicas, o sea, cualquier persona que para las dependencias evaluadas fuera un especialista.

Los lapidarios argumentos de la oposición acabaron con el proyecto del Consejo Nacional de Evaluación

de Políticas públicas. Sólo quedó en pie el CONEVAL, que se limita a la evaluación de políticas sociales. Es poco probable que en el corto plazo el gobierno federal haga otro intento de crear un organismo evaluador nacional.

COMENTARIOS FINALES

Son varios los temas de reflexión que se derivan del análisis de las experiencias mexicanas de PROGRESA y el CNEPP. Uno de ellos es el de la independencia de las evaluaciones; todo el mundo coincide en que éstas deben independientes pero, ¿qué significa independencia cuando hablamos de evaluación de políticas públicas?

En principio parece plausible la idea de que la independencia de la evaluación debe garantizarse en el marco de la división de poderes; si la política es diseñada e implementada por el ejecutivo parece razonable que las evaluaciones queden a cargo del legislativo, por ejemplo. Ésta es una tesis que se enmarca muy bien en la matriz democrática del análisis de políticas públicas en nuestros países. Sin embargo, cuando nos aproximamos al debate las respuestas ya no parecen tan claras y hay varios aspectos que ameritan mayor reflexión.

Por un lado, el riesgo de que la evaluación quede empuñada en el juego de los colores partidistas existe en ambos sentidos: si es cierto que cabe la posibilidad de que una evaluación a cargo del mismo poder que diseña e implementa la política pública resulte favorable a sus intereses políticos por aquello de que “el que paga manda”, también existe el mismo riesgo cuando el poder que evalúa está controlado por colores opositores. El arma arrojada puede volar en ambas direcciones y, dado el caso, cercenar un programa público que debía mantenerse.

Por otra parte, aún no se acaba de comprender del todo la naturaleza y funciones de la evaluación; como se ha indicado, éstas suelen ser más claras en los círculos públicos cercanos a los sectores académicos que la promueven, usualmente en ciertas dependencias del ejecutivo (como las encargadas de la política social, por ejemplo) que en otras, a cargo de áreas menos visibles de la política pública, o bien en los otros poderes públicos. En el caso del CNEPP mexicano fue muy clara la confusión de algunos opositores entre la evaluación y la auditoría; su argumento de que no hacía falta un nuevo organismo encargado de la evaluación de la política pública cuando ya existía la Auditoría Superior de la Federación es por demás explícito.

Por lo demás, la experiencia de PROGRESA ejemplifica muy bien la posibilidad cierta de que el poder que diseña e implementa procure una buena evaluación; más aún, como el propio organismo evaluador se encargó de destacar, la eficaz evaluación del programa fue posible gracias a que desde el diseño y las fases tempranas de la implementación se contemplaron los métodos e instrumentos necesarios para una buena evaluación.

Desde luego, en modo alguno se trata de afirmar que las evaluaciones deben quedar a cargo de los poderes que diseñan e implementan. La alternativa de enmarcar la evaluación en la división de poderes quizá siga siendo la mejor, pero parece indispensable que el poder que evaluará tenga claridad sobre la naturaleza y funciones de la evaluación; que esté convencido de su necesidad y que no la confunda con otras funciones como la auditoría o el control de gestión; en todo caso, ambos poderes, el evaluador y el evaluado, tendrían que coincidir en su concepto de evaluación, así como en los mecanismos, instrumentos y métodos apropiados para realizarla.

Parece más promisorio la tesis de que el ente responsable de la evaluación sea un organismo autóno-

mo del poder ejecutivo, a la manera de los bancos centrales, por ejemplo. En realidad éstos son un buen ejemplo de la posibilidad cierta de desvincular no sólo la evaluación, sino también el diseño y la implementación de un área de la política pública que se presume crucial para la buena marcha de la economía, del cotidiano juego de intereses y colores políticos.

Ahora bien, no parece haber ninguna razón lo suficientemente fuerte para oponerse a que el poder ejecutivo haga sus propias evaluaciones. En realidad, si retomamos el hilo conductor del análisis de políticas públicas recordaremos que una idea fundamental es que el diseño, la implementación y la evaluación forman parte del mismo proceso y están indisolublemente ligados entre sí. De hecho, es deseable que los ejecutores de la política pública hagan sus propias evaluaciones, cuyo interés estaría fuera de duda toda vez que la propia experiencia de la implementación les genera conocimiento de primera mano útil para todas las etapas del proceso.

El tema de la naturaleza y objetivos de la evaluación ciertamente precisa de mayor reflexión y análisis. Podemos pensar en objetivos un tanto diferentes según se trate de una evaluación realizada por el ejecutivo o por el legislativo; en el primer caso, es probable que la evaluación se aproxime más al objetivo de la retroalimentación, como insumo imprescindible de la mejora continua; en el segundo, tal vez se enfatice el objetivo de la rendición de cuentas. En todo caso, podemos pensar en evaluaciones paralelas del ejecutivo y del legislativo, o bien del órgano evaluador autónomo, cuyos resultados no tendrían que ser muy diferentes entre ellos si realmente se llevaron a cabo con objetividad e imparcialidad.

Al lado del tema de la independencia de la evaluación hay otros igualmente relevantes, como el de la especialización. No es ajeno al logro de una eficaz evaluación de PROGRESA el hecho de que el gobierno mexicano haya recurrido a una firma de con-

sultoría especializada en investigaciones de política alimentaria para realizarla. La política pública es tan amplia como la vida misma y la probabilidad de que una misma agencia evaluadora sea capaz de realizar buenas evaluaciones en todos o en muchos de sus campos es más bien baja. En Estados Unidos existen firmas especializadas en prácticamente cualquier área de la política pública; usualmente no se limitan a la evaluación, sino que se extienden a todo el proceso (formación de la agenda, diseño, implementación, evaluación) pero de una determinada área de la política pública. Naturalmente, con el tiempo y las sucesivas experiencias van acumulando un gran acervo de conocimientos que emplean en cada nueva experiencia, como ha sido el caso del IFPRI y el PROGRESA mexicano.

En América Latina, donde la evaluación de políticas públicas –y el análisis de políticas en general– es-

tán mucho menos desarrollados, todavía no existen suficientes agencias especializadas en los distintos campos, y ésta es probablemente la razón de que el gobierno mexicano haya recurrido a una firma estadounidense para evaluar el PROGRESA. Por lo demás, uno de los cuestionamientos al CNEPP hacía referencia a la integración de su Comisión Ejecutiva Evaluadora, que incluía seis consejeros de los cuales cuatro serían expertos en política social, uno más en lo que podríamos llamar política industrial y el último en seguridad pública. Los opositores se preguntaban por las demás áreas de la política pública. Así, parece necesario que una eventual nueva propuesta de creación de un organismo evaluador de cobertura nacional incluya una comisión ejecutiva lo suficientemente amplia para dar cuenta de la pluralidad de áreas de la política pública, que son por lo menos tantas como carteras existen en el gabinete del ejecutivo.

REFERENCIAS BIBLIOGRÁFICAS

Araya Umaña, Sandra, Caminos recorridos por las políticas educativas de género, Instituto de Investigación para el mejoramiento de la Educación Costarricense, Universidad de Costa Rica, 2003.

Canto Sáenz, Rodolfo, Políticas públicas. Más allá del pluralismo y la participación ciudadana, en *Gestión y Política Pública*, Vol. IX, No. 2, Centro de Investigación y Docencia Económica, México, segundo semestre de 2000.

Hintze, Susana, Reflexiones sobre el conflicto y la participación en la evaluación de políticas sociales, revista *Reforma y Democracia* No. 21, CLAD, Caracas, 2001.

Lahera P. Eugenio, Política y políticas públicas, CEPAL, Serie Políticas Sociales No. 95, Santiago, 2004.

Osuna Llana, José Luis, Carmen Vélez Méndez, Ana Cirera León y Juan Murciano Rosado, Programación y evaluación pública: un triángulo complejo, en *Ekonomiaz*, revista vasca de economía, No. 60, Vol. I, tercer cuatrimestre, 2005.

Peña Cedillo, Jesús, Evaluación de políticas públicas en derechos humanos. Una introducción al caso venezolano (1994-1999), *Latin American Studies Association*, 2000.

Román del Río, Carlos, Otro estado, en *E-valoración*, revista de la Sociedad Española de Evaluación, Año 0 No. 1, octubre de 2006.

Skoufias, Emmanuel, PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico, Research Report 139, International Food Policy Research Institute, Washington, DC.

Subirats Humet, Joan, Catorce puntos esenciales sobre evaluación de políticas públicas con especial referencia al caso de las políticas sociales, en *Ekonomiaz*, revista vasca de economía, No. 60, Vol. I, tercer cuatrimestre, 2005.

Vaitsman, Jeni, Roberto W. S. Rodrigues y Rômulo Paes-Sousa, El Sistema de Evaluación y Seguimiento de las políticas y programas sociales: la experiencia del ministerio de Desarrollo Social y Lucha contra el Hambre del Brasil, UNESCO, Policy Papers/17.

Zaltsman, Ariel, Desarrollo de la capacidad de evaluación. Experiencia con la institucionalización de monitoreo y evaluación en cinco países latinoamericanos: Argentina, Chile, Colombia, Costa Rica y Uruguay, Grupo de Evaluación Independiente, Banco Mundial, 2006.

PANORAMA DE LA EVALUACIÓN DE LOS PROGRAMAS SOCIALES EN MÉXICO



Sindy Agustina González Tijerina
Estudiante de Doctorado en Políticas Públicas
Escuela de Graduados en Administración Pública y Política Pública
Instituto Tecnológico y de Estudios Superiores de Monterrey

INTRODUCCIÓN

La evaluación de los programas sociales en México es un tema relativamente reciente, en construcción y constante perfeccionamiento. La estrategia de política social implementada en la actual administración del presidente Felipe Calderón titulada Vivir Mejor tiene un elemento que la hace diferente a estrategias pasadas, las acciones y políticas contenidas en ella serán sometidas a evaluación constante de acuerdo a la normatividad y legislación vigente. Estas evaluaciones serán públicas y tienen como propósito la mejora constante de los programas y un uso más eficiente de los recursos públicos.

En el año 2005 entro en función el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) el cual es el organismo encargado de coordinar la evaluación de las políticas y programas de desarrollo social que ejecuten las dependencias públicas y al mismo tiempo es encargado de medir los niveles de pobreza. El CONEVAL considera a la evaluación de la política social, y de la política pública como “una herramienta fundamental para mejorar constantemente su desempeño y conocer cuáles de

las acciones son o no efectivas para resolver los grandes problemas sociales y económicos que aquejan al país”.

En el presente trabajo se presentara un breve panorama de la evaluación de los programas sociales en México. En la siguiente sección se abordan los inicios de la evaluación de los programas sociales, la sección tres expone las características y funciones del CONEVAL. La sección cuatro presenta los diferentes tipos de evaluación de programas sociales que propone el CONEVAL. En la última sección se presentan algunas consideraciones finales.

INICIO DE LA EVALUACIÓN DE LOS PROGRAMAS SOCIALES EN MÉXICO

En México durante mucho tiempo imperó la ausencia de una cultura de evaluación sustentada en análisis y metodologías rigurosos, aunado a escasos incentivos en transparencia y rendición de cuentas en el sector público porque el país estuvo gobernado por un solo partido político desde la década de los 20. En la

década de los 90 se presentan cambios importantes en el control político y la transición a la democracia, se hacen algunas evaluaciones aisladas de programas sociales en el país, sin embargo estas iniciativas buscaban ejercer solamente control interno y no tanto rendir cuentas (Castro, López-Acevedo, Beker y Fernández ,2009). De manera sistemática y rigurosa hasta 1997 se evalúa el Programa de Educación, Salud y Alimentación (Progresá) hoy Oportunidades, el Banco Interamericano de Desarrollo (BID) dona dinero al programa pero requiere de una evaluación externa.

En 1998, el Congreso introduce un requerimiento especial para los programas federales que proveen subsidios, las Reglas de Operación de los Programas ROP (Castro et al, 2009). Las ROP son el instrumento normativo y regulatorio que permite a los programas su implementación y funcionamiento, procurando ser eficientes, eficaces e incluyentes, buscan facilitar la evaluación de los beneficios económicos y sociales y asegurar la coordinación de acciones con otras dependencias y entidades (Fuentes, 2009).

