

# **Agencia de Evaluación: innovación social basada en la evidencia**

**Rafael Pinilla Pallejà**

Documento de trabajo 91/2006



## Rafael Pinilla Pallejà

Médico de la Sanidad nacional, psicólogo y economista. Ha trabajado para el Ministerio de Sanidad, la Consellería de Sanitat de la Generalitat Valenciana y, en la actualidad, para el Ministerio de Administraciones Públicas. Ha sido profesor asociado de Análisis económico en la Universidad de Valencia y es investigador colaborador del Instituto Valenciano de Investigaciones Económicas.

Ha publicado diversos artículos y libros relacionados con la evaluación de la calidad de vida, la renta básica de ciudadanía y la política basada en la evidencia. Actualmente es editor jefe de la revista *Basic Income Studies* y promueve un proyecto de investigación dirigido a evaluar en España, mediante un experimento con asignación aleatoria, la efectividad y eficiencia de distintas políticas de garantía de una renta de ciudadanía.

Ninguna parte ni la totalidad de este documento puede ser reproducida, grabada o transmitida en forma alguna ni por cualquier procedimiento, ya sea electrónico, mecánico, reprográfico, magnético o cualquier otro, sin autorización previa y por escrito de la Fundación Alternativas

© Fundación Alternativas

© Rafael Pinilla Pallejà

ISBN: 84-96653-05-6

Depósito Legal: M-20907-2006

## Contenido

<b>Resumen ejecutivo</b> .....	<b>5</b>
<b>Introducción</b> .....	<b>7</b>
<b>1. El reto de la Agencia de Evaluación</b> .....	<b>10</b>
1.1 Recomendaciones para una mayor utilización .....	10
1.2 La institucionalización de la evaluación en España .....	11
<b>2. Evaluación e innovación institucional</b> .....	<b>15</b>
2.1 Estructura y dinámica de las instituciones sociales .....	15
2.2 Innovación social intencional .....	17
<b>3. Hacia la política basada en la evidencia</b> .....	<b>19</b>
3.1 El ejemplo de la medicina basada en la evidencia .....	19
3.2 El proceso de la política basada en la evidencia .....	20
<b>4. Lecciones de los experimentos sociales</b> .....	<b>25</b>
<b>5. La búsqueda de estrategias políticas robustas</b> .....	<b>28</b>
5.1 Decisiones bajo incertidumbre .....	28
5.2 Estructura FASE .....	31
5.3 Institucionalización de la evaluación de políticas .....	33
<b>6. Recomendaciones para la puesta en marcha de la Agencia de Evaluación</b> .....	<b>35</b>
6.1 Elementos para un plan estratégico de transición a la innovación social basada en la evidencia .....	37
6.2 A modo de conclusión .....	38
<b>Índice de Gráficos</b> .....	<b>40</b>
<b>Bibliografía</b> .....	<b>41</b>

**Siglas y abreviaturas**

CC AA	Comunidades autónomas
EPDS	Encuesta sobre Pobreza y Desigualdades Sociales del País Vasco
EE UU	Estados Unidos de América
FASE	Factores, Acciones, Simulación, Evaluación
I+D	Innovación y desarrollo
ISBE	Innovación social basada en la evidencia
MAP	Ministerio de Administraciones Públicas
MBE	Medicina basada en la evidencia
MDRC	Manpower Demonstration Research Corporation
OCDE	Organización para la Cooperación y el Desarrollo Económicos
PBE	Política basada en la evidencia

## Agencia de Evaluación: innovación social basada en la evidencia

**Rafael Pinilla Pallejà**

Médico de la Sanidad nacional

La Agencia de Evaluación se creó como respuesta a la creciente demanda ciudadana de eficacia y eficiencia en la gestión de los recursos públicos. El reto de la Agencia es conseguir que la evaluación se utilice de forma normal para mejorar la eficacia de las políticas públicas. Para ello hay que promover evaluaciones con el mayor rigor científico, y tratar de que el estamento político las acepte y participe en la evaluación.

La adopción de la evaluación sistemática de los programas y políticas públicas representa un cambio importante para el proceso político tradicional. La incorporación de la evaluación al proceso político tradicional puede dar lugar a un proceso de “política basada en la evidencia” (PBE) que mejore la competencia de la práctica política.

Tomando en consideración las mejoras que los procedimientos científicos pueden aportar a la racionalidad de las decisiones políticas y la metodología de análisis de la robustez, este trabajo incluye las siguientes recomendaciones:

- Para afrontar con éxito la etapa de puesta en marcha de la Agencia se requiere un plan estratégico que guíe su funcionamiento y actividades durante los primeros tres o cuatro años.
- El plan estratégico debe tener dos objetivos esenciales: a) demostrar capacidad científica para llevar a cabo evaluaciones con el máximo rigor, y b) lograr la participación de los responsables políticos y la aceptación social de las evaluaciones.
- Durante la primera etapa, será necesario llevar a cabo una cuidadosa selección de los programas a evaluar. Criterios a tener en cuenta para esta selección: a) impacto presupuestario o social; b) visibilidad política; c) carácter innovador e incertidumbre de los efectos; d) carácter multisectorial; e) metodología emblemática (evalua-

ción experimental); f) participación de las comunidades autónomas; g) posibilidad de resultados decisivos en menos de cuatro años; y h) existencia de consenso político y social en los objetivos del programa.

- Es importante elaborar y validar en breve un sistema básico de información para la evaluación de la efectividad de las políticas que sea de aplicación general. Este sistema debería incorporar indicadores que midan los logros alcanzados en términos de capacidades humanas individuales.
- Dar prioridad a la realización de evaluaciones con metodología experimental.
- Es importante disponer de un procedimiento de análisis de la robustez que permita realizar evaluaciones previas *ex ante* y elaborar los diseños de contraste empírico de acuerdo con ello.

## Introducción

Con el envío al Congreso de los Diputados del Proyecto de Ley de Agencias en julio de 2005 se puso en marcha el proceso de creación de la Agencia de Evaluación de Políticas Públicas y Calidad de los Servicios. La disposición adicional primera de dicho Proyecto de Ley autoriza al Gobierno para “crear la Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios, adscrita al Ministerio de Administraciones Públicas (MAP), cuyo objeto es la promoción y realización de evaluaciones de las políticas y programas públicos, favoreciendo el uso racional de los recursos públicos y el impulso de la gestión de la calidad en los servicios” (MAP, 2005:17).

La creación de la Agencia trata de dar respuesta a la creciente demanda ciudadana de eficacia y eficiencia en la gestión de los recursos públicos. De acuerdo con las recomendaciones del Informe de la Comisión de Expertos, la misión de la Agencia debería ser “contribuir a la mejora de la acción y resultados en la actividad del sector público y sus instituciones, mediante la articulación de un sistema público de evaluación de políticas públicas y de mejora de la calidad de los servicios, en un marco de transparencia y ética institucional” (Garde, 2005:85).

Aunque el nombre y la misión de la Agencia aluden tanto a políticas como a servicios, este trabajo se centrará esencialmente en la problemática de la evaluación de políticas, entendiendo por tales los sistemas de programas públicos con los que se trata de alcanzar objetivos sociales concretos, por ejemplo, políticas de empleo, de salud, de lucha contra la pobreza, de vivienda, de educación, etc.

En todos estos casos existe una finalidad social a lograr, y el Gobierno articula una serie de programas que llevan aparejado el gasto de recursos presupuestarios. La evaluación de políticas plantea preguntas como las siguientes:

- ¿En qué medida los recursos que se destinan a estas políticas están logrando los resultados deseados? O dicho de otro modo, cuál es la **efectividad** o eficacia real de los programas públicos existentes.
- ¿En qué medida podría mejorarse la efectividad aumentando el gasto destinado a estos programas? O bien, cuál es la **efectividad marginal**, es decir, cómo mejoraría el efecto al aumentar los recursos de un programa, o cómo puede empeorar el resultado si disminuimos o eliminamos un programa.

- ¿Podría lograrse una mayor **eficiencia** en la utilización de los recursos públicos? ¿Se puede conseguir más efecto con el mismo presupuesto, o el mismo efecto con menos recursos?

No cabe duda de que disponer de conocimiento fiable sobre la efectividad, la efectividad marginal y la eficiencia de los programas públicos facilitaría enormemente la toma de decisiones políticas. Por ello la evaluación de políticas ha de considerarse como una “herramienta indispensable en la actuación de los poderes públicos” (Monnier, 1995:12). Sin embargo, la evaluación está lejos de ser una práctica habitual en la Administración pública y todavía es muy escasa la evidencia científicamente fiable sobre efectividad y eficiencia de los programas públicos. Esta escasez se debe en gran parte a las dificultades de obtener evidencia científicamente fiable sobre la efectividad de las políticas públicas. Pero, también, a la reticencia de los responsables políticos a la evaluación rigurosa de sus propuestas.

En cuanto a la dificultad de obtener una evidencia fiable, Schmidt sintetiza las dificultades de la investigación evaluativa: a) la dificultad para elegir medidas de efecto adecuadas, sin las cuales no es posible decidir si un programa ha tenido éxito; b) la dificultad para medir todos los costes y beneficios de un programa, tanto los directos como los indirectos, sin lo que no es posible valorar adecuadamente la eficiencia; y c) la atribución causal del impacto a una intervención política, sin la cual no es posible evaluar una política. Este último es el problema fundamental de la evaluación (Schmidt, 2001).

En cualquier caso, el rigor en el método para poder atribuir correctamente los efectos del programa es una condición necesaria, pero no suficiente, para la utilidad de una evaluación. Es también fundamental modificar el modo tradicional de proceso político basado en creencias por un proceso político basado en la evidencia y favorable a la innovación social. Este cambio es necesario porque para garantizar la utilidad es preciso tomar decisiones sobre qué criterios utilizar para valorar la efectividad de los programas y asegurar que la evaluación responda a las preguntas que más interesan a los responsables de la toma de decisiones, ya sea gobernantes o legisladores, y a los beneficiarios del programa.

Estos dos últimos aspectos, la elección de indicadores de efecto y de las preguntas relevantes a evaluar, no son cuestiones estrictamente técnicas. No pueden resolverse únicamente con metodología científica. Se necesita la participación y la iniciativa de responsables políticos y representantes de los beneficiarios en el proceso de evaluación. De hecho, existe un amplio consenso en la bibliografía en la necesidad de lograr la participación de los afectados y el compromiso de los responsables políticos en el proceso de evaluación. Este es posiblemente el mayor reto de la Agencia: lograr la participación voluntaria y el compromiso de los responsables políticos e incluso que sean ellos quienes tomen la iniciativa. Se analizará esta cuestión en el capítulo 1.

Si la misión de la Agencia es mejorar la acción y resultados del sector público mediante la evaluación, de lo que se trata es de promover la innovación en la gestión de los recursos



públicos. Es decir, la evaluación debería servir para decidir la sustitución de programas ineficientes por otros más eficientes, así como para conseguir mejores resultados con la misma cantidad de recursos públicos; en definitiva, promover la innovación social basada en la evidencia (ISBE). En el capítulo 2 se establecerá brevemente un marco conceptual de innovación social conveniente para entender apropiadamente el papel de la Agencia de Evaluación como organización promotora de un proceso intencional de mejora institucional.

La idea de mejorar la práctica aprendiendo de la evidencia científica es la base del método científico. Pero trasladar la evidencia científica a la práctica conlleva una serie de dificultades. La medicina basada en la evidencia (MBE), que se ha extendido rápidamente desde 1995, constituye un buen ejemplo del que aprender para llevar al campo de la política la utilización sistemática de la evidencia en la toma de decisiones. Se tratará de este tema en el capítulo 3.

En el capítulo 4 se expondrán las principales enseñanzas de la experiencia internacional en evaluación experimental de programas sociales. La evaluación mediante contraste de grupos con asignación aleatoria es el modelo ideal de evaluación científica. Sin embargo, la experimentación social plantea serias dificultades prácticas y éticas. Desde el punto de vista científico, la metodología experimental se considera el modo más fiable de evaluar el efecto de una política. Pero, dado que no siempre es posible llevar a cabo experimentos, se han desarrollado métodos no experimentales que pueden ser de utilidad cuando no es factible o razonable utilizar procedimientos de asignación aleatoria y para proporcionar información complementaria. Una de las funciones de la Agencia deberá ser el impulso de las metodologías de evaluación más rigurosas, estimulando, cuando sea factible, la evaluación experimental.