En 1999 el Congreso de la Unión requiere por ley una evaluación anual externa de todos los programas federales para lograr mayor transparencia y rendición de cuentas a partir del 2000. En el año 2002 se emitió el Acuerdo de Requisitos Mínimos que establecía los elementos mínimos que debían contener las evaluaciones externas con el propósito de perfeccionarlas (Hernández, 2006).

En 2004 se introduce la Ley General de Desarrollo Social (LGDS) y esta legislación crea al CONEVAL el cual entra en función a partir del 2005 y es el organismo encargado de evaluar los programas sociales y medir la pobreza. La LGDS en 2004 considera por primera vez en nuestro país la evaluación sistemática de la política de desarrollo social y de los programas sociales con la el propósito de corregirlos, modificarlos, ampliarlos o terminarlos.

La LGDS establece que la planeación del desarrollo social incluirá además de los programas federales a los programas municipales, planes y programas estatales, regionales e institucionales; designa a la Comisión Nacional de Desarrollo Social como el instrumento de coordinación de los programas federales en los 3 niveles de gobierno y a la Comisión Intersecretarial de Desarrollo Social como el instrumento de coordinación de las acciones del Ejecutivo Federal con las Secretarías.

CONEVAL: CARACTERÍSTICAS Y FUNCIONES

El CONEVAL fue creado por decreto por parte del Ejecutivo Federal en el año 2005 y en la LGDS, artículo 81, se establece que el CONEVAL es un organismo público descentralizado, con personalidad jurídica, patrimonio propio, autonomía técnica y de gestión de conformidad con la Ley Federal de las Entidades Paraestatales y que las competencias generales que tiene a su cargo son:

- a. Normar y coordinar la evaluación de las Políticas y Programas de Desarrollo Social, que ejecuten las dependencias públicas; y
- b. Establecer los lineamientos y criterios para la definición, identificación y medición de la pobreza, garantizando la transparencia, objetividad y rigor técnico en dicha actividad.

Las atribuciones específicas del CONEVAL se mencionan en el Decreto de creación del mismo y entre las más importantes figuran: revisar el cumplimiento de los objetivos de los programas sociales; establecer criterios y lineamientos para las metodologías de evaluación sobre la política y los programas de desarrollo social; definir los requisitos que deberán cumplir los organismos evaluadores independientes; establecer

lineamientos para realizar estudios y diagnósticos relacionados con la conveniencia, viabilidad y eficiencia de todo programa nuevo de evaluación del desarrollo social; dar a conocer los resultados de las evaluaciones; emitir sugerencias y recomendaciones sobre la política y los programas de desarrollo social; impulsar y fortalecer la cultura de la evaluación; actuar como órgano de consulta y asesoría en materia de evaluación de programas sociales y medición de pobreza, de las dependencias y entidades de la Administración Pública Federal, de las autoridades estatales y municipales; establecer los criterios y lineamientos para la definición, identificación y medición de la pobreza; entre otras¹.

En marzo del 2007 el CONEVAL, la Secretaría de Hacienda y Crédito Público (SHCP) y la Secretaría de la Función Pública (SFP) publican los Lineamientos Generales para la Evaluación de los Programas Federales los cuales regulan la evaluación de los programas federales, la elaboración de la matriz de indicadores de cada programa federal, los sistemas de monitoreo, así como la elaboración de los objetivos estratégicos de las dependencias y entidades de la Administración Pública Federal los cuales deberán estar alineados con el Plan Nacional de Desarrollo y los programas que se deriven del mismo. Los Lineamientos introducen los criterios técnicos para la evaluación de programas federales y estandarizan los procedimientos para mejorar la formulación de los indicadores de desempeño (Castro et al, 2009).

En los Lineamientos se prevé que el Programa Anual de Evaluación para cada ejercicio fiscal será establecido conjuntamente por la SHCP, SFP y el CONEVAL y tiene como propósito instaurar el calendario de ejecución de las evaluaciones y los tipos de evaluación de los programas federales en operación y de los pro-

gramas que vayan a comenzar, vincular el calendario de evaluaciones con el calendario de programación y elaboración del presupuesto. Solamente los programas federales sujetos a reglas de operación aprobados en el presupuesto de Egresos de la Federación serán sujetos de evaluación y otros que en su caso determinen de manera conjunta el CONEVAL, SHCP y SFP.

En cuanto a los resultados y seguimiento de las recomendaciones derivadas de las evaluaciones los Lineamientos establecen que “las dependencias y entidades deberán dar a conocer de forma permanente a través de sus respectivas páginas de Internet, los documentos y resultados de todas las evaluaciones externa” y, “deberán dar seguimiento a los aspectos susceptibles de mejora de los programas federales derivados de las evaluaciones realizadas, conforme al convenio de compromisos de mejoramiento de la gestión para resultados que celebren”.

Los principales logros que los Lineamientos han permitido obtener en el sector público son: a) medidas específicas para la aplicación de los marcos lógicos de los programas presupuestarios, b) adopción de un lenguaje de monitoreo y evaluación en la esfera federal y c) provisión de metodologías estandarizadas de monitoreo y evaluación (Castro et. al 2009).

En cuanto a los desafíos para maximizar la utilidad de los Lineamientos en un proyecto del Banco Mundial México: Análisis de los Lineamientos generales para la evaluación de los Programas Federales de la Administración Pública Federal² se señalan: a) fortalecimiento y coordinación de las funciones y responsabilidades de las instituciones involucradas en la implementación de los Lineamientos, b) mejorar el vínculo entre el Plan Nacional de Desarrollo y los objetivos de los Sectores y Programas, c) mayor uso de los resultados de las evaluaciones, y d) fortaleci-

1 Ver todas las atribuciones específicas del CONEVAL en el Decreto por el cual se regula <http://www.coneval.gob.mx/contenido/quienes/1814.pdf>

2 <http://www.coneval.gob.mx/contenido/quienes/2937.pdf>

miento de la estrategia de información confiable y de calidad para resultados.

En octubre del 2008 la SHCP, SFP y el CONEVAL emiten el Mecanismo para el cumplimiento a los aspectos susceptibles de mejora derivado de los informes y evaluaciones externas a los programas federales, el cual tiene por objeto establecer el procedimiento general para dar seguimiento a los aspectos susceptibles de mejora (ASM) derivados de las evaluaciones del 2007, definir responsables y articular los resultados de las diferentes evaluaciones (consistencia y resultados, diseño y desempeño) para optimizar el impacto de los programas.

En la siguiente sección se presentan los tipos de evaluación que contempla el CONEVAL y en qué consiste cada uno.

TIPOS DE EVALUACIÓN DE PROGRAMAS SOCIALES

Los tipos de evaluación que se contemplan en los Lineamientos son evaluación de consistencia y resultados, evaluación de indicadores, evaluación de procesos, evaluación de impacto y evaluaciones específicas y se señala que para realizar la evaluación de consistencia y resultados se deberán considerar los términos de referencia que determinen de manera conjunta la SHCP, SFP y el CONEVAL, y respecto al resto de las evaluaciones los términos de referencia serán elaborados conforme a las características particulares de cada evaluación.

A. EVALUACIÓN DE CONSISTENCIA Y RESULTADOS

La evaluación de consistencia y resultados busca

analizar el diseño y desempeño de los programas federales para mejorar su gestión y medir sus resultados. En la página del CONEVAL está publicado el Modelo de Términos de Referencia para la Evaluación de Consistencia y Resultados, el cual proporciona una guía sobre el contenido de la evaluación.

Esta evaluación se divide en seis temas: diseño, planeación estratégica, cobertura y focalización, operación, percepción de la población objetivo y resultados. En total esta evaluación contiene 100 preguntas, cada tema contiene preguntas específicas de tipo binario (sí, no), las respuestas se deben sustentar con evidencia documental y/o de análisis y justificar por qué se respondió de tal o cual manera. Para cada uno de los seis temas el evaluador deberá identificar las fortalezas y oportunidades, debilidades y amenazas e incluir evidencia que sustente el análisis. Además para cada debilidad y amenaza se debe proponer una recomendación que ayude a solucionar el reto identificado. Finalmente se presenta una conclusión por tema y una conclusión global del programa.

B. EVALUACIÓN DE DISEÑO

El propósito de la evaluación es analizar el diseño del programa con el fin que las recomendaciones del análisis retroalimenten la gestión del programa. Se divide en seis temas: características del programa, contribución a los objetivos estratégicos, matriz de indicadores, población potencial y objetivo, vinculación con las reglas de operación, y coincidencias, complementariedades y duplicidades. Consta de 34 preguntas tipo binario (sí y no) y son exactamente las mismas preguntas que el tema de diseño de la evaluación de consistencia y resultados. También se realiza un análisis de fortalezas, oportunidades, debilidades y amenazas, y de igual manera para cada debilidad o amenaza se debe proponer una recomendación que ayude a enfrentar el reto identificado.

C. EVALUACIÓN DE INDICADORES

Consiste en analizar mediante trabajo de campo la pertinencia y alcance de los indicadores de un programa federal para el logro de los resultados. Para este tipo de evaluación los términos de referencia serán elaborados conforme a las características particulares de la evaluación.

D. EVALUACIÓN DE PROCESO

Consiste en analizar mediante trabajo de campo si el programa lleva a cabo sus procesos operativos de manera eficaz y eficiente y si contribuye al mejoramiento de la gestión. Para este tipo de evaluación los términos de referencia serán elaborados conforme a las características particulares de la evaluación.

E. EVALUACIÓN DE IMPACTO

En los Lineamientos se establece que la evaluación de impacto consiste en identificar mediante metodologías rigurosas el cambio en los indicadores de resultados atribuibles a la ejecución del programa, en otras palabras permiten medir los efectos netos que tiene un programa en sus beneficiarios con respecto a individuos que tienen características similares pero que no reciben los beneficios del mismo. No hay una metodología o términos de referencia específicos para este tipo de evaluación, lo que se establece en los Lineamientos es que éstos deberán ser aprobados SHCP, SFP y el CONEVAL. No todos los programas pueden llevar a cabo evaluaciones de impacto debido al grado de dificultad y las condiciones necesarias que debe cumplir un programa, por lo cual es necesario desarrollar un análisis de factibilidad de la evaluación de impacto y enviarla a CONEVAL. A partir

de la revisión del análisis de factibilidad el CONEVAL, la SHCP y SFP determinarán si es posible desarrollar la evaluación.

F. EVALUACIÓN ESPECÍFICA

Así se llama a las evaluaciones no comprendidas en los Lineamientos y que se realizan mediante trabajo de gabinete o de campo. El CONEVAL por ejemplo coordinó la Evaluación Específica de Desempeño (EED) 2008-2009 a 131 programas presupuestario. La Evaluación Específica de Desempeño muestra el avance en el cumplimiento de los objetivos y metas de los programas, a partir de una síntesis de la información contenida en el Sistema de Evaluación de Desempeño (SED) y mediante el análisis de indicadores estratégicos y de gestión que permiten conocer los resultados de la aplicación de los recursos públicos presupuestarios. Esta evaluación se aplicará de manera anual.

G. EVALUACIÓN ESTRATÉGICA

Se aplican a un programa o conjunto de programas en torno a las estrategias, políticas e instituciones. Para este tipo de evaluación los términos de referencia serán elaborados conforme a las características particulares de cada evaluación.