En el capítulo 5 se expondrá una orientación metodológica innovadora para superar simultáneamente las dos grandes dificultades de la innovación social basada en la evidencia: la necesidad de incorporar en el proceso de innovación a los responsables políticos y las limitaciones de los métodos experimentales en el campo de las ciencias sociales. La ciencia social tiene mucho que aportar al proceso político mejorando el rigor de la evaluación, pero no debería tratar de imponer a los políticos soluciones supuestamente óptimas. Los científicos sociales deberíamos ser más modestos si queremos que nuestra aportación al proceso político sea más eficaz. Dado que la realidad social es esencialmente impredecible y está sujeta a profunda incertidumbre, científicos sociales y políticos necesitamos colaborar estrechamente para buscar estrategias robustas, renunciando al logro de un óptimo social que ni existe, ni puede existir.

Finalmente en el capítulo 6 sintetizaré lo que puede ser una serie de recomendaciones prácticas de utilidad para la puesta en práctica de la Agencia de Evaluación.

## 1. El reto de la Agencia de Evaluación

El gran reto de la Agencia de Evaluación es lograr que las evaluaciones que promueva se utilicen de forma habitual para mejorar las instituciones y las políticas públicas, esto es, aprovechar la oportunidad que representa su creación para poner en marcha un proceso sistemático de innovación social basada en la evidencia (ISBE).

La actividad evaluadora ha crecido de forma muy importante en las últimas décadas en los países desarrollados. Sin embargo, existe un amplio consenso internacional entre los especialistas sobre la escasa utilización de las evaluaciones, así como en la necesidad de recomendar la participación de los responsables políticos y administrativos en el proceso evaluador para tratar de corregir este problema de infrautilización (Alvira Martín, 1991; Monnier, 1995; OCDE, 1992; Pazos, 1996; Vedung, 1997; Schmidt, 2001). Por tanto, el gran reto de la Agencia no es tanto promover y extender la cultura de evaluación en la Administración española, como lograr que las evaluaciones que se lleven a cabo sean utilizadas de forma habitual para mejorar la efectividad y la eficiencia de las políticas y programas públicos.

Cualquier tipo de evaluación no conduce necesaria, ni directamente, a mejorar la realidad. La evaluación sólo es una oportunidad para aprender a mejorar. Pero lo que realmente importa es convertir esa oportunidad en una mejora real. Para lograr la utilización es necesario que las evaluaciones realizadas sean metodológicamente rigurosas, de modo que los resultados sean fiables. Pero el rigor no es suficiente, es preciso, además, que las evaluaciones sean aceptadas y utilizadas por los responsables políticos para mejorar la eficacia de las instituciones públicas. El propósito fundamental de este trabajo es suministrar algunas ideas para que la Agencia de Evaluación afronte con éxito este reto desde su puesta en marcha. Como se ha comentado, la mayoría de los expertos en evaluación coinciden en recomendar la participación en el proceso evaluativo de los responsables políticos y administrativos de los programas a evaluar. Se trata casi siempre de recomendaciones un tanto vagas, pero es ilustrativo enumerar las más frecuentes.

### 1.1 Recomendaciones para una mayor utilización

- Parece evidente que si se desea que la evaluación conduzca a la mejora de la efectividad de los programas públicos es esencial lograr la participación activa de los responsables políticos, tanto del poder ejecutivo como del legislativo. La evaluación sería

probablemente mejor valorada y utilizada por los responsables políticos si se logra que la contemplen como una herramienta a su servicio, en vez de como un medio de exigirles responsabilidades.

- La utilización de la evaluación requiere un compromiso de impulso político duradero tanto desde el Gobierno como desde la oposición, porque con frecuencia los procesos de evaluación resultan costosos y se dilatan en el tiempo desde el momento de la gestación hasta la obtención de resultados. Una forma evidente de materializar el compromiso duradero con la evaluación es su institucionalización. En este sentido, la creación de la Agencia Estatal de Evaluación puede ser un paso muy importante.
- Para lograr ese compromiso político duradero es también muy conveniente establecer un consenso amplio entre las principales fuerzas políticas y sociales sobre la necesidad de evaluar las políticas públicas mediante procedimientos objetivos y rigurosos. En este sentido, los objetivos constitucionales asumidos por el conjunto de la sociedad pueden ser un buen punto de partida para establecer un consenso; por ejemplo, en la elaboración de un conjunto de indicadores útiles para la evaluación de los efectos de las políticas, así como en la elaboración del tipo de preguntas relevantes a las que debe responder una evaluación concreta.
- La participación de los responsables políticos no debería significar un menor rigor metodológico. Es necesario compatibilizar el procedimiento científico riguroso con el procedimiento democrático. Ambos procedimientos son esenciales.
- Es necesario que la evaluación pueda llevarse a cabo a un coste razonable en relación con el beneficio potencial que cabe esperar de ella. El coste de la evaluación es razonable si representa un porcentaje pequeño del presupuesto total de un programa.
- La evaluación debe ser capaz de ofrecer resultados relevantes para la toma de decisiones a tiempo de su utilización. Para ello es muy conveniente la obtención de resultados intermedios y cierta flexibilidad para modificar el propio proceso de evaluación en función de los resultados intermedios.
- En países con estructura federal o descentralizada, la evaluación cobra especial relevancia y se debe poner especial atención a la coordinación de esfuerzos entre las distintas administraciones.

## 1.2 La institucionalización de la evaluación en España

Las primeras evaluaciones en España se llevan a cabo en 1991 y son demasiado generales y justificativas (Pazos, 1996). Desde entonces se ha avanzado en distintos ámbitos, sobre todo

como consecuencia del estímulo procedente de las instituciones europeas (Garde, 2005). Pero no cabe duda de que la creación de la Agencia de Evaluación representa una oportunidad decisiva para institucionalizar en la Administración española la evaluación sistemática de políticas y programas. Los inicios de la Agencia pueden ser determinantes para asentar rápidamente la cultura evaluativa, pero sólo si la Agencia consigue demostrar la utilidad política de la evaluación.

Los primeros pasos de la Agencia deberían considerarse, por tanto, como una etapa de transición desde la situación actual de “evaluación insuficiente” a la institucionalización de la evaluación propiamente dicha. Por ello, “es preciso acometer la evaluación de políticas públicas y calidad de los servicios configurando un marco estratégico articulador en el que se aborden las carencias detectadas y se logre que la evaluación y la calidad de lo público se inserten en el proceso de decisión y de gestión desde sus inicios” (Garde Roca, 2005:73). En esta fase de transición es clave demostrar que la evaluación puede ser una herramienta eficaz para facilitar la toma de decisiones y mejorar la eficiencia en la administración de los recursos públicos. En este sentido puede ser conveniente que los responsables de la Agencia elaboren un plan estratégico para la institucionalización de la evaluación que oriente las actividades de la Agencia en sus primeros pasos. Este plan debería tener en cuenta la importancia de seleccionar adecuadamente las primeras evaluaciones y la necesidad de ofrecer a los responsables políticos un entorno de incentivos favorable para que sometieran voluntariamente a evaluación los programas que de ellos dependen.

### **A) Criterios para seleccionar las primeras evaluaciones**

Antes de establecer un conjunto de criterios adecuados para la selección de políticas a evaluar es preciso tener en cuenta las siguientes consideraciones:

- En España existe más experiencia en evaluación de la calidad de los servicios que en evaluación de políticas, pero es de la mayor importancia que la Agencia afronte desde el primer momento la evaluación de políticas y programas para demostrar la utilidad de la evaluación para la toma de decisiones.
- Al seleccionar las evaluaciones a realizar será preferible dar prioridad a la calidad sobre la cantidad. No se puede evaluar todo. Hay que seleccionar algunas políticas y sería muy conveniente para la consolidación de la Agencia realizar algunas evaluaciones que puedan ser emblemáticas por el rigor metodológico y la fiabilidad de los resultados. En este sentido, lo más eficaz sería promover al menos una evaluación de tipo experimental. Es preciso recordar que una evaluación tiene éxito cuando proporciona respuestas fiables, aunque el resultado de la evaluación sea la recomendación de rechazar una propuesta de reforma.
- Al seleccionar los programas a evaluar habría que dar prioridad a la visibilidad política, esto es, atreverse a la evaluación de algún programa nuevo que resalte entre el conjunto de políticas del Gobierno.

- Dada la estructura descentralizada del Estado, conviene que la Agencia se proponga como objetivo la evaluación de programas y políticas públicas en cooperación con otras administraciones, en especial con las comunidades autónomas (CC AA).
- En la elección de la política o programa a evaluar es de la máxima importancia que exista suficiente consenso político y social en algunos de los objetivos a lograr, de modo que la evaluación ayude a decidir sobre programas alternativos para alcanzar objetivos de consenso. De este modo, la Agencia de Evaluación contribuirá a la toma de decisiones de consenso y podrá ganar reconocimiento como herramienta útil para los responsables políticos de distintas tendencias.

Teniendo en cuenta las recomendaciones para favorecer la utilización política de los resultados de la evaluación y el contexto de institucionalización de la evaluación en España, se establecerá posteriormente un listado de los criterios que pueden ser de utilidad para seleccionar algunos programas cuya evaluación se considera prioritaria en la etapa inicial de la Agencia de Evaluación.

## **B) Incentivos para favorecer la evaluación voluntaria**

La mejor forma de lograr la utilización de la evaluación es que los propios responsables expongan sus programas a evaluación. O, todavía mejor, que incorporen la evaluación en el diseño de los programas. Para ello es preciso que perciban claramente que someter los programas a evaluación tiene para ellos más beneficios que perjuicios.

¿Qué consecuencias puede esperar el responsable que se expone a la evaluación? Que la evaluación sirva para mejorar o para paralizar dependerá, en gran medida, de las consecuencias esperadas. Si predomina la percepción de consecuencias de tipo negativo –exigencia de responsabilidad, penalización por las infracciones, crítica política destructiva, etc.–, es probable que los responsables tiendan a eludir la evaluación o la utilicen de forma defensiva para justificar los programas existentes. Desgraciadamente, el proceso político tradicional en el sistema político español ha dado mucho más peso al papel de la crítica y a la exigencia de responsabilidades que a otros incentivos de carácter más positivo. Por tanto, cabe esperar que la oferta de participación en los procesos de evaluación que promueva la Agencia sea recibida inicialmente por los responsables de los programas a evaluar con una comprensible reticencia.

Para que la evaluación sirva de forma efectiva a la mejora de instituciones y políticas debería contribuir a modificar este panorama en el que predominan los incentivos negativos. Es importante reducir el riesgo de castigo e incrementar la probabilidad de recompensa para los responsables políticos que decidan exponerse a un proceso de evaluación de políticas.

Por ejemplo, la exposición de un programa a la evaluación formal de la Agencia podría servir como salvaguarda contra la crítica demagógica. Además, debería ser una oportu-

nidad para el reconocimiento del mérito. Por otra parte, para asentar la cultura de la evaluación sería conveniente que, al menos durante un tiempo, la exposición a la evaluación fuese una opción voluntaria y demostrar que quienes se exponen a ella pueden esperar mejores resultados que quienes no lo hacen.

## 2. Evaluación e innovación institucional

La evaluación de políticas debería conducir a la innovación en la administración y gestión de los recursos públicos, es decir, a sustituir programas poco eficientes por otros relativamente más eficientes, y conseguir mejores resultados con la misma cantidad de recursos públicos. Este proceso será posible en la medida en que la evaluación produzca información empírica fiable y esta información sea utilizada para la toma de decisiones políticas.

En el proceso político tradicional muchas decisiones políticas se toman en función de principios y creencias ideológicas. La introducción de la evaluación en el proceso de toma de decisiones representa un cambio importante porque significa conceder un mayor peso a la evidencia. En último término, la evaluación debería conducir a un proceso de cambio social, de innovación social basada en la evidencia: toda política pública es, de hecho, una teoría del cambio social (Monnier, 1995:82). Para entender apropiadamente el papel de la Agencia de Evaluación como organización promotora de un proceso intencional de mejora institucional, será de ayuda establecer un marco conceptual para entender la innovación social como parte del proceso político (Pinilla Pallejà, 2006).