H. EVALUACIÓN DE NUEVOS PROGRAMAS

Los Lineamientos establecen que las dependencias y entidades deberán elaborar por un lado un diagnóstico que justifique la creación de nuevos programas que se pretendan incluir en el proyecto de pre-

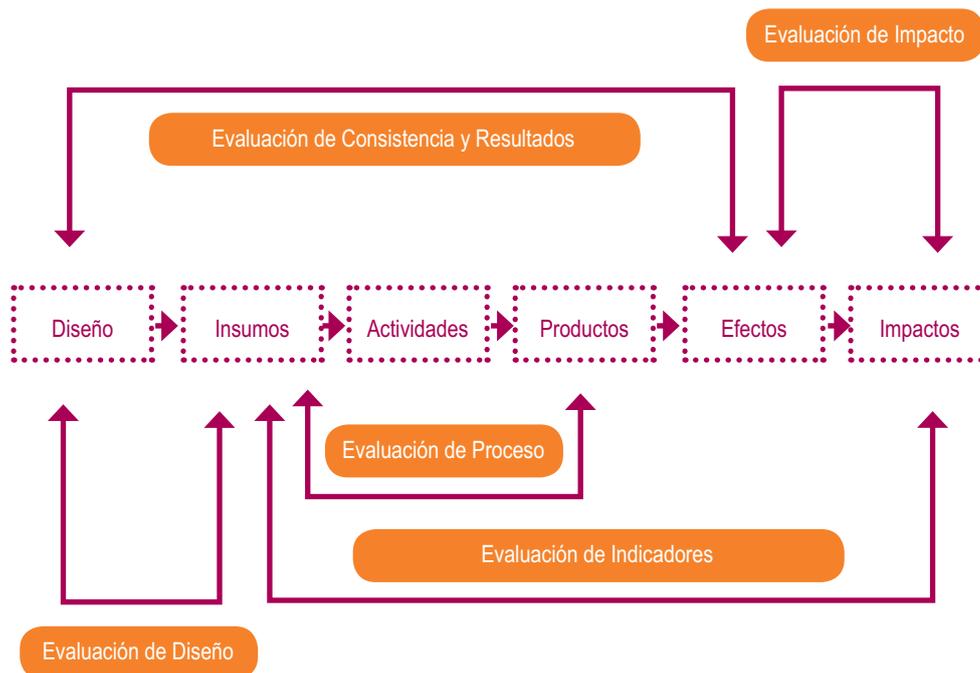
supuesto anual o bien si se quiere ampliar o modificar el programa existente, y por otro lado la matriz de indicadores del programa propuesto. El diagnóstico debe especificar de qué manera el programa que se propone contribuye a los objetivos estratégicos de la dependencia o entidad. El diagnóstico y la matriz de indicadores deberán entregarse a la SHCP y el CONEVAL para la revisión.

En el primer año de operación de los programas nuevos se realizará una evaluación de diseño, así como un análisis de funcionamiento y operación de los mecanismos de selección de beneficiarios, entrega o distribución del apoyo, registro de operaciones pro-

gramáticas o presupuestarias, rendición de cuentas, transparencia y difusión de información. Los resultados de la evaluación y el análisis permitirán mejorar el diseño y la operación de los programas nuevos en su siguiente ejercicio. Las dependencias y entidades deberán elaborar un proyecto de evaluación para sus programas nuevos y presentarlos para que los aprueben la SHCP, SFP y el CONEVAL, y una vez aprobado el proyecto incluirlo en el Programa Anual de Evaluación.

En el siguiente diagrama se presenta la lógica de la articulación de las evaluaciones de los programas sociales que contempla el CONEVAL.

Lógica de la articulación de las evaluaciones de programas sociales que contempla el CONEVAL



Fuente: Elaboración propia con base en el artículo Evaluation Capacity Development. Mexico's M&E System: Scaling up from the sectoral to the national level (Castro et al. 2009)

CONCLUSIONES

El reto ahora en México es generar por un lado una cultura de evaluación de programas sociales sistematizada en el seno de la administración pública, administración que por mucho tiempo obró en la oscuridad sin rendir cuentas ni dar explicaciones sobre cómo y por qué gastaba los recursos públicos de una manera y no de otra, acostumbrada a no responder preguntas, lejos de los ojos críticos de la sociedad, sociedad que ahora es más participativa y exigente, que reacciona ante las decisiones del gobierno, que tiene voz y poder en la definición del futuro compartido y más

que importarle quién la gobierna, le importa quién la gobierna de la mejor manera.

Afortunadamente los primeros pasos en el tema de la evaluación de programas sociales en el país ya se dieron, actualmente se llevan a cabo diferentes tipos de evaluaciones, sin embargo queda la duda hasta que punto la información que se desprende de dichas evaluaciones es efectivamente utilizada por las autoridades correspondientes para retroalimentar el funcionamiento de los programas y perfeccionar los puntos débiles del mismo.

REFERENCIAS BIBLIOGRÁFICAS

Castro, M., G. López-Acevedo, G. Beker, X. Fernandez. (2009) Evaluation Capacity Development. Mexico's M&E System: Scaling up from the sectoral to the national level. Independent Evaluation Group. ECD Working Paper Series No. 20.

Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) www.coneval.gob.mx.

Consejo Nacional de Evaluación de la Política de Desarrollo Social (2008). Informe de evaluación de la política de desarrollo social en México, 2008.

<http://www.coneval.gob.mx/contenido/home/2234.pdf>.

Decreto por el que se regula el Consejo Nacional de Evaluación de la Política de Desarrollo Social. <http://www.coneval.gob.mx/contenido/quienes/1814.pdf>.

Estrategia Vivir Mejor http://www.presidencia.gob.mx/vivirmejor/vivir_mejor.pdf.

Fuentes, G. (2009). Reglas de operación de los programas del Gobierno Federal: Una revisión de su justificación y su diseño. Centro de Estudios Sociales y de Opinión Pública de la Cámara de Diputados, LX Legislatura. Documento de Trabajo núm. 71.

Hernández, G (2006). M&E de programas sociales en México. Hacia la Institucionalización de los Sistemas de Monitoreo y Evaluación en América Latina y El Caribe: Actas de una Conferencia del Banco Mundial / Banco Interamericano de Desarrollo. Editores: Ernesto May, David Shand, Keith Mackay, Fernando Rojas, Jaime Saavedra.

Lineamientos Generales para la Evaluación de los Programas Federales de la Administración Pública Federal en México.

http://www.coneval.gob.mx/coneval2/htmls/evaluacion_monitoreo/HomeEvalMonitoreo.jsp?id=tr_eval_consist_result.

Modelo de términos de referencia para la evaluación de consistencia y resultados http://www.coneval.gob.mx/contenido/eval_mon/906.pdf.

Medina, Alejandro. México: Análisis de los Lineamientos generales para la evaluación de los Programas Federales de la Administración Pública Federal. Proyecto de Referencia: P101567. Líder del proyecto: Gladys López-Acevedo. Unidad de Pobreza y Género Reducción de la Pobreza y Gestión Económica Región de América Latina y el Caribe.

<http://www.coneval.gob.mx/contenido/quienes/2937.pdf>.



ALGUNAS CUESTIONES CONCEPTUALES Y METODOLÓGICAS EN TORNO A LA EVALUACIÓN DE IMPACTO

José María Larrú
Universidad CEU San Pablo

INTRODUCCIÓN

La evaluación de impacto está adquiriendo cada vez más relevancia en el ámbito académico y en la toma de decisiones de políticas públicas. Algunos autores (Leonhardt 2008) han llegado a escribir que la economía del desarrollo ha salido de su ostracismo e irrelevancia gracias al comienzo de la elaboración de evaluaciones de impacto aleatorizadas que, bajo el método experimental, permiten evidencias fuertes sobre lo que funciona y lo que no en las intervenciones de desarrollo (económico y humano). El artículo pretende alcanzar tres objetivos. En primer lugar, aclarar algunos términos que a menudo se usan indistintamente (específicamente resultados, efectos e impactos). En segundo lugar, contribuir a valorar de forma equilibrada las ventajas y limitaciones que tienen las evaluaciones de impacto experimentales o bajo el método de “experimentos aleatorios controlados” (también conocidos por su acrónimo inglés RCTs, de random control trials). La identificación directa de que sólo este método produce conocimiento científico y

por tanto válido, se cuestiona y se pone en relación con la utilidad y enjuiciamiento que toda evaluación debe generar. Por último, se hace una breve presentación del Fondo Español para Evaluaciones de Impacto (SIEF), como apuesta de mayor calado –político y presupuestario- que se ha realizado desde el gobierno español, haciendo notar que su mayor potencial –recibir lecciones aprendidas y articular la información disponible de forma amable y sintética para la toma de decisiones- es una dimensión del SIEF que aún está por implementar. Estos tres objetivos, coinciden con la estructura del resto del artículo, que termina recogiendo las principales conclusiones.

ACLARANDO ALGUNOS TÉRMINOS

En esta sección me propongo aclarar el sentido de algunos términos referidos a la evaluación de impacto que generan confusión. Las fuentes de confusión son dos: por una parte están las debidas a las traduccio-

nes de los términos ingleses; por otra parte, diversos organismos internacionales han utilizado términos iguales para hacer referencia a conceptos distintos.

EL CONCEPTO DE IMPACTO

Creo que el punto de partida debe ser el Glosario del CAD (CAD 2002). En él se define “impacto” como “los efectos de largo plazo positivos y negativos, primarios y secundarios, producidos directa o indirectamente por una intervención para el desarrollo, intencionalmente o no”. Es decir, se adopta un enfoque extenso y holístico. El contexto será altamente condicionante para el juicio evaluativo. Este enfoque es el adoptado por redes internacionales como la “Red de Redes en Evaluaciones de Impacto” (Network of Networks on Impact Evaluation, NONIE (2009:ix).

Algunos autores y organizaciones de desarrollo, como la International Initiative for Impact Evaluation (3ie), el Abdul Latif Jameel Poverty Action Lab (J-PAL) asociado al MIT o Innovations for Poverty, asociada a la Universidad de Yale, “practican” evaluaciones de impacto entendiendo por ellas las que son realizadas bajo diseños experimentales (los que utilizan un grupo de población sobre la que se interviene y otra sobre la que no, a modo de control). Es decir, existe una definición más estrecha de evaluación de impacto. Aquella que en la que evaluar el impacto es responder a la pregunta del contrafactual: cuál sería la situación de la población intervenida de no haber intervenido. Es decir, buscan acercarse al conocimiento causal de la forma más precisa posible. Para algunos esta precisión, será la “científica”, al poder atribuir los efectos netos de una intervención con un máximo grado de validez interna. Es decir, si logramos eliminar los sesgos que perturban la medición del efecto producido por la intervención (especialmente el sesgo de selección, es decir, el producido por la autoselección de beneficiarios o, de forma más general, cuando el efecto de una intervención está correlacionado con el

hecho mismo de intervenir) tendremos una medición insesgada, válida dentro del sistema de intervención. Si eso es “ciencia”, conocimiento insesgado de la realidad medible, las evaluaciones bajo diseños experimentales serán “científicas” y las otras no.

Recordemos que la evaluación no es una disciplina que nazca con el deseo de “hacer ciencia”, sino con la de generar información útil para aprender, controlar procesos y poder tomar decisiones con mejor información. La evaluación pretende levantar la información más ética, útil y rigurosa posible. Por tanto pretende realizar juicios válidos (lógicamente) y verdaderos (epistemológicamente), pero no se va a “obsesionar” por si eso es el fin de la ciencia o no. Las políticas se orientan por muchos valores (desde el “interés general” a “ganar unas elecciones”) entre los que no creo que se encuentre la “cientificidad”.

En resumen, la concepción “canónica” (por más antigua) del impacto establecida por el CAD como uno más de los criterios de evaluación, tiene un alcance extenso y holístico. El contexto es un factor determinante del análisis.

En la concepción “estrecha” más moderna el impacto, toma algunas connotaciones diferenciales: se ciñe preferentemente a los diseños experimentales y, en segundo rango, a los cuasi-experimentales; se centra en la medición de los efectos netos de la intervención; toma en serio los sesgos de selección y aborda de forma prioritaria la atribución, por ello utiliza los grupos de tratamiento y control como opción metodológica más perfecta; el concepto de impacto será asociado al de respuesta al contrafactual.

Debemos aclarar entonces que, para la concepción “estrecha” de impacto, responder al contrafactual es el fin primario de una evaluación. Pero utilizar grupos de control no es la única manera de obtener esta respuesta. Se pueden utilizar otros métodos como el análisis de experimentos naturales o la simulación

computacional si disponemos de series históricas de datos. Dicho esto, es cierto que si pretendemos evaluar el impacto de una intervención de desarrollo que se va a desempeñar en el presente y durante unos años en adelante, una forma altamente eficiente de emprender el diseño de la evaluación es adoptando un enfoque experimental que permita recoger información cuantitativa y cualitativa tanto de un grupo sobre el que se interviene como sobre otro “idéntico” en promedio excepto en el “tratamiento” de la intervención. Si el tamaño muestral es lo suficientemente grande (y eso lo determina el tipo de intervención y el número de variables independientes), podré obtener una medición insesgada del efecto promedio debido exclusivamente a la intervención y no a otros posibles factores. En otras palabras, ahora el contexto es que el evaluador pretende “aislar” lo más posible. Siendo poco riguroso, pero con intención de ser gráfico, la evaluación de impacto “estrecha” quiere “un número sin contexto”, mientras que la definición “clásica” del CAD hace del contexto precisamente su objeto de análisis en profundidad. No le importa obtener “un juicio de los hechos en su contexto, sin número”.

Un detalle importante es que para el sentido “amplio” de la evaluación de impacto, es posible diseñar la intervención únicamente ex –post. Pero el sentido estrecho de evaluación de impacto, requiere –en la mayoría de los casos- un sistema de información que empiece a funcionar desde el comienzo mismo de la intervención.

Creo que no tiene sentido preguntarse qué definición es mejor. Simplemente son dos abordajes diferentes. Para algunas evaluaciones, la información requerida de la evaluación de impacto deberá ser cuantitativa, “exacta” y con atribución (por ejemplo, cuando la pregunta por la eficacia de una intervención –¿funciona este subsidio o no?- y su eficiencia respecto a otras alternativas). El contexto deberá “aislarse” lo más posible. Por el contrario, en otras evaluaciones lo que

interese sea precisamente el contexto: Por ejemplo, si el promotor de una evaluación pregunta ¿en qué circunstancias y con qué características poblaciones esta intervención ha mostrado ser eficaz para este grupo concreto? En este caso, el acento queda puesto en la heterogeneidad de los efectos. El diseño experimental devuelve un valor promedio de la intervención insesgado. Pero no puede responder a la heterogeneidad de los efectos, a no ser que se complete con un análisis cualitativo o con un enfoque de evaluación basada en la teoría.

Por último, no hay que olvidar que en muchas de las intervenciones (y quizá en la mayoría de las políticas públicas, no así en sus programas o proyectos) el diseño experimental simplemente no es posible, con lo que el sentido del “impacto” en la evaluación ha de ser necesariamente el clásico del CAD.

IMPACTO Y RESULTADOS

Otra confusión frecuente es la referida a evaluación de resultados. Primero hay error de traducción. A menudo no se cae en la cuenta que los documentos oficiales sobre la eficacia de la ayuda (por ejemplo la Declaración de París y la Agenda para la Acción de Accra) se expresan en términos de orientación hacia resultados y no “de” resultados.

Pero también es cierto que “resultados” es un término que ha sido traducido tanto para outputs (quizá porque productos suene muy “economicista”) como para outcomes. Creo que la principal fuente de confusión (al menos en España) ha sido que el lenguaje del Marco Lógico ha acuñado “resultados” como el fruto de las actividades programadas, de forma que los “resultados” formulados y sobre los que se da cuenta en el informe de seguimiento y final, han sido siempre el grado de cumplimiento de las actividades previstas.

Mi opción es traducir los outcomes como los efectos o cambios en la vida de las personas (con independencia de la atribución a la intervención o no). El Glosario del CAD que dedica una sección a la gestión “basada en” resultados, define el término como “Producto, efecto o impacto (intencional o no, positivo y/o negativo) de una intervención para el desarrollo. Términos conexos: efecto directo, repercusión, impacto” y “Cadena de resultados” como “la secuencia causal de una intervención para el desarrollo que estipula la secuencia necesaria para lograr los objetivos deseados, comenzando con los insumos, pasando por las actividades y los productos, y culminando en el efecto directo, el impacto y la retroalimentación. En

algunos organismos, los destinatarios (o el « alcance ») forman parte de la cadena de resultados. Términos conexos: supuestos, marco de resultados” (CAD 2002: 33).

Queda aquí claramente mostrada como la confusión está más que justificada.

La oficina comunitaria de cooperación para el desarrollo, EuropeAid, aborda esta cuestión en sus metodologías de evaluación. Para esta agencia, existe un matiz temporal entre resultados (efectos iniciales o de corto plazo) y los impactos que se producen a más largo plazo.

Figura 1. Significados polisémicos de impacto y resultados entre la OCDE y la Comisión Europea.

Riesgo de malos entendidos		
	OCDE	CE
Impacto	... efectos a largo plazo producidos por una intervención para el desarrollo...	Un término general utilizado para descubrir los efectos de una intervención en la sociedad...
Outcome (traducido como efecto directo)	Los efectos, a corto o a medio plazo probables o logrados por los productos de una intervención.	Los impactos a más largo plazo, por lo general expresados como consecuencias socio económicas en el sentido amplio de la palabra.
Resultado	... productos, efecto o impacto (intencional o no, positivo y/o negativo)	El impacto inicial de una intervención...

Fuente: Glosario de evaluación de EuropeAid: http://ec.europa.eu/europeaid/evaluation/methodology/glossary/glo_es.htm

La polisemia se agudiza cuando se comparan el significado de producto, resultado e impacto entre la guía

de evaluación de EuropeAid, el glosario del CAD y las directrices del Presupuesto comunitario. Como refleja

Figura 2. Diversidad de acepciones de producto, resultado e impacto.

Lógica de intervención: comparación de los conceptos clave							
Nivel	Esta guía		OCDE CAD		CE DG Presupuesto		
1	Producto	Efecto	Producto	Resultado	Producto		
2	Resultado		Outcome		Resultado	Resultado	Impacto
3	Impacto		Impacto		Outcome		
1	Bienes y/o servicios producidos/entregados por la intervención						
2	Cambio inicial atribuible a la intervención pública						
3	Cambios a más largo plazo atribuibles a la intervención pública						

Fuente: Glosario de evaluación de EuropeAid: http://ec.europa.eu/europeaid/evaluation/methodology/glossary/glo_es.htm

la Figura 2, en el ámbito comunitario se mantiene la asociación de resultados al corto plazo (nivel 2 de la comparación) aunque incluye la atribución del efecto, cosa que no ocurre con la acepción del CAD. Nuestra opción ha sido la de reservar la atribución para el impacto ni con nuestra definición “estrecha” de impacto.

En conclusión y de aquí en adelante, el sentido de los términos que emplearemos es el siguiente:

productos (outputs) serán los indicadores que informen del grado de cumplimiento de las actividades previstas en la formulación del Marco Lógico de un proyecto o programa;

resultados o efectos (outcomes) serán los cambios que las intervenciones han producido en la vida de las personas, con independencia de que puedan atribuirse de forma exclusiva o causal a la intervención; así una evaluación de resultados indagará en los cambios habidos en la vida de las personas utilizando testimonios recogidos en entrevistas o indicadores cuantitativos observables (mejoras en la infraestructura

del hogar, grado de morbilidad de los miembros de una familia; desnutrición observada a través de medidas antropométricas, etc.); esta evaluación informará de la contribución de la intervención a dicho cambio pero suspenderá el juicio sobre la atribución causal;

impacto será la respuesta a un contrafactual con capacidad para atribuir el cambio medido a la intervención realizada.

La figura 3 expresa dos vías para encadenar las fases de una intervención. La primera recoge el sistema de gestión predominante hasta ahora en la Cooperación Española en el que el protagonismo recae sobre el informe final. Éste debe dar cuenta de los productos obtenidos y justificar de forma precisa el gasto incurrido para obtener dichos productos. Pero apenas informa sobre la sostenibilidad futura de estos productos ni de su utilidad real en la vida de las personas. Pongamos un ejemplo ilustrativo. El informe final de un programa de formación de microempresarios revelará el número de capacitaciones realizadas sobre un número determinado de “beneficiarios”. Pero desconocemos

si esta formación ha servido para las microempresas estrenadas fruto de la formación sobrevivan dos o tres años y si los receptores de las capacitaciones mejoraron su situación personal en cuanto ingresos o calidad de vida. Es decir, podemos estar apoyando procesos que no dan fruto en términos de desarrollo.

Siguen siendo pobres. Capacitados, pero pobres. Por no ahondar en que han podido existir efectos perjudiciales como un aumento de la carga familiar o personal de trabajo, menos tiempo libre, menor calidad de vida por mayor estrés y preocupación por sacar adelante un negocio que no es rentable.

Figura 3. Encadenamiento de productos y de resultados.



Fuente: elaboración propia.

LA PREGUNTA PRECEDE AL MÉTODO

Lo primero que debemos preguntarnos cuando diseñamos una evaluación de impacto, no es qué herramientas debemos usar o qué diseño metodológico emplear, si es el experimental o cuasi-experimental. La supuesta “jerarquía de métodos”, si es que existe, no es más una clasificación funcional bajo criterios de “cientificidad”. Pero el valor de la ciencia y sus criterios de verdad no tienen por qué coincidir con los de una evaluación. La evaluación busca un juicio práctico que ayude a responder a una (o varias) pregunta concreta. La utilidad es básica y ésta queda marcada por el contexto socio-político de cada momento. La ciencia, busca responder de forma lo más universal posible a la verdad, entendida como adecuación a la reali-

dad, con independencia (hasta donde sea posible) del contexto.

En la discusión que tuvo lugar en El Cairo en un seminario organizado por NONIE (29-marzo-2 abril, Perspectives on Impact Evaluation) varios expertos fueron inquiridos a diseñar tres intervenciones muy diversas. La mayoría de ellos comenzó su intervención apuntando que el diseño que se iba a comentar a continuación era secundario. Lo primero era preguntarse por qué se promueve esta evaluación, quién se beneficiará de ella, y cómo será utilizado y entendido el “impacto” (Chambers 2009). Ravallion (2009) sintetizó las preguntas clave previas al diseño concreto de una evaluación de impacto en las siguientes:

- ▶ ¿Por qué se ha realizado esta intervención?
- ▶ ¿Cuál es el conocimiento actual sobre la cuestión y cuáles son las incógnitas más importantes (knowledge gaps)?
- ▶ ¿Qué contrafactual es el que resulta ahora más idóneo? ¿Cuál es realmente el que le interesa al hacedor de políticas?
- ▶ ¿Cuáles son los efectos deseados, sobre qué periodo y cuáles los potencialmente indeseados?
- ▶ ¿Cuáles son los parámetros más relevantes a estimar (los intencionalmente tratados o los realmente tratados)?
- ▶ ¿Cuáles son los mecanismos de transmisión de efectos más probables?
- ▶ ¿Cuáles son las opciones metodológicas más adecuadas, dadas las restricciones existentes de recursos, tiempo e información?

Como vemos son preguntas que deben preceder a toda investigación. Pero refuerzan la idea anterior y el título de esta sección. Las metodologías (y sus correspondientes herramientas cuantitativas y cualitativas) son posteriores a las respuestas estratégicas (más políticas) de la evaluación. Es decir, el diseño debe responder a la utilidad que se le va a dar a la evaluación. Si ante una intervención lo principal que se desea conocer es dónde poder gastar recursos limitados (eficiencia), el método será diferente a si lo que predomina es cómo aumentar la escala de una programa, o detallar los efectos indeseados (directos o secundarios) que provoca para minimizarlos. Lo más probable (siempre el contexto puede matizar esta valoración) es que para responder al primera pregunta lo primero que debiéramos considerar es un diseño experimental (o experimento aleatorio con control, Randomized Control Trial). Pero el RCT no es

la mejor opción para predecir si los efectos se mantendrán al ampliar la escala de una intervención concreta. En este caso y en el de los efectos secundarios y heterogéneos en la población meta, los diseños con abundancia de herramientas cualitativas y participativas pueden ser más útiles y necesarios.

Una falacia que conviene advertir es adjudicar “linealmente” las categorías de: experimental-cuantitativo-riguroso-científico-válido-verdadero frente a lo no experimental-cualitativo-flexible-opinático-inválidomenos creíble.

Karlan (2009) y White (2009), entre otros, han señalado que el diseño experimental bajo los RCT no son la única forma de evaluar (nadie lo sostiene a mi modo de ver) y no se nutre sólo de cálculos sin ninguna teoría o ninguna evidencia cualitativa. Por otra parte, los diseños participativos son también muy capaces de generar datos y cálculos estadísticos que midan el impacto (Mayoux y Chambers 2005; Chambers 2007).