### 2.1 Estructura y dinámica de las instituciones sociales

Debemos a North la conveniente distinción conceptual entre instituciones y organizaciones sociales (North, 1993). Las instituciones son las reglas del juego y las organizaciones los jugadores. Las instituciones incluyen tanto las leyes formales como las reglas de comportamiento informal y los medios para que se cumplan. Las instituciones definen la estructura de incentivos dentro de la cual se desarrolla el comportamiento social. Así, la Agencia de Evaluación es una organización cuyo propósito es la mejora del funcionamiento de las instituciones (normas, políticas y programas), contribuyendo de ese modo al cambio social positivo. Como agente social, la Agencia será una organización, pero su objetivo consiste en institucionalizar la evaluación, esto es, hacer de la evaluación la norma.

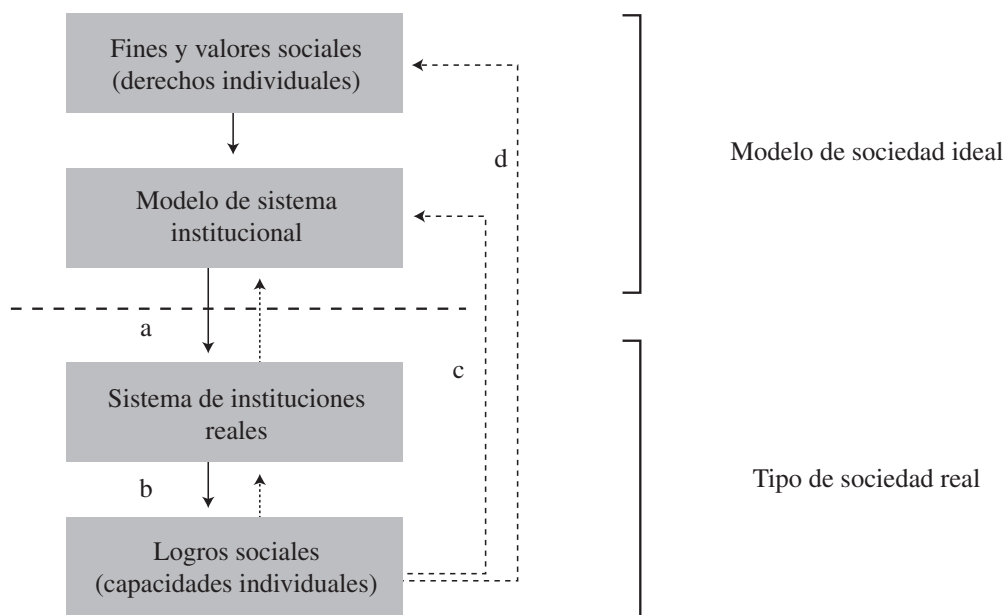
Las instituciones sociales tienen un propósito o una finalidad, de modo que es posible evaluar objetivamente en qué medida logran dicho propósito. Podemos también imaginar y proponer cambios para lograr instituciones más eficaces en el logro de los fines socialmente deseables. Una primera función de la evaluación es comprobar si efectivamente los

cambios propuestos funcionan en la realidad. Para poder llevar a cabo esta evaluación es preciso determinar con la mayor claridad posible cuáles son los fines últimos que la sociedad desea alcanzar con determinada reforma política, mediante qué modelo de sistema institucional se propone hacerlo, a qué sistema de instituciones reales da lugar ese modelo en el lugar y el tiempo concreto en que se lleva a cabo la reforma y cuáles son finalmente sus logros, tanto en relación con los objetivos específicos de la reforma como en relación con otros fines considerados generalmente como socialmente valiosos.

El Gráfico 1 representa en forma de diagrama las relaciones entre lo que se puede denominar un “modelo de sociedad ideal” y un “tipo de sociedad real”. Un “modelo de sociedad ideal” se compone de un conjunto de finalidades y valores sociales ampliamente compartidos y de un modelo institucional básico. Una parte importante del modelo de sociedad ideal de una comunidad política está recogido en la Constitución. Las constituciones incluyen el conjunto de finalidades y valores en forma de declaración de derechos individuales, y el sistema de instituciones y normas básicas con las que la sociedad se dota a sí misma para el logro de los fines sociales deseados. El modelo de sistema institucional deriva, en gran parte, de los fines y valores que la Constitución aspira a lograr, proteger o garantizar.

Sin embargo, las previsiones institucionales requieren adaptaciones al mundo real. En cada lugar y tiempo concretos un mismo modelo institucional ideal puede traducirse en

**Gráfico 1. Estructura institucional ideal y real**



Fuente: Pinilla Pallejà (2006).



un “tipo de sociedad real” diferente. En la práctica, el “tipo de sociedad real” puede alejarse mucho del modelo ideal deseado, debido a las restricciones que impone la propia realidad: los recursos del medio ambiente natural, la eficiencia de la tecnología disponible, la capacidad real de los ciudadanos para ejercer de forma efectiva los derechos que la Constitución les reconoce, etc. El sistema de instituciones reales influirá finalmente en el conjunto de logros sociales que pueden evaluarse en términos de capacidades individuales efectivas.

Se utilizarán los términos efectivo y efectividad como sinónimos de eficacia real, esto es, el grado en que un modelo de sociedad ideal se traduce en la realidad en logros sociales. El criterio de realidad implica que la efectividad de un modelo social para producir logros sociales deseables en un tipo de sociedad real es el criterio determinante para juzgar la “adecuación” de un modelo. Este criterio puede aplicarse al conjunto del modelo y a las instituciones concretas que forman parte de él. Como seres humanos reales que vivimos en una comunidad política nos interesa la evaluación de la efectividad de los distintos modelos para alcanzar metas definidas de desarrollo humano. El diagrama del Gráfico 1 puede ser de utilidad para contextualizar la evaluación de la efectividad de los modelos y sus instituciones. La idea principal es que, en último término, la efectividad ha de evaluarse por los logros sociales alcanzados medidos como capacidades individuales de los ciudadanos.

Es fundamental que los logros sociales alcanzados se midan por las capacidades individuales de los ciudadanos, porque la efectividad de un sistema institucional también depende –a veces de forma crucial– de las capacidades de sus ciudadanos. Cualquier modelo institucional funcionará mejor en una sociedad cuyos ciudadanos gocen de un alto nivel de capacidades, mientras que el mismo modelo tendrá una efectividad menor en las sociedades en las que una parte significativa de la población padezca de limitaciones importantes en sus capacidades básicas. Por otra parte, las capacidades humanas que se consideran básicas en una sociedad concreta tampoco son independientes del sistema institucional. Una sociedad más desarrollada permite y garantiza una mayor libertad individual, pero también exige mayores capacidades básicas para que el ejercicio de esa mayor libertad sea efectivo. Las capacidades también son importantes porque las instituciones van cambiando a lo largo del tiempo de acuerdo con las preferencias expresadas de una u otra forma; por ejemplo, mediante votaciones en procesos electorales o migraciones.

## 2.2 Innovación social intencional

Es evidente que las instituciones que configuran los tipos de sociedad real existentes son manifiestamente mejorables, tanto en eficiencia material como en garantía de capacidades y satisfacción de las necesidades humanas básicas. Pero las instituciones reales son resultado de un proceso, normalmente largo, en el que se han dejado atrás modelos peores. Y, si no tenemos mejores instituciones, es porque resulta difícil avanzar en el logro

de las metas más elementales del desarrollo humano mediante un proceso político basado exclusivamente en creencias. Por ello es importante añadir a ese proceso el conocimiento que proporciona la evidencia científica obtenida a partir de la evaluación rigurosa de los programas y políticas públicas.

En todos los países del mundo existe una diferencia entre el modelo social ideal al que se aspira y los logros del tipo de sociedad real. La corrección de las diferencias entre modelo ideal y tipo real es un imperativo para la política de los países que se presumen democráticos. Denominaremos “innovación social intencional” al proceso de diseño y evaluación de reformas institucionales con el fin de mejorar la efectividad en el logro de los fines ideales socialmente deseables. Este es el fin último de la evaluación de programas y políticas públicas.

La innovación social intencional no debería ser muy diferente de la innovación tecnológica. Al igual que la tecnológica, la innovación social puede tener un gran impacto en la mejora de la calidad de vida de las personas. Pero la innovación social requiere superar un doble filtro, el científico-técnico y el político. Por ello es todavía más necesario llevar a cabo un proceso de innovación social basada en la evidencia (ISBE) que trata de combinar de la mejor manera posible la información objetiva proporcionada por los procedimientos científicos y la intuición y experiencia de gobernantes políticos y legisladores para lograr la doble legitimidad, científica y democrática.

La legitimidad científica exige establecer un sistema objetivo, metodológicamente riguroso, de evaluación de los logros sociales. Por otra parte, la legitimidad político-democrática exige articular un procedimiento reglado de participación social efectiva en el proceso de evaluación, tanto para la selección del diseño como para establecer los indicadores de evaluación de los logros y para la interpretación de los resultados. En definitiva, la innovación social no puede desarrollarse de espaldas a la evolución de los sistemas de valores y creencias de las personas.

Al igual que en los proyectos de I+D (innovación y desarrollo tecnológico), la innovación social basada en la evidencia requiere invertir tiempo y cuantiosos recursos económicos para proporcionar resultados valiosos. Pero también, al igual que en los proyectos de I+D, los resultados pueden ser extraordinariamente rentables a medio y largo plazo. Dado el éxito que los procedimientos científicos han cosechado en el ámbito de la técnica, cabría esperar que la metodología científica se utilizase también de forma rutinaria para diseñar y evaluar proyectos de innovación social. Sin embargo, la experiencia en el uso de metodología científica aplicada a la innovación social es escasa y relativamente reciente. Sólo en las últimas décadas se está abriendo paso la evaluación sistemática de los efectos de las políticas públicas. La creación de la Agencia de Evaluación representa, en este sentido, una gran oportunidad que es preciso aprovechar.

### 3. Hacia la política basada en la evidencia

La idea de mejorar la práctica aprendiendo de los resultados reales y la evidencia científica es la base del método científico. Pero trasladar la evidencia científica a la práctica conlleva una serie de dificultades que es necesario superar. La medicina basada en la evidencia (MBE), que se ha extendido rápidamente en la práctica médica, constituye un modelo de éxito del que es posible aprender si deseamos llevar a la práctica de la política la utilización sistemática de la evaluación en la toma de decisiones.

#### 3.1 El ejemplo de la medicina basada en la evidencia

“La medicina basada en la evidencia (MBE) (...) es la utilización consciente, explícita y juiciosa de la mejor evidencia científica clínica disponible para tomar decisiones sobre el cuidado de los pacientes individuales. La práctica de la medicina basada en la evidencia significa la integración de la maestría clínica individual con las mejores evidencias clínicas externas disponibles” (Sackett *et al.*, 1997). En medicina, la evidencia procede de la investigación científica básica y clínica. A partir de la evidencia clínica, la práctica médica sustituye pruebas diagnósticas y tratamientos previamente aceptados por otros nuevos que son más potentes, exactos, eficaces y seguros. Este proceso de innovación basado en la investigación es lo que caracteriza a la medicina científica. Pero el médico clínico siempre tiene que tomar decisiones sobre pacientes concretos y para esa práctica es también esencial la maestría individual que el profesional de la medicina adquiere con la experiencia. Del mismo modo, la integración de la mejor evidencia evaluativa en la política requerirá siempre de la maestría que el profesional de la política adquiere con la experiencia.

En la MBE, uno de los procedimientos fundamentales es el cribado sistemático de los innumerables artículos publicados en revistas científicas; en esencia, se comparan la metodología experimental utilizada con el “patrón oro” del ensayo clínico. Este cribado ayuda a los profesionales clínicos a mejorar su competencia clínica. En el caso de la PBE este tipo de cribado no sería muy operativo, ya que no existe, ni seguramente puede existir, un patrón semejante de calidad de la investigación empírica en política social y económica. Aun así, ha sido posible poner en marcha procedimientos de revisión sistemática de la literatura sobre ensayos experimentales y cuasi experimentales en política social como la *Campbell Collaboration*, que recopila en Internet una gran cantidad de experiencias en evaluación científica de proyectos de intervención social (Davies y Boruch, 2001).