Rogers (2008, 2009) diferencia entre evaluaciones de intervenciones sencillas, complicadas y complejas en función de la implementación de la intervención. Cada una de ellas elabora una forma de causalidad diferente, y complementaria si es el caso. La sencilla se caracteriza porque la lleva a cabo una sola organización. El impacto se centra en “la causa” más probable que ha producido los efectos. Responde predominantemente a la pregunta de qué funciona.

La intervención complicada está formada por múltiples componentes (objetivos específicos en terminología del EML). Varias organizaciones trabajan de forma contractual bajo responsabilidades bien definidas. El análisis causal es múltiple, secuencial y/o simultáneo y en diferentes niveles. Podemos conocer qué funciona, para quién y bajo qué circunstancias.

La intervención compleja es aquella en la que inter-

vienen varias organizaciones relacionándose en un marco informal de asociación y cooperación (o competencia). Pensemos en la multitud de organizaciones que trabajaron en la reconstrucción de las zonas afectadas por el tsunami de 2005. En este caso, las rela-

ciones causales pueden ser recurrentes y “emergen” efectos dinámicos (es decir, el todo es mayor la suma de sus partes). La respuesta buscada es qué es lo que está funcionando en este contexto y circunstancias determinadas.

Figura 4. Evaluación de intervenciones simples, complicadas y complejas.

Intervención	Fines de la evaluación	Cadena causal	Ejemplo
SIMPLE (un OE y un agente involucrado lidera la ejecución de la intervención).	Saber si funciona y en qué contexto. Mejorar un programa o política ya en marcha.	A -> X No A ->no X	Transferencias condicionadas
COMPLICADA (varios OEs y varios agentes involucrados repartidas las responsabilidades de forma clara y contractual).	Valorar los impactos de un programa (múltiples componentes y agentes involucrados).	A, B, C -> X A->B B->C pero no necesariamente A->C; A^B -> X B^noC->X etc.	Construcción de infraestructuras
COMPLEJA (múltiples objetivos y acciones dinámicas; efectos dinámicos e intervenciones emergentes respondiendo a diversas necesidades y oportunidades).	Entender los procesos y cadenas causales; mejorar programas en marcha y adoptar ajustes dinámicos y contextuales. Varios efectos sobre múltiples causas producidos por varios agentes que interactúan y modifican los efectos.		Comisión anti-corrupción

Fuente: Elaboración propia, basada en Rogers (2009).

En resumen, la pregunta principal en la evaluación de impacto es por qué se promueve esta evaluación, quién la usará y qué tipo de evidencia necesita para dar respuesta a su iniciativa. Aclaradas estas cuestiones, se determinará a quién se debe involucrar (para Chambers es innegociable que los afectados-usuarios-beneficiarios estén directamente involucrados y participando) y por qué método optar. En tercer lugar se selecciona la combinación de herramientas cuantitativas que sea más oportuna para cada caso.

No debemos olvidar que estas respuestas no deberían obtenerse únicamente de forma ex -post, o una vez terminada o ya iniciada la intervención. La recogida de información es clave y esta debe ser levantada antes, durante y después de la intervención. En los RCT, es donde quizá esto queda más evidenciado dada la gran cantidad de datos cuantitativos a recoger y que el tamaño muestral queda determinado por el poder muestral (y margen de error) que pretendemos alcanzar.

LA APORTACIÓN ESPAÑOLA: EL SIEF

Hasta la fecha, la Cooperación Española no ha promovido ni realizado de forma directa ninguna evaluación de impacto en el sentido estricto que aquí venimos utilizando.

Sin embargo, en 2007 decidió crear el Fondo Español de Evaluaciones de Impacto (SIEF en acrónimo inglés Spanish-World Bank Trust Fund for Impact Evaluation)¹.

Es una iniciativa que se enmarca dentro del Grupo de Desarrollo Humano del Banco Mundial y que está dotada con 10,4 millones de dólares para el periodo 2007-2010. España ha establecido, en diálogo con el Grupo de Desarrollo del Banco Mundial, un listado de 72 países prioritarios donde realizar evaluaciones de impacto, la mayoría bajo metodología aleatorizada o experimental.

La organización y gestión del fondo queda sintetizada en la Figura 5.

Figura 5. Organigrama del SIEF.



Fuente: elaboración propia a partir de la web del SIEF.

¹ Véase <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:21407362~menuPK:384336~pagePK:148956~piPK:216618~theSitePK:384329,00.html>

Como refleja el gráfico, el fondo está organizado en tres niveles. En el nivel directivo, las decisiones de máximo nivel se toman de forma consensuada entre el Vicepresidente de la Red de Desarrollo Humano del Banco Mundial y –por parte española- la Secretaría de Estado de Cooperación Internacional. La gestión diaria queda a cargo de la Oficina del Economista Jefe de la Red de Desarrollo Humano y su equipo de apoyo. En el nivel técnico, encargado de la selección y aprobación de los fondos destinados a cada evaluación, ayudan a la economista jefe un cuerpo de seis expertos.

El SIEF tiene encomendada tres finalidades: realizar evaluaciones de impacto, realizar capacitaciones sobre las metodologías de evaluación de impacto y difundir los aprendizajes surgidos de las evaluaciones realizadas. A continuación describimos brevemente cada una de las tres.

Para realizar evaluaciones de impacto, se han establecido 72 prioridades geográficas y once sectoriales o temáticas. Las prioridades geográficas se distribuyen entre 29 países sudafricanos, 20 latinoamericanos, 8 del norte de África y Oriente Medio, 6 del Sur de Asia, 6 del Este asiático y 3 de Europa Central y Oriental. Todos los países prioritarios de la Cooperación Española están dentro de esta extensa lista, elaborada bajo criterios ad hoc². Por su parte, las once prioridades temáticas son las siguientes:

1. Soberanía alimentaria y lucha contra el hambre
2. Educación
3. Protección Social y laboral
4. Salud
5. HIV/SIDA, Malaria y otras enfermedades

2 Al menos, en la web no aparece ningún tipo de argumento técnico para esta selección, con lo que puede deducirse que ha sido la discrecionalidad política quien ha fijado éstas prioridades.

6. Igualdad de género
7. Salud Sexual y Reproductiva
8. Niñez y Juventud
9. Discapacitados
10. Agua y Saneamiento
11. Vivienda

Como puede comprobarse, el conglomerado nutrición-salud-educación, que está en la base del concepto de Naciones Unidas de seguridad alimentaria, queda ampliamente recogido pues las prioridades 1,2,4,5,6,7 y 10 tienen cabida en dicho concepto³.

El SIEF ha establecido a su vez tres tipos de fondos o convocatorias.

► Las evaluaciones innovadoras o de “resultados rápidos” (quick wins) que fueron las primeras en aprobarse⁴. Están organizadas en torno a siete temas comunes. Se presentaron 24 propuestas y fueron seleccionadas 13. Cuentan con un presupuesto conjunto de 1,5 millones de dólares.

► Convocatoria de evaluaciones dentro de los “clusters” seleccionados. Mediante concurso, han sido seleccionadas 26 Evaluaciones dentro entre las 126 propuestas, que se han agrupado en torno a 6 clusters temáticos: prevención del SIDA, transferencias

3 La definición global de seguridad alimentaria (o alimentaria y nutricional), oficializada unánimemente por los Jefes de Estado y de Gobierno de los países miembros de la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) durante la Cumbre Mundial de la Alimentación (1996) indica que existe seguridad alimentaria “Cuando todas las personas tienen en todo momento acceso físico, social y económico a los alimentos suficientes, inocuos y nutritivos que satisfagan sus necesidades energéticas diarias y preferencias alimentarias para llevar una vida sana y activa”.

4 Véase el detalle de esta modalidad en http://sitere-sources.worldbank.org/INTISPM/Resourses/383704-1184250322738/SIEF_QuickWins_Program.pdf

condicionadas en efectivo, educación básica, control de la malaria, contrataciones de desempeño en salud y, mercados de trabajo activos y para jóvenes. Está dotada con 5,6 millones de dólares.

► La última en implementarse ha sido el Fondo de Innovación. Dotado con 2,6 millones de dólares, está dividido en dos ventanillas. La primera es de Desarrollo Humano y deben realizarse sobre intervenciones dentro de los 9 sectores seleccionados que inciden de forma estrecha sobre el desarrollo humano. Bajo esta ventanilla se están realizando 5 evaluaciones. La segunda ventanilla, de Desarrollo Social, financia actualmente 8 evaluaciones dentro del ámbito temático del acceso a agua, saneamiento y vivienda⁵.

La segunda misión del SIEF es realizar cursos y sesiones formativas que aumenten la capacidad local para promover y, en su caso, realizar evaluaciones de impacto rigurosas. Desde su creación en 2007 se han llevado a cabo nueve cursos en ciudades de todo el mundo y están programados otros 3 para 2010. En Madrid se celebró el curso entre el 23-28 de junio de 2008 con 184 participantes⁶.

La tercera misión encomendada al SIEF es la de difundir los resultados obtenidos en las distintas evaluaciones de impacto. Para ello están previstas actividades como estudios de política de alta calidad que sinteticen la evidencia generada sobre aquello “que funciona” para alcanzar resultados en el ámbito de desarrollo humano; acceso público a las bases de datos y traducción al español de los resultados de las

5 El detalle de las 12 evaluaciones del Fondo de Innovación puede consultarse en: http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1256762343506/SIEF_Innovation_Fund_Budget_Allocation_17nov09.pdf

6 Los detalles del programa y materiales ofrecidos pueden consultarse en la web del SIEF: <http://web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/ORGANIZATION/EXTHDNETWORK/EXTHDOFFICE/0,,contentMDK:22383415~pagePK:64168445~piPK:64168309~theSitePK:5485727,0.html>

investigaciones. Lamentablemente, hasta la fecha y que conozca, no hay disponible ningún producto que pueda incluirse dentro de esta tercera misión, lo cual es una pena ya que es la forma más directa de cómo la Cooperación Española se beneficia del SIEF. Dado que no hay ningún evaluador español dirigiendo las evaluaciones de impacto en marcha, la apropiación de los conocimientos generados por los demás debería ser la prioridad para rentabilizar la inversión del SIEF. En 2009, se destinó a Washington a un miembro de la Unidad de Evaluaciones de la DGPOLDE, para poder tener un seguimiento más cercano y directo del Fondo. Sería deseable que esta dimensión de producción de buenas y malas prácticas y resultados puedan reflejarse en decisiones estratégicas e intervenciones de desarrollo de probado mayor impacto. Evaluaciones que no conducen a cambios operativos y estratégicos no sirven de mucho, y la utilidad –junto con la credibilidad– son las dos características esenciales de las evaluaciones (y que las diferencian en parte de las investigaciones).

CONCLUSIONES

En este artículo se ha propuesto diferenciar una concepción más amplia y canónica de impacto (la contenida en el Glosario del CAD) de otra más estrecha o estricta que es aquella que responde al contrafactual. También debe diferenciarse resultados (en el sentido frecuente de productos siguiendo la terminología del Enfoque del Marco Lógica habitual en la formulación de las intervenciones de cooperación para el desarrollo) de resultados o efectos (outcomes) que se utiliza en los documentos doctrinales de la eficacia de la ayuda como la Declaración de París o de Accra.

Por otra parte, la respuesta al contrafactual puede realizarse bajo diversas metodologías y diseños evaluativos, no debiendo ceñirse exclusivamente al experimental o aleatorizado. La pregunta evaluativa

precede al método para lograr una respuesta útil y creíble. El método experimental, como cualquier otro diseño, tiene ventajas y limitaciones (véase una discusión más amplia en Larrú 2007, 2009). Una opción, preferente para algunos evaluadores e investigadores, es responder al contrafactual bajo un diseño de RCT. Pero existen otras posibilidades, en función del contexto o de la complejidad de la intervención, como propone Rogers.

Por último, España a través del Fondo de Evaluaciones de Impacto (SIEF) ha hecho una apuesta muy

fuerte por este tipo de evidencias. Creado en 2007 y bien dotado económicamente, se echa de menos una mayor presencia de evaluadores españoles realizando evaluaciones y, sobre todo, que la dimensión de difusión de buenas prácticas a instrumentos de desarrollo exitosos o erróneos tenga una mayor entidad, pues hasta la fecha no se ha producido un ejercicio sistemático en esta materia que enriquezca a la comunidad evaluadora española, ni a los tomadores de decisiones implicados en la política de cooperación y/o de desarrollo internacional.

REFERENCIAS BIBLIOGRÁFICAS

CAD (2002) Glosario de los principales términos sobre evaluación y gestión basada en resultados. Evaluation and Aid Effectiveness N°6. OECD-DAC. París.

CHAMBERS, R. (2007) "Who Counts? The Quiet Revolution of Participation and Numbers", IDS Working Paper 296.

CHAMBERS, R. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

KARLAN, D. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

LARRÚ, J.M. (2007) "La Evaluación de Impacto: ¿qué es, cómo se mide y qué está aportando en la Cooperación al Desarrollo", Colección Escuela Diplomática N°12, 109-133.

LARRÚ, J.M. (2009) La ayuda al desarrollo, ¿reduce la pobreza? Eficacia y evaluación de la cooperación para el desarrollo. Biblioteca Nueva. Madrid.

LEONHARDT, D. (2008) "Making Economics Relevant Again", New York Times, feb, 20.

MAYOUX, L. & CHAMBERS, R. (2005) "Reversing the Paradigm: Quantification, Participatory Methods and Pro-Poor Impact Assessment", Journal of International Development 17 (2), 271-298.

NONIE (2009) Impact Evaluations and Development. NONIE Guidance on Impact Evaluation. Frans Leeuw & Jos Vaessen (authors). The Network of Networks on Impact Evaluation, Washington.

RAVALLION; M. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

ROGERS, P. (2008) "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", Evaluation 14(1), 29-48.

ROGERS, P. (2009) "Designing Impact Evaluations: Different Perspectives", International Initiative for Impact Evaluation Working Paper 4. [también publicados en Journal of development Effectiveness 1(3)].

WHITE, H. (2009) "Some Reflections on Current Debates in Impact Evaluation", 3ie Working Paper 1. [también publicado en Evaluation, Vol. 16, No. 2, 153-164].

EL ESTUDIO DE LA EVALUABILIDAD COMO CONTRIBUCIÓN A EVALUAR CON CALIDAD

Itziar Arispe

Investigadora, Agencia Internacional de la Energía Atómica

RESUMEN

El artículo pretende analizar como los estudios de evaluabilidad son compatibles en la práctica de las evaluaciones actuales. Primero presenta la evolución del marco normativo en evaluación, y el concepto de evaluabilidad técnica y política. Seguidamente se enumeran algunas de las razones por las que los estudios de evaluabilidad no son habituales en las organizaciones internacionales y por último se analizan posibles escenarios para fomentar su práctica.

La evaluación es una disciplina que ha pasado por diferentes fases evolutivas. Hoy en día, podemos afirmar que nos encontramos en un momento donde desde las instituciones públicas y organismos internacionales se valoran las ventajas de la evaluación y se promueve activamente la existencia de un campo normativo y de su profesionalización.

En este sentido, cabe destacar el auge que ha experimentado el desarrollo de códigos normativos en el campo de la evaluación. Bustelo (1998) identifica que la primera función de los códigos deontológicos es que permiten la introducción y desarrollo de temas éticos. La segunda función es la creación de un senti-

do de pertenencia e identificación a una determinada comunidad profesional.

La existencia de códigos de evaluación en los países ha ido intrínsecamente relacionada con un mayor grado de institucionalización y consolidación profesional de la función de evaluación (Ballart 1993). A la cabeza se encuentran países como EEUU, Canadá, Noruega, Australia, Japón y Gran Bretaña, los cuales se caracterizan por ser países con amplia experiencia en políticas sociales educativas, y a su vez, por ser importantes donantes de la cooperación para el desarrollo, tanto a nivel nacional, como a través de su participación en organizaciones supranacionales.

Existen múltiples diferencias entre los códigos de evaluación nacionales. Por ejemplo, Los del Joint Committee on Educational Evaluation, la Sociedad de evaluación Australasiática, la Alemana y la Británica, son códigos enfocados al ámbito de la evaluación, y como tal, prescriben con detalle los requisitos necesarios para conducir una evaluación de calidad. Sin embargo, los principios de la Asociación Americana de Evaluación, de la Asociación Canadiense, y la Sociedad Francesa son de carácter más general y orientativo. En cualquier caso, todos ellos coinciden

en dos aspectos fundamentales: 1) Consolidar la evaluación como un componente para el fortalecimiento de aquello que se está evaluando, 2) Dar pautas sobre cómo desempeñar la práctica de la evaluación de una manera ética.

El impulso de los códigos normativos en evaluación ha servido de base para que se promueva también el desarrollo de políticas de evaluación a nivel institucional. En realidad, lo que se promueve es que dichas políticas sean escritas y que esa 'tangibilidad' les otorgue una mayor oficialidad, puesto que en aquellas instituciones donde se practica la evaluación, las políticas ya existen, bien sean explícitas y escritas, o no escritas y sujetas a procedimientos más ad hoc. Es decir, las evaluaciones en una institución siempre están guiadas por una determinada política institucional de hacer evaluaciones, ya sea del conocimiento o no de sus propios protagonistas (Trochim 2009).

A pesar de que no existen formas comúnmente aceptadas de qué elementos deben contener las políticas de evaluación en las instituciones, se puede afirmar que puesto que su propósito es fortalecer el ejercicio de la evaluación a través de guiar la realización de las evaluaciones, se centran más en el cómo que en el qué. De esta manera, muchas instituciones donde se practica regularmente la evaluación dan por hecho que a priori, todo es evaluable. Sin embargo, valorar previamente si la intervención en particular es evaluable puede resultar indispensable para poder llevar a cabo una evaluación de calidad.

El concepto de evaluabilidad se desarrolla en la década de los 70 a través de autores como Wholey, Rutman y Cronbach. En sus inicios, el concepto de evaluabilidad estaba más orientado a valorar la evaluabilidad técnica de la potencial evaluación. Así, Cronbach resumía el criterio de evaluabilidad en cuatro aspectos clave: 1) incertidumbre previa ante una cuestión, 2) coste de la información, 3) existencia de información recogida, y 4) posibilidad de aprovechar la informa-

ción en análisis y acciones (Cronbach et al, 1980).

Dichos criterios, presumen la utilización racional de los resultados de la evaluación por los tomadores de decisión, en especial, los de la esfera política. En la práctica, esto no siempre es posible. Chemlinsky (1998) amplía dichos criterios con el fin de incorporar los efectos de carácter político que pueden afectar el éxito de una evaluación. A su juicio existen 8 razones para sí rechazar una evaluación, donde se observan aspectos técnicos y otros más políticos: 1) si no se va a poder realizar de una forma convincente por una limitación razonable de recursos económicos, 2) si hay poca o ninguna base de conocimiento sobre la que construir y el tema es nuevo para el equipo, 3) si no se pueden recopilar datos originales y no hay datos ya existentes disponibles, 4) si no puede ser terminada a tiempo para que resulte útil, 5) si no va a añadir nada nuevo, ni para el programa ni para la práctica de la evaluación, 6) si no hay una posición o postura respecto a la política o programa sobre la que ejecutar las recomendaciones potenciales, 7) si el interés público no puede contrarrestar el resto de las consideraciones anteriores, y 8) si las preguntas propuestas fueran demasiado numerosas, amplias, triviales, sesgadas y no pudieran ser modificadas.

Las consecuencias de la no evaluabilidad técnica son claras. Tampoco es que exista gran uniformidad de criterio entre lo que puede constituir una evaluación técnicamente viable, pero se puede afirmar que a priori es complicado llevar a cabo una evaluación si no se sabe qué se quiere conseguir, qué se considera exitoso o cual es la situación de partida. Igualmente, sin tiempo ni recursos económicos, o sin posibilidad de obtener información, la evaluación tiene muchas posibilidades de fracasar, en el sentido de no cumplir los objetivos para los que fue concebida, bien sea la rendición de cuentas, el aprendizaje, la incorporación de lecciones aprendidas en futuros programas o los que sean.

Igualmente importante a la evaluabilidad técnica, es la evaluabilidad política. La existencia de una cultura institucional que facilite la evaluación es clave para que la evaluación sirva a los objetivos previstos. Dotar a la evaluación de recursos humanos y financieros necesarios, proporcionar la información pertinente a los evaluadores, garantizar que los evaluadores puedan trabajar con independencia sin presiones sobre lo que la evaluación 'debería decir', y contar con un sistema de incorporación de los resultados de la evaluación en la gestión institucional son sólo algunos de los aspectos que podrían a priori presumir que la institución está 'dispuesta', ella o sus programas, a ser objeto de evaluación.

Sin embargo a pesar de que en teoría el análisis de ambos tipos de evaluabilidad facilitarían la realización de evaluaciones de calidad, en la práctica dichos estudios no son habituales. A continuación destaco algunas de las situaciones que dificultan la realización de estudios de evaluabilidad en organismos internacionales:

- ▶ El carácter obligatorio de las evaluaciones. Tal y como menciona Merino (2007), las evaluaciones en la mayoría de los organismos internacionales son de carácter obligatorio mientras que los estudios de evaluabilidad son a lo sumo recomendables. Esto implica que puesto que la evaluación se hará sí o sí, la realización de estudios de evaluabilidad no son relevantes.
- ▶ Los criterios de control que determinan cuándo evaluar. En la mayoría de organizaciones internacionales el criterio que determina cuándo realizar evaluaciones no sigue patrones de evaluabilidad, sino que sigue un parámetro en su mayoría de presupuesto (proyectos de más de X presupuesto), de duración (proyectos de una duración mayor a X periodo de tiempo), o simplemente cíclico (todos los proyectos se evalúan cada X años).

▶ La inevitable rendición de cuentas (accountability) de las evaluaciones. A pesar de que la mejora a través del aprendizaje es un propósito habitual en evaluación, las evaluaciones también clarifican responsabilidades y dan cuenta de la gestión de lo que se evalúa, bien sea a nivel meramente informativo, o con carácter determinante para la continuidad del programa. Sin embargo, aprendizaje y rendición de cuentas tienen audiencias distintas. Así, el aprendizaje va ligado generalmente a los responsables del programa evaluado, y la rendición de cuentas a los donantes y público en general. Un resultado negativo de un estudio de evaluabilidad indica que no se reúnen ni siquiera las condiciones para evaluar, por lo que deja translucir de antemano una gestión posiblemente ineficiente.

▶ El escaso poder de decisión de los evaluadores en aspectos importantes de la evaluación. A pesar de que los evaluadores asumen una gran responsabilidad al conducir una evaluación, en realidad, la capacidad de incidir en aspectos importantes para el éxito de la misma con frecuencia están fuera de su control. De la misma manera que el presupuesto y el tiempo para hacer una evaluación suelen ser parámetros poco negociables, el estudio de evaluabilidad es un 'extra' en ambos aspectos que no suele estar contemplado en sus términos de referencia.

De las cuatro situaciones anteriores se desprende que en la evaluación actual que se efectúa en las organizaciones internacionales no es necesario que los responsables del programa evaluado vean el valor añadido que la evaluación puede ejercer sobre los programas, y por tanto los esfuerzos que suponen los estudios de evaluabilidad dejan de ser prioritarios.

Según Dahler Larsen (2007) la estimación de evaluabilidad ha sido en la actualidad reemplazada por ideas, normas y prescripciones que están basadas en el valor de la evaluación en general, más que en las características específicas de la situación a evaluar. Es decir, la estimación de la evaluabilidad ha per-

dido fuerza en favor del establecimiento de capacidad y culturas de evaluación.

Sin embargo, el fortalecimiento de la evaluación a través del campo normativo y la evaluabilidad no tienen por qué ser excluyentes, puesto que los dos aspectos pueden contribuir a evaluar con una mayor calidad. La solución pasa por hacerlos compatibles. Dicha compatibilidad puede hacerse de diferentes maneras. Un primer escenario consiste en expandir la obligatoriedad de la evaluación a los estudios de evaluación previos. Así se garantizaría la evaluabilidad como parte del proceso de evaluación, aunque como toda norma obligatoria, persistiría el problema de que su potencial uso no sea valorado ni comprendido.