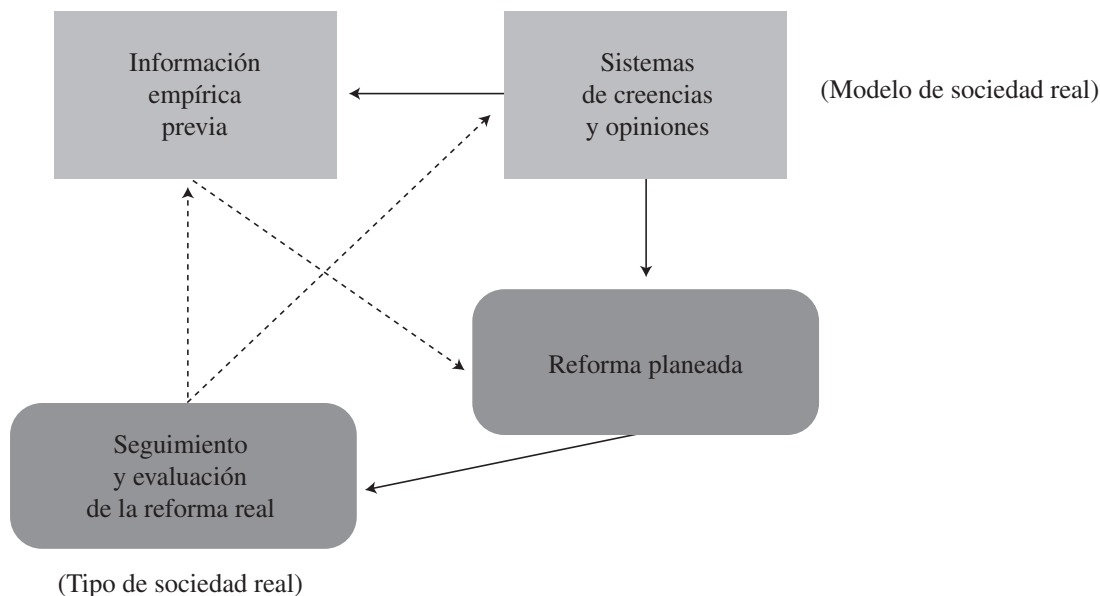
En el caso de la política social y económica lo más importante es ir modificando la práctica de la política basada en ideología, que consiste en evidencia sesgada, en puntos de vista individuales, prejuicios y supuestos ideológicos, para promover, en cambio, una práctica política que busque activamente la mejor evidencia disponible. Al igual que en la MBE, se trata de que los profesionales de la política aprendan a buscar, interpretar y utilizar la evaluación científica de políticas públicas para mejorar su competencia política. Si no existe evidencia fiable procedente de experiencias realizadas previamente, es esencial que el político promueva la evaluación de la efectividad de las normas y programas que se ponen en marcha. Ante la escasez de una evidencia previa fiable, mayor motivo para tratar de asegurar al menos que la práctica actual genere evidencia útil para las decisiones que habrá que tomar en el futuro.

Una contribución esencial de la MBE ha sido ayudar a los profesionales clínicos a formular preguntas de modo que puedan ser respondidas por la evidencia clínica publicada en revistas científicas o generada por nuevas investigaciones relevantes. Del mismo modo, una de las contribuciones de la política basada en la evidencia consiste en asistir a los responsables políticos para que formulen las preguntas relevantes de modo que la evaluación de políticas las pueda responder. Los procedimientos científicos pueden ayudarnos a mejorar un modelo institucional evaluando la efectividad de distintas alternativas programáticas para lograr los fines deseados. La evaluación científica es deseable en la medida en que puede incrementar la tasa de aprendizaje y de innovación social. Para ello es esencial crear un clima favorable a la evaluación rigurosa de los efectos de las políticas públicas con el fin de valorar su efectividad y proponer mejoras. Pero para que la evaluación pueda llevarse a cabo eficazmente, es importante que las leyes y las reformas institucionales se formulen de modo que sean evaluables.

### **3.2 El proceso de la política basada en la evidencia**

El Gráfico 2 representa el proceso característico de la política tradicional, basada en creencias, ideologías y opiniones. En una propuesta de reforma política tradicional el sistema ideológico domina todo el proceso. La información empírica disponible apenas puede influir en las creencias, ya que las evaluaciones y sus resultados tienden a diseñarse e interpretarse para confirmar las ideas propias y justificar las decisiones tomadas en el pasado. Las reformas se planean de acuerdo con las soluciones que prescribe el sistema de creencias y la reforma planeada se lleva a la realidad con mayor o menor éxito. En el caso de que se lleve a cabo un seguimiento y evaluación de los efectos de la reforma, la información obtenida tendrá una utilidad limitada, porque, en el mejor de los casos, sólo se evalúan los objetivos específicos; normalmente no se prevé el contraste de la reforma con otras alternativas posibles, ni se evalúan los posibles efectos colaterales en el conjunto de la sociedad. En consecuencia, la tasa de aprendizaje es baja y con frecuencia se obtiene la impresión de que la evaluación de un programa no ha sido de mucha utilidad.

**Gráfico 2. El proceso de la política basada en creencias y opiniones**

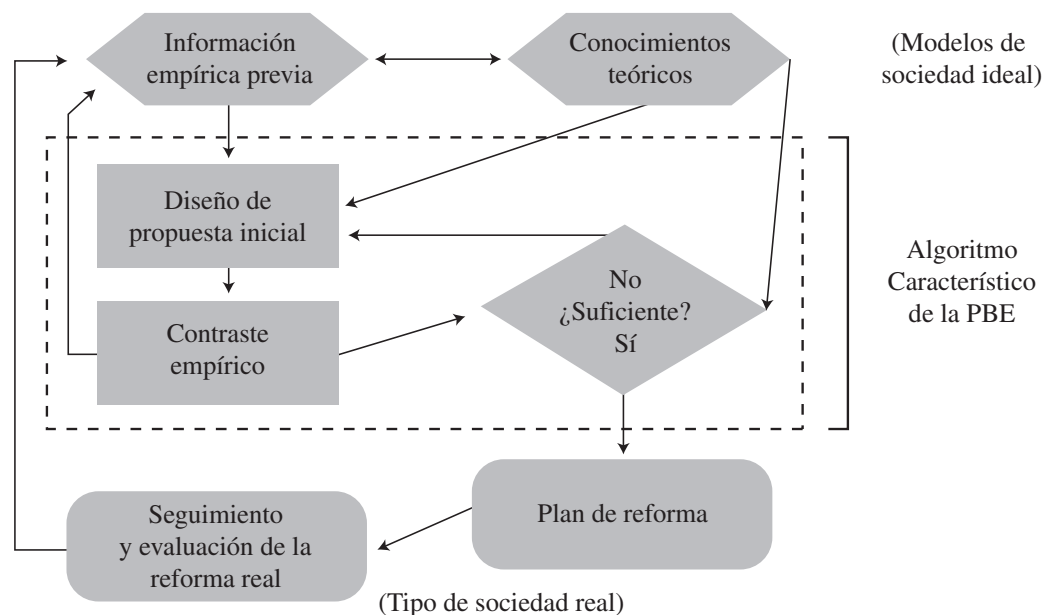


Fuente: Pinilla Pallejà (2006).

En el proceso de política basada en ideología, domina el sentido “descendente”, de arriba hacia abajo. El resultado de las evaluaciones tiende por ello a ser interpretado de forma justificativa, ya que en ningún momento se pretende comparar el programa implantado con otras políticas alternativas. Por ello, el problema de la utilización difícilmente se resolverá con propuestas vagas de promover la participación o formulando un nuevo paradigma de “evaluación pluralista” que trate de adaptar la evaluación al proceso en que realmente se toman las decisiones políticas (Monnier, 1995). Este tipo de propuestas pecan de una cierta ingenuidad, al no tener en cuenta el carácter fuertemente descendente del proceso político tradicional y el desafío que para ese proceso representa una evaluación científica de la efectividad de las políticas.

El Gráfico 3 presenta el proceso característico de la PBE. Con el fin de ordenar el proceso en una secuencia lógica, podemos dividir el proceso de elaboración de una PBE en tres fases: diseño de la propuesta de reforma, contraste empírico del proyecto y evaluación (Davies, 2004). Si se comparan ambos gráficos se puede apreciar que lo característico de la política basada en la evidencia es que estas tres fases, que constituyen la esencia del algoritmo científico, se insertan dentro del proceso político tradicional. En ningún caso lo sustituyen. El algoritmo científico tiene como finalidad esencial el aprendizaje a partir del contraste con la realidad. Sólo añade calidad al proceso de diseño institucional.

**Gráfico 3. El proceso de la política basada en la evidencia**



Fuente: Pinilla Pallejà (2006).

En una **primera fase**, que puede denominarse de **diseño de una propuesta inicial**, se utiliza la información empírica previa y los conocimientos teóricos disponibles para concretar una o más propuestas alternativas a la política existente que se desea reformar. La presencia de alternativas es clave, porque es preciso comparar la propuesta con una alternativa de referencia. La alternativa puede ser no hacer nada, pero es preciso especificar que esa posibilidad es una alternativa política, porque con frecuencia no es factible políticamente no hacer nada si existe preocupación social. En la medida en que el diseño de la propuesta se va concretando se aprecian las carencias informativas y la necesidad de obtener información específica para completar un proyecto de reforma, lo cual permitirá mejorar las herramientas de recogida de información y completar un proyecto de reforma ajustado a la realidad. En esta fase es esencial la participación de responsables políticos para:

- determinar las alternativas políticamente relevantes;
- anticipar qué resultado se considerará un éxito o un fracaso y la información necesaria para establecer ese juicio de valor;
- establecer un consenso amplio sobre los fines deseables de la política que se va a evaluar. Este consenso debería ser amplio, incorporando en la medida de lo posible sensibilidades políticas y sociales distintas de las del Gobierno.

Una vez que el proyecto de reforma parece suficientemente concreto entramos en la **segunda fase**, que podemos denominar de **contraste empírico**. En esta fase se trata de someter a prueba la reforma en una muestra representativa de sujetos –potenciales beneficiarios– para comprobar en qué medida los cambios que se proponen logran los resultados deseados. Eventualmente puede tratarse de un programa piloto o de la fase inicial de implementación de un programa, o de la reforma parcial de un programa ya en marcha en una muestra de beneficiarios. Siempre que sea posible, es preferible utilizar un diseño de tipo experimental en el que los sujetos sean asignados a los grupos de tratamiento y control de forma aleatoria. En esta fase se requieren instrumentos de medida diseñados para valorar la diferencia de los efectos en ambos grupos. Dependiendo de los resultados, la reforma puede ser aprobada por los responsables políticos, puede ser descartada o puede requerir ajustes. Si no es factible llevar a cabo un experimento, lo que sucede con frecuencia en investigación social, siempre se puede recurrir a diseños cuasi experimentales o incluso no experimentales. Pero nunca debería descartarse demasiado rápido la posibilidad de llevar a cabo un verdadero experimento.

Hay que tener en cuenta que los resultados de un experimento social nunca permiten valorar todos los efectos que produciría la reforma en la realidad. Al intervenir sólo sobre una muestra, no es posible obtener información completa de los efectos de conjunto. Por ejemplo, en un experimento de garantía de rentas, cabe esperar que los sujetos del experimento modifiquen su oferta de trabajo, pero como se trata de una muestra reducida, este cambio de comportamiento no producirá ningún efecto en el mercado de trabajo en su conjunto. En cambio, una política de rentas que afecte a toda la población podría tener un efecto global en el mercado de trabajo, ya que al modificar las ofertas individuales de toda la población se modificaría también la oferta agregada de mercado, provocando efectos de conjunto que un experimento no puede detectar. Por este motivo, la evaluación empírica de los efectos de una política social no termina con el experimento, debe continuar con la evaluación de la reforma misma. La fase de contraste empírico resulta, sin embargo, esencial para establecer los beneficios potenciales de una reforma innovadora en comparación a las posibles alternativas, y para establecer con precisión un sistema de indicadores que nos permita valorar de forma rápida y precisa los efectos de la reforma final. Es decir, la fase de contraste contribuye de modo decisivo a la calidad del seguimiento y evaluación final del programa.

Finalmente, y suponiendo que se ha decidido llevar a cabo la reforma, entraríamos en la **tercera fase de seguimiento y evaluación**. Sólo en esta fase podremos valorar exactamente los efectos completos de la reforma en el conjunto de la sociedad, y por ello es conveniente diseñar un plan de evaluación detallado que permita corregir posibles desviaciones sobre lo previsto. En esta fase hay que valorar también los efectos generales de la reforma sobre otros fines sociales valiosos que no son necesariamente los objetivos del programa. Para ello se puede recurrir a distintas estrategias de introducción gradual, de modo que cada paso pueda evaluarse y utilizar la información obtenida para decidir cómo dar el siguiente. Es preciso resaltar que la calidad de la evaluación en la fase de segui-

miento será mucho mayor si se ha llevado a cabo la segunda fase de contraste empírico experimental o cuasi experimental sobre una muestra reducida o en las etapas iniciales de la implantación de la reforma. Una de las funciones esenciales de la fase de contraste es preparar un seguimiento de calidad, porque ayuda a mejorar la calidad de la información, su interpretación inequívoca y la posibilidad de utilizar el sistema de seguimiento como herramienta de retroalimentación que permita detectar y corregir las desviaciones del programa.

En cualquier caso, aunque los procedimientos científicos pueden mejorar la calidad del proceso político, no pueden ni deben sustituirlo. La PBE sólo pretende utilizar los procedimientos científicos para reforzar y mejorar la eficacia en el proceso político democrático. Los responsables políticos siempre tendrán que tomar decisiones asumiendo algún grado de incertidumbre.

En este marco conceptual parece claro que el papel fundamental de la Agencia de Evaluación como organización creada para institucionalizar la evaluación de políticas consiste en incorporar dentro del ciclo de decisión política tradicional el algoritmo característico del método científico, promoviendo, de este modo, la política basada en la evidencia. La finalidad de la PBE es mejorar la tasa de aprendizaje derivada del proceso evaluador para incrementar la efectividad de las instituciones políticas en general. La evaluación de las políticas es perfectamente posible sin recurrir a un proceso de política basada en la evidencia, pero la experiencia internacional nos dice claramente que la utilidad de la evaluación resulta ser escasa si no se demuestra capaz de modificar el modelo tradicional de toma de decisiones políticas, ya que por su carácter “descendente” es refractario a la información procedente de la actividad evaluadora.



## 4. Lecciones de los experimentos sociales

La evaluación mediante diseños de carácter experimental puede considerarse como el “patrón oro” de la evaluación científica. Sin embargo, la aplicación de la experimentación para resolver cuestiones sociales plantea serias limitaciones prácticas y éticas. La experiencia internacional en experimentación social, realizada esencialmente en Norteamérica, resulta muy instructiva para entender tanto las posibilidades como las limitaciones de esta metodología.

La metodología experimental se considera el método más fiable de evaluar el efecto de una política, “el único medio de zanjar las disputas” (Campbell y Stanley, 1966). Sin embargo, la aplicación de los métodos experimentales a la investigación social ha venido acompañada con frecuencia de grandes entusiasmos y profundas decepciones. Así sucedió con los primeros experimentos que trataron de evaluar los efectos de los programas de impuesto negativo sobre la renta en EE UU en los años setenta. Después de años de investigación, y de gastar más de 100 millones de dólares, los resultados llegaron cuando ya no interesaban políticamente, y tanto los defensores del impuesto negativo como sus detractores interpretaron los resultados como un apoyo a sus tesis, de modo que el aparatoso despliegue experimental no sirvió realmente para zanjar ninguna disputa.

El fracaso inicial de los primeros experimentos sociales en EE UU se ha esgrimido reiteradamente por algunos de los más reconocidos expertos europeos en evaluación para descartar la aplicabilidad del paradigma experimental en la evaluación de programas y políticas públicas en nuestro continente (Monnier, 1995; Vedung, 1997). Sin embargo, este descarte fue probablemente prematuro. En EE UU, el fracaso de las primeras evaluaciones experimentales no llevó a los investigadores a renunciar al método, sino a aprender de los errores cometidos imaginando formas mucho más económicas y prácticas de realizar la evaluación experimental, de forma que proporcionaran respuestas útiles a cuestiones decisivas. En esta labor de demostración de la factibilidad y utilidad de la metodología experimental para la evaluación de programas públicos ha tenido un papel destacado la MDRC (*Manpower Demonstration Research Corporation*), una organización sin fines de lucro que ha organizado y dirigido más de 30 estudios de evaluación en los que han participado más de 300.000 personas, insistiendo siempre en la asignación aleatoria como un elemento clave para poder establecer sin sesgos el efecto neto de los programas evaluados (Gueron, 2000:1).

La perseverancia de la MDRC y los científicos sociales americanos en la defensa del método experimental ha sido finalmente reconocida por la Administración americana. Tanto

el Departamento de Trabajo como el de Salud y Asuntos Sociales han hecho suyas las recomendaciones de dos paneles de expertos de la Academia Nacional de las Ciencias, que concluían que “la asignación aleatoria es la aproximación más fiable para determinar el efecto neto de los programas sociales” (Gueron, 2000:5). Hoy día, las peticiones de financiación de programas públicos a favor del empleo y la formación deben ir acompañadas de una evaluación rigurosa, preferentemente mediante la utilización de técnicas experimentales. Se ha reconocido de este modo que “la credibilidad y utilidad de una evaluación dependen en gran medida del método empleado, y se admite que para toda evaluación importante conviene proceder con el mayor rigor, si no se quiere despilfarrar recursos, tiempo y energía, y concluir un estudio que, en lugar de aportar elementos para la adopción de decisiones, sólo produzca debates sin fin sobre cuestiones de método insolubles” (Gueron, 1992:299). En general, “de las observaciones recogidas en diferentes países, resulta que los estudios de demostración que utilizan selección aleatoria figuran entre los medios de evaluación más eficaces” (OCDE, 1992:241).

Conviene recordar lo realmente importante para mejorar la racionalidad de la toma de decisiones: el **efecto neto** –qué cantidad de resultado podemos atribuir al programa en comparación con el no programa– y el **efecto relativo** –consiste en la comparación de un programa con un programa alternativo–. La ventaja decisiva del método experimental consiste en comparar dos o más grupos experimentales, a los que se aplica uno o varios programas alternativos cuyo efecto se desea evaluar, con un grupo control al que no se aplica el programa y que sirve de contrafáctico, esto es, trata de mostrar qué hubiera pasado si el programa no se hubiera aplicado al grupo experimental. Para garantizar que todos los grupos son comparables se asigna los sujetos a los grupos mediante un procedimiento aleatorio. El procedimiento aleatorio es esencial porque garantiza que todos los sujetos han tenido la misma probabilidad de recibir cualquiera de los tratamientos –o el no tratamiento– y, por tanto, las diferencias encontradas pueden atribuirse al efecto del programa evaluado y no a otras variables extrañas.

Dado que no siempre es posible llevar a cabo experimentos, por razones de costes, de imposibilidad técnica, o por motivos éticos, se han desarrollado métodos no experimentales que tratan de crear grupos de comparación de características similares utilizando distintos procedimientos estadísticos para controlar los posibles sesgos (por ejemplo, el método de emparejamiento selecciona para cada sujeto del grupo experimental un sujeto control, asegurando que sea similar para toda una serie de características observables, construyendo así un contrafáctico de comparación). Estos métodos pueden ser de utilidad cuando no es factible o razonable utilizar procedimientos de asignación aleatoria y también para proporcionar información complementaria que ayude a explicar las razones del efecto de un programa (Riccio y Bloom, 2001). Pero cuando se ha evaluado un mismo programa con métodos experimentales y no experimentales, las diferencias encontradas sugieren que los métodos no experimentales no son lo bastante buenos como para sustituir a los experimentales cuando estos últimos son posibles (Bloom *et al.*, 2002).

Lo que podemos y debemos aprender de esta experiencia es que merece la pena hacer un esfuerzo para llevar a cabo evaluaciones con el mayor rigor metodológico. El método experimental ha de mantenerse como la máxima aspiración, pero no cabe entusiasmarse con sus posibilidades. La experimentación social es difícil de realizar y la información que proporciona, aunque fiable, es limitada. La política puede apoyarse en la ciencia, pero la ciencia nunca podrá dirigir la política. Las sociedades humanas son demasiado complejas y el futuro es intrínsecamente incierto. Por ello, en el siguiente capítulo se abordará cómo incorporar del mejor modo posible la evaluación científica al proceso político.

## 5. La búsqueda de estrategias políticas robustas

Gran parte de las reticencias políticas hacia la experimentación social se debe a los excesos de optimismo de los investigadores sociales americanos sobre las posibilidades de llevar a cabo una política racional guiada por el conocimiento científico. Además de parecerse a una propuesta inaceptable de sustitución de política democrática por dirigismo científico, se trataba de una expectativa totalmente fuera de la realidad.

La complejidad de las sociedades humanas y la incertidumbre esencial sobre el futuro obligan a reubicar la metodología científica en una posición mucho más humilde. Dado que la realidad social es esencialmente impredecible y está sujeta a profunda incertidumbre, científicos sociales y políticos necesitamos colaborar estrechamente para buscar estrategias robustas, renunciando a la búsqueda de una utopía social científicamente óptima que, ni existe, ni puede existir.

Si queremos incorporar en el proceso de innovación a los responsables políticos, además de tener en cuenta las limitaciones de los métodos experimentales en el campo de las ciencias sociales, es necesario tener en cuenta las peculiares dificultades de la toma de decisiones políticas: elevado grado de incertidumbre sobre los sucesos futuros, escaso margen de maniobra para la modificación de las políticas, brevedad del ciclo político en contraste con la necesidad de tomar decisiones que sólo pueden ser eficaces a largo plazo, etc. Para enfrentar con éxito este tipo de dificultades adaptaremos al campo de la innovación social una original metodología de planificación estratégica desarrollada por expertos de la RAND Corporation (Lempert *et al.*, 2003). Se trata de una metodología muy apropiada porque propone un procedimiento eficaz de colaboración sistemática entre científicos y responsables políticos, reconociendo que esta colaboración es esencial para encontrar soluciones efectivas a algunos de los grandes problemas de la humanidad.

Lempert y colaboradores parten del siguiente dilema: cuando pensamos en buscar soluciones a los grandes problemas de la humanidad, parece evidente que no pueden resolverse de hoy para mañana, que se necesitará bastante tiempo, pero si queremos que ese tipo de problemas se resuelvan alguna vez, tenemos que empezar a actuar hoy. El resultado a medio y largo plazo de lo que hagamos hoy determinará un futuro mejor o un futuro peor. El problema es que el ciclo político es mucho más corto, y nadie puede predecir con certeza el resultado a largo plazo de nuestras acciones de hoy, porque hay infinidad de sucesos impredecibles que pueden modificar el resultado final. ¿Cómo elegir entonces las estrategias más apropiadas para que los grandes problemas encuentren algún día solución?

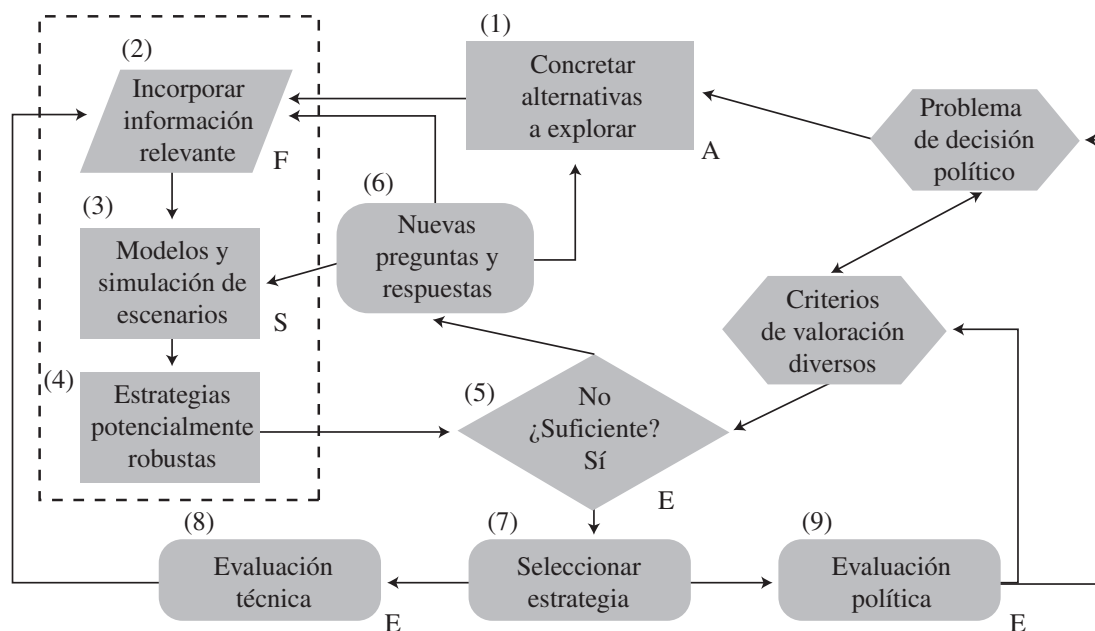
La idea que proponen Lempert, Popper y Bankes es sencilla: dado que no sabemos lo que va a pasar, busquemos estrategias robustas, esto es, estrategias que puedan dar un resultado razonablemente bueno “pase lo que pase” (Lempert *et al.*, 2003). Y para ello proponen un procedimiento de análisis de estrategias políticas que parece muy apropiado para poner en práctica la PBE.