Un segundo escenario consiste en prestar una mayor atención a cuando y cómo introducir la práctica de la evaluación en una institución si se quiere que se valore su potencial uso. Es decir, priorizar desde las instituciones que exista una apreciación entre los responsables del programa sobre lo que la evaluación puede aportar antes de introducir normas sobre su obligatoriedad. De esta manera, si se valora su potencial uso, habrá una mayor disposición a poner los medios necesarios para optimizarla, entre ellos, la estimación de evaluabilidad.

Este segundo escenario, probablemente más sostenible y completo no es tarea fácil debido a que en las

organizaciones internacionales las evaluaciones tienden a servir a varios propósitos al mismo tiempo que a su vez implican a diferentes audiencias, incluyendo algunas de mayor peso que los responsables de los programas que se evalúan. Aún así, el futuro de la evaluación como disciplina capaz de aportar valor a la gestión de los programas puede ser una buena razón para intentarlo.

CONCLUSIÓN

El desarrollo de códigos normativos ha favorecido el fortalecimiento de la evaluación a través de guiar cómo realizar evaluaciones de una manera ética y de calidad. El concepto de evaluabilidad no tiene por qué ser incompatible con la normatividad de la evaluación puesto que contribuye a que la evaluación pueda servir con una mayor precisión al propósito para el que fue concebida.

Actualmente nos encontramos con que en muchas organizaciones no es prioritario el que sus responsables vean el potencial valor añadido de hacer una evaluación, y por tanto, los estudios de viabilidad han dejado de ser relevantes. El entendimiento y apreciación por parte de los responsables de programa sobre el potencial beneficio de la evaluación en su trabajo son claves para fomentar la práctica de los estudios de evaluabilidad en sus procedimientos.

REFERENCIAS BIBLIOGRÁFICAS

Ballart X. 2003 Evaluación de políticas (Marco conceptual y organización institucional), Revista de Asuntos Políticos abril junio 1993.

Bustelo 1998 Deontología de la evaluación: el modelo de los códigos éticos anglosajones. Gestión y Análisis de Políticas Públicas. Enero Agosto 1998.

Chemlinsky E. 1998 The Role of Experience in Formulating Theories of Evaluation Practice, American Journal of Evaluation. Vol 19 N1.

Crombach L.J et al 1980 Toward reform of program evaluation, San Francisco CA, Jossey Bass.

Dahler Larsen P. 2007 ¿Debemos evaluarlo todo? O de la estimación de la evaluabilidad a la cultura de la evaluación, Evaluación de Políticas Públicas, ICE Mayo Junio 2007 N 836.

Merino M. 2007 La evaluabilidad: De instrumento de gestión a herramienta estratégica en la evaluación de políticas públicas, Agencia de Evaluación y Calidad, Ministerio de Administraciones Públicas.

Trochim W 2009 Evaluation Policy and Evaluation Practice, New Directions for Evaluation, American Evaluation Association N 123.

ENTREVISTA A

D^a. Oneida Álvarez .

Coordinadora del Programa Master Latinoamericano en Evaluación de Políticas Públicas.



1. ¿QUÉ LE LLEVÓ A APOYAR, DESDE EL PRINCIPIO, LA PUESTA EN MARCHA DEL MASTER LATINOAMERICANO EN EVALUACIÓN DE POLÍTICAS PÚBLICAS QUE LE PLANTEÓ EL ENTONCES PRESIDENTE DE LA SEE, D. CARLOS ROMÁN DEL RÍO?

En primer lugar, para todos los que hemos tenido el privilegio de compartir espacios académicos con el Dr. Carlos Román, resultaba obvio que su proyecto de Master debía responder a una idea bien pensada y a una concepción técnica y metodológicamente acertada, porque el elevado rigor profesional era una de sus características distintivas. Ello justifica que no tuviera que meditar mucho para aceptar su invitación, desde la perspectiva científica. Fue convincente y preciso al fundamentar la necesidad del referido Postgrado, la utilidad que del mismo podría derivarse para los participantes, y el impacto que era factible esperar como derivación de sus resultados. Adicionalmente, mis estrechos vínculos afectivos con Carlos convertían esa tarea en un placer, más que en una faena, no sólo por el aprendizaje seguro de nuevos conocimientos y experiencias, sino también por la delectación que significaba para sus AMIGOS trabajar junto a tan excepcional personalidad.

A partir de esa aceptación apriorística, inicié un proceso reflexivo durante el transcurso de la primera edición del Master y sus resultados sucesivos, que me han inducido conscientemente a continuar apoyando decididamente el mismo, porque se han constatado las justificaciones que tuvo el entonces presidente de la SEE para proponer su desarrollo. Ahora comparto esas razones más consecuentemente, y explicaré mis argumentos, desde diferentes ángulos, a lo largo de estas reflexiones que la Revista me ha brindado la oportunidad de compartir con sus lectores.

2. ¿COMO COORDINADORA DEL MASTER IBEROAMERICANO EN EVALUACIÓN DE POLÍTICAS PÚBLICAS, CUÁLES FUERON LOS ASPECTOS QUE TUVO EN CUENTA PARA EL DISEÑO DE SU PLAN DE ESTUDIO?

Es necesario precisar que el primer Plan de Estudio respondió a la concepción original del Master que tuvo Carlos Román, pero al finalizar el desarrollo de su etapa presencial, tuvimos la oportunidad de relatarle los aciertos y debilidades detectados por los miembros del claustro y los estudiantes, al tiempo que intercambiamos ampliamente sobre los cambios que se estaba produciendo en las políticas públicas en Latinoamérica. Con todos esos elementos se realizaron, atendiendo a sus sugerencias, algunos ajustes para la segunda edición, que no significaron modificaciones al Plan, sino a los contenidos de los módulos o cursos que lo integraban, así como a la bibliografía y el perfil de algunos docentes, para adecuarlos mejor a los requerimientos de la nueva coyuntura. Es precisamente ahora, después de haber transitado por una experiencia de tres ediciones, que hemos sometido a un juicio más colectivo el diseño del plan de estudio, partiendo de tres aspectos fundamentales:

- los inéditos retos a que se enfrenta hoy el Estado en Latinoamérica para impulsar los procesos de desarrollo, a partir de la compleja situación que impone la crisis internacional aún imperante;
- las modificaciones que se van produciendo en el nivel y el perfil profesional de los estudiantes que optan por este Postgrado;
- la necesidad de incrementar la articulación y la armonía de los diferentes módulos para potenciar el carácter sistémico e integrador del proceso de evaluación de las políticas públicas.

Pueden adicionarse otros argumentos, no menos atendibles, de carácter metodológico, que persiguen modificar la relación entre el tiempo presencial y no presencial de docencia, así como aprovechar mejor las posibilidades que brinda la plataforma virtual de la UNIA y otras tecnologías de la informática y las comunicaciones para la enseñanza de postgrado en la sede iberoamericana de La Rábida.

3. DADA SU EXPERIENCIA DOCENTE E INVESTIGADORA, ¿QUÉ CARENCIAS CONCRETAS VIENE A CUBRIR EL PROGRAMA Y QUÉ LO HACE DIFERENTE?

El Programa del Master puede contribuir a enfrentar carencias de dos ámbitos diferentes, íntimamente vinculados: de carácter formativo y cultural. En los estudios que hemos realizado sobre DNA (determinación de necesidades de aprendizaje) de los funcionarios públicos de Latinoamérica se han detectado múltiples insuficiencias en sus conocimientos sobre Evaluación de Políticas, Programas y Proyectos del sector público a diferentes niveles, así como de aquéllos que se desarrollan en coordinación con organismos internacionales, ONGs o con el sector privado. Adicionalmente, en muchos de los cursos que se han diseñado en otras instituciones no se ofrece una visión integradora de esta disciplina científica, que en ocasiones se continúa enfocando desde paradigmas superados, criterios de gestión atrasados, indicadores de medición poco efectivos y escasa atención a la evaluación de diseño e impactos. De otra parte, los materiales bibliográficos más actualizados para apoyar los cursos provienen en diversas ocasiones de estudios o experiencias de realidades diferentes a las de los países de América Latina y el Caribe.

El Master Latinoamericano en Evaluación de Políticas Públicas que se imparte en La Rábida, bajo la dirección académica del actual Presidente de la Sociedad Española de Evaluación, Dr. José Luís Osuna, se ha propuesto subsanar esas limitaciones, y en ello estriba una de sus diferencias

respecto a otros Programas. Además, pone especial énfasis en la divulgación de las más avanzadas experiencias prácticas realizadas por prestigiosos evaluadores de disímiles latitudes en sectores y actividades variadas, y con una metodología ampliamente participativa. Con ello favorece la cultura y la ética de evaluación desde una perspectiva que estimula la transparencia, y coadyuva a la promoción de un mejor y mayor vínculo de la ciudadanía con los procesos de diseño y rendición de cuentas de las políticas públicas.

No solo los contenidos, sino también los métodos de enseñanza hacen singular este Programa, que se desarrolla con una combinación acertada de teoría/práctica, trabajo individual/grupal, e instrucción problematizada, poniendo en el estudiante y sus capacidades de acción/reflexión el centro del proceso de enseñanza-aprendizaje.

4. EN QUÉ MEDIDA CREE EN LA OPORTUNIDAD DE ESTE TIPO DE FORMACIÓN ESPECIALIZADA; DESDE LA PERSPECTIVA DE LA MEJORA EN LA GESTIÓN PÚBLICA?

El avance de la gestión pública depende de múltiples factores, pero sin dudas los niveles de calificación de los recursos humanos encargados del diseño de las estrategias y políticas, su aplicación y su evaluación constituyen un factor de especial connotación. De ahí que los Programas de Formación que fertilicen esos objetivos desempeñan, en mi opinión, un papel trascendente en dicha dirección. La formación y superación no sólo debe contemplar los aspectos más generales que atañen a la gestión, sino además los más especializados, como los de evaluación. Concretamente este perfil posibilita que funcionarios y ciudadanos, promotores y beneficiarios de los proyectos, gobernantes y gobernados se sientan copartícipes y corresponsables de las estrategias y políticas, del uso de los recursos y del control de éstos. El Programa de nuestro Master, sin dudas, brinda una oportunidad para pasar de una gestión centralizada, tecnocrática y burocrática a un nuevo estilo de dirección estatal/sectorial/territorial más participativo y efectivo.

5. ¿QUÉ PERFIL PRESENTAN LOS ALUMNOS QUE ACCEDEN AL MASTER? RECIENTEMENTE HA FINALIZADO LA III EDICIÓN DEL MISMO, EN ESTOS TRES AÑOS, ¿SE HA PRODUCIDO UNA EVOLUCIÓN DEL PERFIL DE LOS ALUMNOS?

El perfil de los estudiantes que acceden al Master ha ido variando desde la primera edición, porque se ha logrado una composición más diversa en cuanto a los países de procedencia, las especialidades profesionales y las funciones laborales. En el Curso 2009-2010 hemos tenido alumnos de la subregión andina, del cono sur, y de Centroamérica, graduados de muy heterogéneas carreras universitarias y muchos de ellos funcionarios de nivel medio y medio-alto de las administraciones públicas e instancias políticas a nivel central y territorial en América Latina. Esa heterogeneidad ha enriquecido el intercambio de culturas, conocimientos y experiencias de todo tipo, pero muy especialmente en los temas vinculados a la gestión y evaluación de las políticas públicas, y ha propiciado el acercamiento, tolerancia, entendimiento, y mayor confianza entre identidades no siempre próximas por razones de distinta índole. Ello puede considerarse un resultado marginal, pero no por ello poco importante, de este Programa de Master. En su seno se van generando Redes de trabajo a favor de una cultura común de Evaluación Latinoamericana, que podrá tener en un futuro cercano su expresión en formalizaciones de alta capacidad para potenciar la transparencia, eficacia y la eficiencia de la gestión pública en la región.

6. HAGA DE “ABOGADA DEL DIABLO”, ¿CUÁLES SON LAS DIFICULTADES Y LAS OPORTUNIDADES DE PARTICIPAR EN EL MASTER, DESDE EL PUNTO DE VISTA DEL PROFESORADO?