## 5.1 Decisiones bajo incertidumbre

La evaluación científica ayuda a tomar decisiones reduciendo la incertidumbre, pero no la puede eliminar. Podemos estar bastante seguros de que siempre habrá problemas humanos que requerirán tomar decisiones bajo condiciones de profunda incertidumbre. Problemas en los que “el analista no sabe, o las distintas personas que tienen que tomar parte en la decisión no pueden ponerse de acuerdo en: a) cuál es el modelo apropiado para describir las relaciones entre los factores causales relevantes; b) las distribuciones de probabilidad utilizadas para representar la incertidumbre sobre variables o parámetros clave; y/o c) cómo valorar la deseabilidad de los resultados de las distintas alternativas” (Lempert *et al.*, 2003). La mejor respuesta a la incertidumbre profunda será una estrategia que, en vez de ser óptima para un futuro particular que no podemos predecir, pueda funcionar bien en gran cantidad de futuros plausibles. Y es más probable que una estrategia pueda funcionar bien en una variedad de futuros diferentes si tiene la capacidad de adaptarse a lo largo del tiempo a medida que se dispone de nueva información. Por ello, Lempert, Popper y Bankes proponen una metodología de exploración sistemática de escenarios futuros que incluye cuatro elementos clave: a) considerar grandes conjuntos de escenarios; b) buscar estrategias robustas en vez de estrategias óptimas; c) lograr la robustez mediante adaptabilidad; y d) explorar de forma iterativa una multiplicidad de futuros plausibles.

Estos elementos pueden articularse a través de un proceso iterativo en el que los ordenadores ayudan a crear un conjunto amplio de escenarios plausibles. La meta es descubrir alternativas de acción a corto plazo (políticas) que sean robustas frente a un amplio rango de futuros posibles al evaluarlas con una variedad de criterios de valor. Se trata, por tanto, de una infraestructura analítica adecuada para el debate público de propuestas de reforma cuya repercusión social puede ser grande, pero difícil de predecir. La meta de un proceso de decisión robusta es alcanzar el consenso sobre un curso de acción a corto plazo entre una variedad de partidos o grupos de interés. Los modelos de simulación por ordenador juegan un papel importante en la búsqueda de alternativas robustas. Las nuevas tecnologías de la información hacen posible la colaboración entre intuición humana y simulación por ordenador para encontrar estrategias robustas. Los ordenadores pueden utilizarse para manejar, mostrar y resumir grandes cantidades de información y de ese modo facilitar la exploración sistemática de un amplio conjunto de futuros plausibles. La meta a alcanzar es el logro de un consenso sobre las alternativas de corto plazo que pueden dar forma a un futuro humanamente deseable.

**Gráfico 4. Procedimiento de análisis de robustez**



Fuente: Pinilla Pallejà (2006).

El Gráfico 4 ilustra el procedimiento iterativo para analizar la robustez. Ante un problema de decisión política debido a la incertidumbre o a la existencia de criterios de valoración diversos se toma la iniciativa de proceder al análisis de robustez. El primer paso (1) consiste en la concreción de las estrategias alternativas a corto plazo a explorar. Una vez definidas las alternativas es preciso recoger la información relevante para el análisis (2) e incorporarla a los modelos de simulación para generar una gran cantidad de escenarios (3). El comportamiento de las distintas estrategias en el conjunto de escenarios puede definir –o no– alguna estrategia potencialmente robusta (4). El paso decisivo del proceso (5) consiste en tomar la decisión de si se ha alcanzado una solución suficientemente satisfactoria o no. Si los distintos agentes políticos que intervienen en el proceso no llegan a un acuerdo, se reitera el proceso procediendo a generar nuevas preguntas (6), solicitando la incorporación de nueva información, o sometiendo las estrategias alternativas a nuevos sucesos futuros posibles que habrán de incorporarse a los modelos de simulación. El proceso puede reiterarse hasta alcanzar un acuerdo en una estrategia que resulta satisfactoria (7), esto es, la selección de una estrategia que se considera suficientemente robusta. En cualquier caso, las acciones que se lleven a cabo –que constituyen la estrategia elegida– deberán ser evaluadas tanto técnica (8) como políticamente (9), ya que los resultados de la evaluación serán útiles en el análisis de estrategias en el futuro.

Por su naturaleza abierta, los métodos de decisión robusta no pueden demostrar que una estrategia A funcionará siempre mejor que una estrategia B en todos los futuros posibles.

La robustez es un argumento persuasivo basado en la comparación del desempeño de las diferentes estrategias a lo largo de una muestra grande de escenarios. No hay ninguna garantía de que este método produzca un resultado óptimo, del mismo modo que un experimento científico no proporciona una respuesta definitiva. La meta del procedimiento propuesto es facilitar el consenso político en circunstancias de profunda incertidumbre.

## 5.2 Estructura FASE

Quisiera resaltar en el Gráfico 4 los elementos clave que caracterizan el proceso de decisión. Estos elementos se pueden organizar sintéticamente en la estructura FASE. La F representa la información sobre los factores externos que cabe esperar que influyan en los resultados, es decir, los factores relevantes para las alternativas a valorar, pero que los gobernantes no pueden modificar directamente. Para que la simulación sea fiable es preciso incorporar la información disponible sobre todos los factores relevantes. La A representa el conjunto de las acciones variables que los gobernantes podrían llevar a cabo. Cada combinación de acciones posibles constituirá una estrategia política alternativa. En la medida en que siempre hay muchas acciones posibles que pueden combinarse de distintos modos, la variedad de estrategias alternativas puede ser muy grande, y los políticos deben decidir cuáles son relevantes. La S señala el momento de la simulación. Para la simulación es necesario especificar las relaciones entre las variables F y A. Finalmente, la E marca los puntos del proceso en los que se lleva a cabo la evaluación de los resultados de la simulación y de la estrategia política, una vez que se lleva a cabo. La evaluación tiene gran importancia. Distintos tipos de evaluación son necesarios en distintas etapas del proceso.

El proceso de decisión robusta requiere que un equipo de proyecto revise y organice la información relevante. Es útil para ello agrupar los elementos clave del análisis en el cuadro que hemos denominado estructura FASE, y que no es más que una adaptación libre al castellano de la estructura XLMR creada por Lember (Lempert *et al.*, 2003). Así, el equipo que va a explorar un campo de decisión política controvertido debe enumerar y clasificar los elementos que va a considerar de acuerdo con la estructura siguiente:

- **F (factores exógenos):** información sobre los factores fuera del control de los responsables políticos, pero que pueden ser importantes para determinar el éxito de las estrategias.
- **A (acciones a corto plazo):** listado de las acciones que los responsables políticos podrían llevar a cabo en un plazo de tiempo corto. Las estrategias alternativas a explorar deberían poder expresarse como una combinación de estas acciones.
- **S (simulación de las relaciones):** incluye las relaciones de causalidad e interacción entre los distintos factores externos y las acciones, así como los efectos sobre los resultados. Las relaciones deben especificarse en forma de ecuaciones dinámicas entre

variables simples (tipo F, A o E). La fidelidad con la que las ecuaciones reflejen las verdaderas relaciones de causalidad será determinante para la calidad de la simulación. Una configuración particular de S y F representa un estado del mundo.

- **E (evaluación de los resultados):** son los indicadores de resultado que los responsables políticos y las personas afectadas por las estrategias utilizarán o tendrán en cuenta para valorar los resultados de cada alternativa.

Los tres conjuntos de elementos variables (F, A y E) están ligados entre sí por el cuarto conjunto (S), que representa las relaciones entre las variables. Para llevar a cabo el proceso de simulación puede utilizarse un programa de simulación dinámica, que es un tipo de *software* que genera un escenario resultado (E) a partir de la información de entrada (F y A), utilizando el sistema de ecuaciones especificado (S) (López Díaz-Delgado y Martínez Vicente, 2000). La especificación de las ecuaciones del modelo de simulación es una cuestión clave para la que no siempre existe base empírica cuantitativa suficiente al iniciar la tarea. Una parte importante del trabajo de modelado consiste en ajustar las ecuaciones del modelo a la mejor información disponible. En cualquier caso, los propios programas de simulación permiten realizar un análisis de sensibilidad de los resultados a variaciones en la especificación de las ecuaciones, por lo que las carencias informativas no son un obstáculo insalvable.

Para explorar eficazmente múltiples escenarios se puede utilizar un explorador de escenarios, que es un tipo de *software* que analiza de forma sistemática el impacto potencial sobre los resultados (E) de todas las posibles combinaciones de acciones (A) dentro de un rango definido. El explorador de escenarios puede ser programado para descartar directamente las acciones que producen resultados inaceptables de acuerdo con criterios de decisión previamente establecidos. Las alternativas aceptables se pueden mostrar conjuntamente en pantalla ofreciendo un panorama gráfico de los resultados esperados ante una diversidad de escenarios futuros.

El proceso se inicia con los datos y el conocimiento sobre el funcionamiento del mundo real, pero es una metodología para el aprendizaje que permite incorporar nueva información y mejorar la capacidad de intervenir sobre la realidad. De este modo, una propuesta de política representa tan sólo un punto de partida para la innovación social. A partir de una propuesta que puede funcionar, el procedimiento ayuda a imaginar y buscar políticas alternativas rivales que mejoren el desempeño potencial de la propuesta de partida. En este proceso de creatividad, es esencial la participación de una amplia representación de los grupos sociales afectados por la reforma, incluidos los grupos más críticos con la propuesta.

Una de las ventajas del análisis de la robustez es la sencillez del concepto. En condiciones de incertidumbre profunda, los responsables políticos deberían identificar acciones a corto plazo que sean robustas en un amplio rango de futuros plausibles buscando información que desafíe y mejore la robustez de las decisiones. Las instituciones democráticas



son un ejemplo de robustez por su capacidad de adaptación. La capacidad de adaptación al cambio contribuye a la robustez y constituye una ventaja evolutiva decisiva.

El análisis de la robustez puede considerarse como una modalidad de evaluación para facilitar la selección de políticas relevantes y, como cualquier otro tipo de evaluación, un reto esencial de la búsqueda de estrategias robustas es la incorporación de este método al proceso político corriente.

En este contexto es fácil comprender que una decisión política que casi siempre será robusta es invertir en sistemas de evaluación rigurosa. La creación de Agencia para la institucionalización de la evaluación representa en este sentido una acción robusta y una gran oportunidad para la promoción de la innovación social basada en la evidencia.

### **5.3 Institucionalización de la evaluación de políticas**

La institucionalización de la evaluación puede considerarse como una herramienta para mejorar la capacidad de adaptación de los sistemas democráticos. Los sistemas democráticos son capaces de adaptarse a circunstancias imprevistas, pero su capacidad de adaptación depende de la capacidad de elección de los ciudadanos. El funcionamiento eficaz de las instituciones democráticas requiere también ciudadanos más capacitados y mejor informados sobre la efectividad de los programas públicos. Por ello, en la medida en que asegure la capacidad de los ciudadanos para juzgar el valor de las políticas, la robustez de las instituciones democráticas será mayor.

Por el contrario, si las instituciones democráticas no protegen eficazmente la formación y el mantenimiento de las capacidades de los ciudadanos, ellas mismas pueden verse debilitadas y amenazadas. Por ello, es importante institucionalizar sistemas de evaluación de los objetivos generales de la sociedad en términos de capacidades humanas. Y también incorporar medidas de evaluación del proceso y los resultados de políticas concretas, esto es, de los objetivos específicos del programa. Una estrategia política será en general más robusta si mejora la capacidad de adaptación del sistema institucional.

La institucionalización de la evaluación por medio de la Agencia de Evaluación debería tratar de asegurar la utilización del análisis de robustez en las fases previas de diseño de la evaluación. El análisis de robustez puede ser de gran ayuda para seleccionar un número reducido de estrategias potencialmente robustas como programas candidatos para el contraste empírico –de forma experimental o mediante la implantación gradual–. Pero, además, los modelos de simulación creados para explorar la plausibilidad de distintos escenarios futuros facilitan mucho las fases de evaluación posterior, al poner claramente de manifiesto las carencias y necesidades informativas y orientar sobre los diseños más rigurosos para la evaluación.

Ya se comentó anteriormente que en la fase de contraste debía tratarse de emplear la metodología científica más rigurosa, si es posible recurriendo a la metodología experimental con asignación aleatoria de los sujetos a los grupos de tratamiento y control, con el fin de obtener resultados fiables para la toma de decisiones.

Finalmente, en la fase de seguimiento y evaluación final de un programa candidato se debería institucionalizar tanto la evaluación del impacto global –impacto general del programa en términos de calidad de vida y capacidades individuales de la población en general–, como la evaluación del impacto específico –efectividad del programa para alcanzar sus objetivos específicos, entendiendo por objetivos específicos los que justificaban la creación y puesta en marcha del programa–. Para la evaluación del impacto global es preciso normalizar un sistema general de información para la evaluación que debería aplicarse de forma sistemática a todos los programas públicos.

## 6. Recomendaciones para la puesta en marcha de la Agencia de Evaluación

Se ha iniciado este documento haciéndose eco del reto que se le plantea a la Agencia de Evaluación: demostrar su propia utilidad como organización, logrando, en un periodo de tiempo relativamente breve, introducir la evaluación en el proceso normal de decisión política, esto es, la institucionalización de la evaluación. El reto es en realidad doble. Por un lado, consiste en **demostrar capacidad científica** para llevar a cabo evaluaciones con el mayor rigor metodológico y, en segundo lugar, **lograr la participación activa** de los responsables políticos de modo que los resultados de las evaluaciones se utilicen para la toma de decisiones y la mejora de los programas.

En definitiva, se trata de promover un proceso de política basada en la evidencia, por lo que se trata de un desafío ambicioso.

En cuanto al rigor metodológico, la Administración española tiene cierta experiencia en la evaluación de la calidad de los servicios, pero el panorama de evaluación de la efectividad de las políticas públicas es muy insuficiente (Garde Roca, 2005). Por otra parte, en España el debate político parece instalado en la confrontación, con un elevado tono emocional y exceso de crítica poco constructiva. De hecho, el comité de expertos recomienda que la Agencia dependa del Gobierno mejor que del Parlamento, porque entiende que la dependencia del Parlamento se podría instrumentalizar como plataforma de crítica indiscriminada al Gobierno (Garde Roca, 2005). Podría decirse que este no es el ambiente más adecuado para favorecer un consenso amplio de ningún tipo, pero tal vez por esto mismo el recurso a la evaluación científica como medio de fomentar la política basada en la evidencia sea más necesario que nunca.

A pesar del ambiente de confrontación, nos encontramos en una coyuntura en la que parece existir un consenso general en el valor social de los fines del Estado del bienestar, pero también en la necesidad de renovar algunas de sus instituciones (Garde Roca, 1999; Pinilla Pallejà, 2006). Donde no hay acuerdo, sino profundas discrepancias, es en qué reformas pueden ser las más adecuadas para conseguir los efectos deseados. Parece evidente que deberíamos aspirar a que las reformas a introducir sean técnicamente correctas y políticamente aceptables por la mayoría, pero los conocimientos técnicos y los procedimientos democráticos parecen discurrir con frecuencia por cauces separados.

La metodología de la política basada en la evidencia trata de superar este distanciamiento proponiendo una forma de articular la plausibilidad científica con la aceptabilidad política de las propuestas de reforma. Justamente en una situación como la actual, en la que parece haber un notable consenso en los fines del Estado del bienestar y en la necesidad de introducir reformas, pero existen profundos desacuerdos en las reformas a introducir, la evaluación científica puede revelarse como extraordinariamente útil en la medida en que los evaluadores sean capaces de elaborar un diseño experimental o cuasi experimental que permita contrastar las principales políticas alternativas y cuyos resultados puedan ser aceptados por las principales fuerzas políticas. Por tanto, la evaluación de políticas públicas, más que la evaluación de la calidad de los servicios, constituye la gran oportunidad para la Agencia de Evaluación.

La Agencia debe ser especialmente cuidadosa en la primera etapa –tres o cuatro primeros años de funcionamiento–, ya que en función de lo que haga, y de lo que no haga, creará un estilo de funcionar que determinará en gran parte su futuro. La Agencia debe intentar ganar el reconocimiento de las distintas administraciones españolas demostrando su capacidad para llevar a cabo la evaluación de políticas relevantes para que los resultados de las evaluaciones sean útiles para la toma de decisiones y la mejora de normas e instituciones. Con este fin debería diseñar un plan estratégico que asegure que una parte importante de su actividad se concentre en la evaluación de políticas y sirva de guía para transitar la tortuosa senda que conduce desde la política basada en la ideología a la PBE.

Durante la etapa inicial de la Agencia la evaluación de programas existentes con anterioridad debería limitarse a aquellos programas cuyos responsables deseen y soliciten voluntariamente ser evaluados. La Agencia debería seleccionar sólo algunos de estos programas de acuerdo con los criterios que se especifican más abajo.

Para que surja un número suficiente de programas voluntarios y poder elegir entre ellos, conviene crear un entorno favorable a la evaluación en el que el beneficio de ser evaluado exceda al riesgo de exigencia política de responsabilidades. Por ejemplo, el Gobierno podría incentivar la participación voluntaria en los procesos de evaluación, restringiendo la posibilidad de llevar a cabo incrementos presupuestarios en función del compromiso adquirido por cada programa para someterse a un proceso de evaluación rigurosa supervisado por la Agencia.

En el caso de los programas nuevos, sería deseable que todos ellos incorporen un plan de evaluación desde el diseño. La Agencia podría actuar como simple supervisora de la evaluación prevista, pero también en algunos casos podría diseñar y llevar a cabo la evaluación del programa. Para ello, la Agencia debe ser capaz de supervisar todas las nuevas políticas o programas, ya que si no fuese así la exigencia de incorporar un sistema de evaluación en todos los programas nuevos podría convertirse en un elemento retardatario, más que en un incentivo a la innovación social basada en la evidencia.

## 6.1 Elementos para un plan estratégico de transición a la innovación social basada en la evidencia

- Ofrecer los **servicios de evaluación** de la Agencia a todas las administraciones en unas condiciones que los responsables de los programas puedan considerar como especialmente convenientes. Es importante que se reciban suficientes peticiones de evaluación como para poder seleccionar entre ellas las más favorables para la Agencia de acuerdo con un conjunto de criterios explícito.
- En cuanto a **los criterios para seleccionar** qué políticas o programas merecen ser evaluados directamente por la Agencia, hay que señalar que inicialmente no se puede evaluar todo y, por tanto, es importante contar con la colaboración e iniciativa de los responsables. Posteriormente se evaluará a los que se presenten voluntarios, seleccionando un conjunto de programas convenientes para la Agencia. Entre los criterios a valorar para seleccionar los programas a evaluar por la Agencia podríamos citar los siguientes:
  - impacto presupuestario potencial o social importante;
  - visibilidad política;
  - carácter innovador y con efectos en parte desconocidos o poco predecibles;
  - probabilidad de que provoquen efectos en distintos sectores multisectoriales;
  - metodológicamente emblemática; posibilidad de evaluación experimental;
  - participación en el programa de las CC AA;
  - posibilidad de ofrecer resultados decisivos en menos de cuatro años;
  - existencia de consenso político y social en los objetivos del programa.
- Aceptar un compromiso para elaborar un **sistema de información** para la evaluación general en función de los logros entendidos como capacidades humanas individuales de los ciudadanos pertenecientes a la comunidad política. Debería considerarse una prioridad el desarrollo de un sistema que permita el seguimiento y evaluación de los logros de las instituciones, en su conjunto, y el impacto de cada una de ellas en la mejora global.
- Dar prioridad a la inclusión de **metodología experimental** en las evaluaciones cuando se cumplan unos criterios mínimos que la hagan factible: carácter innovador, elevado volumen de gasto, implantación gradual –inicialmente se puede sortear el derecho a recibir el programa–, etc. La metodología experimental confiere mayor calidad a todo el

proceso, en especial el seguimiento y evaluación *ex post* convencionales. Utilizar un contraste experimental incrementa sustancialmente la calidad de la información, la capacidad de interpretación y, consecuentemente, el valor del seguimiento como guía para la toma de decisiones –retroalimentación o *feedback*–.

- Implantar y poner en práctica el **algoritmo de análisis de la robustez**, muy valioso para la evaluación previa y la determinación de los objetivos a evaluar en un diseño experimental. Permite evaluar políticas en un contexto de análisis de impacto global, más allá de un programa aislado. Se trata de un enfoque analítico y sistémico al mismo tiempo. Incorpora los resultados del análisis a un procedimiento de síntesis sistémico y multidisciplinar.
- Aprovechar las evaluaciones más emblemáticas para llevar a cabo **investigación evaluativa**. Por ejemplo, contrastar las similitudes y diferencias del método experimental con otros métodos cuasi experimentales o no experimentales que pueden llevarse a cabo en paralelo.
- Establecer **relaciones de colaboración** con instituciones nacionales e internacionales con experiencia en evaluación y capacidad de investigación evaluativa: universidades, institutos de investigación, MDRC, etc.
- Lógicamente, la Agencia de Evaluación debería predicar la evaluación con el ejemplo, esto es, **sometiéndose a evaluación sus propias actividades**. La Agencia debe demostrar que la inversión en evaluación es efectiva para mejorar el rendimiento de los programas públicos y que el beneficio social logrado con los procesos de evaluación supera generalmente los costes<sup>1</sup>.
- Prever **la difusión y divulgación pública del trabajo de evaluación**, en especial las evaluaciones demostrativas de la efectividad para mejorar las políticas públicas.

## 6.2 A modo de conclusión

La Agencia de Evaluación pretende institucionalizar la evaluación de políticas públicas. La incorporación de la evaluación sistemática de las políticas al proceso político tradicional representa un cambio mayor de lo que suele creerse, puesto que implica modificarlo

---

<sup>1</sup> Demostrar esto no será difícil, porque, aunque la evaluación puede ser costosa, el coste puede ser una proporción muy pequeña del presupuesto del programa. Por ejemplo, la Encuesta de Pobreza y Desigualdades del Gobierno Vasco (EPDS) sólo representa el 0,1% del presupuesto del programa de rentas mínimas de la Comunidad Autónoma Vasca (Gobierno vasco, 2000).

para convertirlo gradualmente en un proceso de PBE. La PBE no significa sustituir a los representantes políticos en el proceso de decisión política, sino de incorporar a ese procedimiento la búsqueda activa de la mejor evidencia científica sobre la efectividad de las políticas públicas.

Es comprensible que el estamento político reciba esta propuesta con cierta reticencia, pero los mejores profesionales de la política comprenderán que los procedimientos de la PBE permiten elaborar, mediante el consenso político fundado en la evaluación empírica, propuestas técnicamente correctas y políticamente aceptables para cualquier futuro plausible.

Para poner en marcha este método se requiere, ante todo, la voluntad de mejorar la calidad de nuestras decisiones políticas. La voluntad política debería traducirse, por un lado, en esfuerzos concretos por mejorar los sistemas estadísticos que suministran la información para cargar los ordenadores en mayor esfuerzo para investigar las relaciones entre las variables relevantes, por otro, y, por último, en un compromiso para diseñar las políticas públicas, estableciendo objetivos operativos mensurables y evaluables.