Considero que formar parte del claustro del Master es ante todo una gran responsabilidad, por las exigencias de rigor que el Programa implica en cuanto a su concepción holística y diseño metodológico, así como por los requerimientos de ordenación académica que la UNIA ha establecido para lograr que sus títulos propios respondan a los estándares de alta calidad que les caracteriza. Alumnado interesado, motivado, preparado y crítico completa la trilogía de interpelaciones.

Al propio tiempo, significa una oportunidad y un privilegio poder ser seleccionado para integrar dicho colectivo docente y compartir estrado con especialistas, investigadores, funcionarios y profesores de alta talla profesional y gran prestigio en sus respectivas especialidades. En ocasiones puede parecer redundante presentar a la mayoría de ellos, diciendo que es un verdadero lujo contar con su presencia en La Rábida para compartir sus saberes y vasta experiencia con los condiscípulos que han logrado ser aceptados para cursar ese Programa de Postgrado. Procedentes de muy diversas esferas y regiones, más de 4 decenas de expertos se conjugan para dar el más reciente nivel de actualización en su perfil concreto de actividad y contribuir a que los estudiantes integren desde disímiles perspectivas y técnicas un juicio sistémico del complejo proceso de evaluación de las políticas públicas, con sus más variados componentes y matices. La mayor dificultad, a mi juicio, pudiera ser no alcanzar el nivel de preferencia de los estudiantes en un contexto de altísima competencia y expectativas, o resultar desfavorablemente evaluado como docentes por quienes han hecho de la aplicación de severos criterios de evaluación una especialidad más de sus capacidades cotidianas.

7. ¿QUÉ VALORACIÓN LE MERECE LA COMUNIÓN DE INTERESES ENTRE LA SOCIEDAD CIVIL (SEE), LA ADMINISTRACIÓN PÚBLICA (AEVAL) Y LA UNIVERSIDAD (UNIA) COMO MECANISMO PARA DAR CABIDA A INICIATIVAS FORMATIVAS COMO LA QUE VD. COORDINA?

Tengo en alta estima el significado y los resultados de esta iniciativa, y de la posibilidad de conjunción de tan complementarios factores para asegurar el éxito del Programa. Para mí ha sido una segunda experiencia, porque ya en una ocasión tuve la suerte de coordinar un proyecto similar en Sudamérica, auspiciado por universidades bolivianas, un instituto internacional, una gran corporación privada y diferentes organizaciones de la sociedad civil. Pero faltó entonces el decidido apoyo de la administración pública y la vocación de excelencia y latinoamericanista que aportan a este Master la AEVAL y la UNIA. Como parte de la Coordinación Académica del proyecto y como ciudadana identificada y comprometida con las necesidades de desarrollo del “nuevo mundo”, agradezco profundamente el empeño de la actual dirección de la Universidad Internacional de Andalucía, y muy especialmente de su sede iberoamericana de La Rábida, a favor de este esfuerzo que tanto puede aportar al perfeccionamiento de las políticas y la gestión pública de nuestros países.

8. POR ÚLTIMO, CON SU AMPLIA EXPERIENCIA ¿CUÁL ES EL PAPEL DE UNA UNIVERSIDAD INTERNACIONAL, CON CLARA VOLUNTAD DE COOPERACIÓN CON AMÉRICA LATINA – VR. GR. LA UNIA – EN EL MARCO DE UN FUTURO PROCESO DE CREACIÓN DE UN ESPACIO UNIVERSITARIO LATINOAMERICANO COMÚN A SIMILITUD DEL EUROPEO.?

Somos protagonistas actualmente de un renovado espíritu integracionista en América Latina y el Caribe, impulsado por muy diversas fuerzas, entre las que destaca la voluntad política de sus gobernantes y otros relevantes actores económicos y socio-políticos. Pero ese proceso no está exento de obstáculos, dificultades y contradicciones. Resulta imprescindible aunar todas las fuerzas que puedan apoyarlo. La creación y fortalecimiento paulatino de un espacio latinoamericano común es una de esas oportunidades, no sólo por lo que puede contribuir a potenciar la unidad de la región, sino además por su contribución al perfeccionamiento de las capacidades de la principal fuerza productiva del área: sus recursos humanos. Los centros de educación superior con su despliegue de actividades docentes, investigativas y proyectos de extensión a las comunidades donde se encuentran, se convierten en potencias dinamizadoras de la “inteligencia social.” Esta es el pivote esencial para enfrentar los retos del desarrollo en las actuales condiciones de sociedades basadas en el conocimiento, de ahí que ratifico en esta oportunidad el criterio que he reiterado a lo largo de mi estancia en La Rábida durante los últimos meses: la contribución que la UNIA está realizando es decisiva para hincar los pilares del desarrollo endógeno de América Latina. Agradezco al Director del Master y a la dirección de la UNIA por permitirme tributar con mis modestos esfuerzos a este loable empeño del espacio universitario común latinoamericano.

EASY-ECO Brussels Conference 2010

17-19 November 2010

**Sustainable Development Evaluations in Europe:
From a Decade of Practices, Politics and Science to
Emerging Demands****Call for papers**

This event is sponsored by



MARIE CURIE ACTIONS

EASY-ECO has been awarded an Official Decade Project by the German Commission for UNESCO.

Status: October 2009

**SOCIEDAD SUIZA DE EVALUACIÓN**

The next congress of the Swiss Evaluation Society will take place on the 9th and 10th of September 2010 in Bern. The topic is "International comparisons and Peer-Reviews : what can we learn as an evaluator or an evaluatee ?". On the afternoon before the meeting a methods workshop will be held. At this occasion, evaluations of innovative or interesting methodological issues will be presented and discussed. The presentations will be in German, English and French.

More information: <http://www.seval.ch/fr/veranstaltungen/kongresse.cfm> (French) or <http://www.seval.ch/de/veranstaltungen/kongresse.cfm> (German).

**THE GLOBAL EVALUATION CONCLAVE
MAKING EVALUATION MATTER****Comienza: 25/10/2010 - Finaliza: 28/10/2010**

El próximo mes de Octubre tendrá lugar en Nueva Delhi la conferencia "Making Evaluation Matter". La conferencia reunirá a profesionales de la evaluación de todo el mundo y servirá para aquellos que quieran poner sus conocimientos en nuevas direcciones y estén interesados en aplicar sus ideas en el contexto sudasiático.

Para más información: www.evaluationconclave.org

INTERNATIONAL SUMMER SCHOOL ON EVALUATION RESEARCH

Comienza: 30/08/2010 - Termina: 3/09/2010



El próximo 30 de agosto tendrá lugar en Lille un Curso de verano sobre Evaluation Research .

Esta Escuela de verano pretende ser una oportunidad ideal para aquellos de la comunidad científica interesados en la evaluación de políticas públicas comparen trabajos científicos sobre la materia y la discusión sobre la naturaleza interdisciplinar de la estudios de evaluación.

El plazo para la recepción de la solicitudes es hasta el 7 de Mayo de 2010.

Para más información, consultar los siguientes enlaces: http://www.sfe-asso.fr/sfe-evaluation.php?menu_id=473 o sfe.universite-ete@orange.fr



9ª CONFERENCIA INTERNACIONAL DE LA SOCIEDAD EUROPEA DE EVALUACIÓN

Comienza 06/10/2010 - 11:37 - Finaliza 08/10/2010 - 11:37

Bajo el tema general "Evaluación en el Interés Público: Participación, Política y Política" la Sociedad Europea de Evaluación ha puesto en marcha una convocatoria de propuestas para su Conferencia Bienal. Este evento se celebrará en la ciudad de Praga en la República Checa, entre el 6 y el 8 de octubre de 2010.

Más información: <http://www.europeanevaluation.org/conferences.htm>

SEMINARIO EUROPEO

Comienza: 16/09/2010 - Finaliza: 17/09/2010

La Comisión Europea destaca la importancia de los métodos de evaluación para las operaciones de los Programas de los Fondos Europeos. Es necesario que los métodos de evaluación estén cada vez más vinculados a los sistemas de supervisión de los Programas.

Esta es la razón por lo que una adecuada herramienta de métodos de evaluación es crucial para el éxito de las actividades de la gestión de los Fondos de la Unión Europea.

Por lo que tendrá lugar en Berlín el Seminario Europeo:

"Monitoring and Evaluation Methods for EU Funds-

How to Measure the Success of Projects and Programmes"



CONGRESO INTERNACIONAL DE EVALUACIÓN III CONFERENCIA RELAC

San José, Costa Rica, 28-30 julio 2010

**Evaluación y participación ciudadana en América Latina: Debates
y posibilidades en un contexto de**

Reformas sociales y cambios políticos

La Universidad de Costa Rica (UCR) y la Red Latinoamericana de Seguimiento Evaluación y Sistematización en América Latina y El Caribe (ReLAC) invitan a usted al Congreso Internacional de Evaluación y a la III conferencia ReLAC, a realizarse del **28 al 30 de Julio de 2010 en San José, Costa Rica.**

TEMAS

Los tres grandes ejes temáticos a trabajar son los siguientes:

- Evaluación y Sistematización en América Latina: situación actual, tendencias y perspectivas.
- Nuevas funciones e institucionalidad de la evaluación.
- Profesionalización de la evaluación en América Latina.

TIPOS DE PRESENTACIONES

El Comité Organizador y el Consejo Académico del Congreso Internacional y de la III Conferencia recibirán con mucho interés las **ponencias, comunicaciones y posters** que aborden temas que forman parte de los contenidos de los ejes temáticos definidos.

- Ponencias: extensión máxima de 500 palabras y el texto final no debe superar las ocho mil (8.000) palabras. Las ponencias serán integradas a paneles de discusión organizados en función de temas comunes.
- Propuestas de paneles o coloquios: reúne de 3 a 5 presentaciones sobre un tema o problemática común. El proyecto de panel o coloquio deben ser presentados por escrito identificando, a su vez, los participantes y breve resumen de sus presentaciones.
- Posters: presentación grafica de una experiencia, investigación o evaluación y relacionada con los ejes temáticos del Congreso y de la Conferencia.
- Talleres de Capacitación: Serán el 26 y 27 de Julio en torno a problemáticas, métodos y técnicas que tengan un carácter innovador.

El plazo de presentación de las ponencias, paneles y posters será hasta el 31 de marzo.

SOCIOS INSTITUCIONALES QUE CONTRIBUYEN AL CUMPLIMIENTO DE LOS FINES DE LA **SOCIEDAD ESPAÑOLA DE EVALUACIÓN** DE POLÍTICAS PÚBLICAS



Fundación Andaluza Fondo de Formación y Empleo
CONSEJERÍA DE EMPLEO



Andalucía Emprende, Fundación Pública Andaluza
CONSEJERÍA DE EMPLEO
CONSEJERÍA DE ECONOMÍA, INNOVACIÓN Y CIENCIA



Centro de Estudios Andaluces
CONSEJERÍA DE LA PRESIDENCIA



www.sociedaddeevaluacion.org

The collage displays several key pages from the website:

- Inicio:** Shows the main navigation menu and a search bar.
- Calendario de Eventos:** Lists upcoming events such as 'CONFERENCIAS DE LA SEE' and 'REUNIÓN SEE-VALUACIÓN'.
- Recursos para Evaluadores:** Provides information on 'Bases de Datos', 'Conferencias', 'Jornadas', 'Estudios y Ensayos', 'Congresos', 'Informes', 'Guías', and 'Publicaciones'.
- Noticias:** Features articles like 'NOTICIA EUROPEA POR LA TRANSPARENCIA PUESTA EN FUNCIONAMIENTO DEL "REGISTRO DE REPRESENTANTES DE INTERÉS"'.
- Inscripción a la Sociedad Española de Evaluación de Políticas Públicas:** Details the process of joining the society, including membership types (Individual, Institutional, Life) and fees.

¡¡¡Visítanos!!!

Edita:



SOCIEDAD
ESPAÑOLA
DE EVALUACIÓN

Financian:



Unión Europea

FONDO SOCIAL EUROPEO



JUNTA DE ANDALUCÍA

Servicio Andaluz de Empleo
CONSEJERÍA DE EMPLEO