## Índice de Gráficos

### Gráficos

Gráfico 1. Estructura institucional ideal y real .....	16
Gráfico 2. El proceso de la política basada en creencias y opiniones .....	21
Gráfico 3. El proceso de la política basada en la evidencia .....	22
Gráfico 4. Procedimiento de análisis de robustez .....	30



## Bibliografía

- Alvira Martín, F. (1991), Metodología de la evaluación de programas, CIS, Madrid.
- Bloom, H. S., *et al.* (2002), Can Non experimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs? MDRC, New York. <http://www.mdrc.org>.
- Campbell, D., y Stanley, J. (1966), Experimental and Quasi-Experimental Designs for Research, Chicago, Rand McNally & Company. Ed en Castellano: Diseños experimentales y cuasiexperimentales en la investigación social, Amorrortu ed., Buenos Aires.
- Davies, P. T. (2004), Is Evidence-Based Government Possible?, Jerry Lee Lecture 2004. Presented at the 4th Campbell Collaboration Colloquium, Washington D.C. (19 February), <http://www.policyhub.gov.uk/home/JerryLeeLecture1202041.pdf>.
- Davies, P. T., y Boruch, R. (2001), The Campbell Collaboration, *British Medical Journal*, 323:294-5.
- Garde Roca, J. A. (1999), Políticas sociales y Estado del Bienestar en España, Informe 1999, Ed. Trotta-FUHEM, Madrid.
- Garde Roca, J. A. (coord.) (2005), La Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas. Reflexiones y propuesta de creación. Informe Comisión Expertos, MAP, Madrid.
- Gobierno Vasco (2000), Encuesta de Pobreza y Desigualdades Sociales, 2000. Departamento de Justicia, Trabajo y Seguridad Social, Servicio de Estudios y Régimen Jurídico, Vitoria-Gasteiz.
- Gueron, J. M. (1992), Evaluación y acción: los programas para el empleo y la formación en los Estados Unidos, en OCDE (1992), La evaluación de programas de empleo y de medidas sociales, Ministerio de Trabajo y Seguridad Social, Madrid.
- Gueron, J. M. (2000), The Politics of Random Assignment. Implementing Studies and Impacting Policy, MDRC, <http://www.mdrc.org/publications/45/workpaper.html>.
- Lempert, R. J., Popper, S. W., y Bankes, S. C. (2003), Shaping the Next One Hundred Years. New Methods for Quantitative, Long-Term Policy Analysis, RAND publications, Santa Mónica, CA. <http://www.rand.org>.
- López Díaz-Delgado, E., y Martínez Vicente, S. (2000), Iniciación a la simulación dinámica. Aplicaciones a sistemas económicos y empresariales, Ariel, Barcelona.
- MAP (2004), Agencia Estatal de Evaluación de la Calidad de los Servicios y de las Políticas Públicas. Comisión para el estudio y creación. Informe 4 octubre 2004, MAP, Madrid.

- MAP (2005), Proyecto de Ley de las Agencias Estatales para la mejora de los Servicios Públicos, MAP, Madrid.
- Monnier, E. (1995), Evaluación de la acción de los poderes públicos, Instituto de Estudios Fiscales, Madrid.
- North, D. C. (1993), Five Propositions about Institutional Change, Economic History 9309001, *Economics Working Paper Archive at WUSTL*.
- OCDE (1992), La evaluación de programas de empleo y de medida sociales, Ministerio de Trabajo y Seguridad Social, Madrid.
- Pazos, M. (1996), La evaluación de programas en la Administración Pública Española: el estado de la cuestión, Instituto de Estudios Fiscales, Madrid.
- Pinilla Pallejà, R. (2006), Más allá del bienestar. La renta básica de ciudadanía como innovación social basada en la evidencia, Icaria, Barcelona (en prensa).
- Riccio, J. A. y Bloom, H. S. (2001), Extending the Reach of Randomized Social Experiments: New Directions in Evaluations of American Welfare-to-Work and Employment Initiatives, New York, MDRC, <http://www.mdrc.org>.
- Sackett, D. L., Richardson, W. S., Rosenberg, W., y Haynes R. B. (1997), Medicina Basada en la Evidencia. Cómo ejercer y enseñar la MBE, Churchill-Livingstone, Madrid.
- Schmidt, C. M. (2001), Knowing What Works. The Case for Rigorous Programme Evaluation, CEPR (Centre for Economic Policy Research), London.
- Vedung, E. (1997), Evaluación de políticas públicas y programas, Ministerio de Trabajo y Asuntos Sociales, INSS, Madrid.

## Documentos de trabajo publicados

- 1/2003. **Servicios de atención a la infancia en España: estimación de la oferta actual y de las necesidades ante el horizonte 2010.** María José González López.
- 2/2003. **La formación profesional en España. Principales problemas y alternativas de progreso.** Francisco de Asís de Blas Aritio y Antonio Rueda Serón.
- 3/2003. **La Responsabilidad Social Corporativa y políticas públicas.** Alberto Lafuente Félez, Víctor Viñuales Edo, Ramón Pueyo Viñuales y Jesús Llaría Aparicio.
- 4/2003. **V Conferencia Ministerial de la OMC y los países en desarrollo.** Gonzalo Fanjul Suárez.
- 5/2003. **Nuevas orientaciones de política científica y tecnológica.** Alberto Lafuente Félez.
- 6/2003. **Repensando los servicios públicos en España.** Alberto Infante Campos.
- 7/2003. **La televisión pública en la era digital.** Alejandro Perales Albert.
- 8/2003. **El Consejo Audiovisual en España.** Ángel García Castillejo.
- 9/2003. **Una propuesta alternativa para la Coordinación del Sistema Nacional de Salud español.** Javier Rey del Castillo.
- 10/2003. **Regulación para la competencia en el sector eléctrico español.** Luis Atienza Serna y Javier de Quinto Romero.
- 11/2003. **El fracaso escolar en España.** Álvaro Marchesi Ullastres.
- 12/2003. **Estructura del sistema de Seguridad Social. Convergencia entre regímenes.** José Luis Tortuero Plaza y José Antonio Panizo Robles.
- 13/2003. **The Spanish Child Gap: Rationales, Diagnoses, and Proposals for Public Intervention.** Fabrizio Bernardi.
- 13\*/2003. **El déficit de natalidad en España: análisis y propuestas para la intervención pública.** Fabrizio Bernardi.
- 14/2003. **Nuevas fórmulas de gestión en las organizaciones sanitarias.** José Jesús Martín Martín.
- 15/2003. **Una propuesta de servicios comunitarios de atención a personas mayores.** Sebastián Sarasa Urdiola.
- 16/2003. **El Ministerio Fiscal. Consideraciones para su reforma.** Olga Fuentes Soriano.
- 17/2003. **Propuestas para una regulación del trabajo autónomo.** Jesús Cruz Villalón.
- 18/2003. **El Consejo General del Poder Judicial. Evaluación y propuestas.** Luis López Guerra.
- 19/2003. **Una propuesta de reforma de las prestaciones por desempleo.** Juan López Gandía.
- 20/2003. **La Transparencia Presupuestaria. Problemas y Soluciones.** Maurici Lucena Betriu.
- 21/2003. **Análisis y evaluación del gasto social en España.** Jorge Calero Martínez y Mercè Costa Cuberta.
- 22/2003. **La pérdida de talentos científicos en España.** Vicente E. Larraga Rodríguez de Vera.
- 23/2003. **La industria española y el Protocolo de Kioto.** Antonio J. Fernández Segura.
- 24/2003. **La modernización de los Presupuestos Generales del Estado.** Enrique Martínez Robles, Federico Montero Hita y Juan José Puerta Pascual.
- 25/2003. **Movilidad y transporte. Opciones políticas para la ciudad.** Carme Miralles-Guasch y Ángel Cebollada i Frontera.
- 26/2003. **La salud laboral en España: propuestas para avanzar.** Fernando G. Benavides.
- 27/2003. **El papel del científico en la sociedad moderna.** Pere Puigdomènech Rosell.
- 28/2003. **Tribunal Constitucional y Poder Judicial.** Pablo Pérez Tremps.
- 29/2003. **La Audiencia Nacional: una visión crítica.** José María Asencio Mellado.
- 30/2003. **El control político de las misiones militares en el exterior.** Javier García Fernández.
- 31/2003. **La sanidad en el nuevo modelo de financiación autonómica.** Jesús Ruiz-Huerta Carbonell y Octavio Granado Martínez.

- 32/2003. **De una escuela de mínimos a una de óptimos: la exigencia de esfuerzo igual en la Enseñanza Básica.** Julio Carabaña Morales.
- 33/2003. **La difícil integración de los jóvenes en la edad adulta.** Pau Baizán Muñoz.
- 34/2003. **Políticas de lucha contra la pobreza y la exclusión social en España: una valoración con EspaSim.** Magda Mercader Prats.
- 35/2003. **El sector del automóvil en la España de 2010.** José Antonio Bueno Oliveros.
- 36/2003. **Publicidad e infancia.** Purificación Llaquet, M<sup>a</sup> Adela Moyano, María Guerrero, Cecilia de la Cueva, Ignacio de Diego.
- 37/2003. **Mujer y trabajo.** Carmen Sáez Lara.
- 38/2003. **La inmigración extracomunitaria en la agricultura española.** Emma Martín Díaz.
- 39/2003. **Telecomunicaciones I: Situación del Sector y Propuestas para un modelo estable.** José Roberto Ramírez Garrido y Juan Vega Esquerrá.
- 40/2003. **Telecomunicaciones II: Análisis económico del sector.** José Roberto Ramírez Garrido y Álvaro Escribano Sáez.
- 41/2003. **Telecomunicaciones III: Regulación e Impulso desde las Administraciones Públicas.** José Roberto Ramírez Garrido y Juan Vega Esquerrá.
- 42/2004. **La Renta Básica. Para una reforma del sistema fiscal y de protección social.** Luis Sanzo González y Rafael Pinilla Pallejà.
- 43/2004. **Nuevas formas de gestión. Las fundaciones sanitarias en Galicia.** Marciano Sánchez Bayle y Manuel Martín García.
- 44/2004. **Protección social de la dependencia en España.** Gregorio Rodríguez Cabrero.
- 45/2004. **Inmigración y políticas de integración social.** Miguel Pajares Alonso.
- 46/2004. **TV educativo-cultural en España. Bases para un cambio de modelo.** José Manuel Pérez Tornero.
- 47/2004. **Presente y futuro del sistema público de pensiones: Análisis y propuestas.** José Antonio Griñán Martínez.
- 48/2004. **Contratación temporal y costes de despido en España: lecciones para el futuro desde la perspectiva del pasado.** Juan J. Dolado y Juan F. Jimeno.
- 49/2004. **Propuestas de investigación y desarrollo tecnológico en energías renovables.** Emilio Menéndez Pérez.
- 50/2004. **Propuestas de racionalización y financiación del gasto público en medicamentos.** Jaume Puig-Junoy y Josep Llop Talaverón.
- 51/2004. **Los derechos en la globalización y el derecho a la ciudad.** Jordi Borja.
- 52/2004. **Una propuesta para un comité de Bioética de España.** Marco-Antonio Broggi Trias.
- 53/2004. **Eficacia del gasto en algunas políticas activas en el mercado laboral español.** César Alonso-Borrego, Alfonso Arellano, Juan J. Dolado y Juan F. Jimeno.
- 54/2004. **Sistema de defensa de la competencia.** Luis Berenguer Fuster.
- 55/2004. **Regulación y competencia en el sector del gas natural en España. Balance y propuestas de reforma.** Luis Atienza Serna y Javier de Quinto Romero.
- 56/2004. **Propuesta de reforma del sistema de control de concentraciones de empresas.** José M<sup>a</sup> Jiménez Laiglesia.
- 57/2004. **Análisis y alternativas para el sector farmacéutico español a partir de la experiencia de los EE UU.** Rosa Rodríguez-Monguió y Enrique C. Seoane Vázquez.
- 58/2004. **El recurso de amparo constitucional: una propuesta de reforma.** Germán Fernández Farreres.
- 59/2004. **Políticas de apoyo a la innovación empresarial.** Xavier Torres.
- 60/2004. **La televisión local entre el limbo regulatorio y la esperanza digital.** Emili Prado.
- 61/2004. **La universidad española: soltando amarras.** Andreu Mas-Colell.
- 62/2005. **Los mecanismos de cohesión territorial en España: un análisis y algunas propuestas.** Ángel de la Fuente.
- 63/2005. **El libro y la industria editorial.** Gloria Gómez-Escalonilla.
- 64/2005. **El gobierno de los grupos de sociedades.** José Miguel Embid Irujo, Vicente Salas Fumás.
- 65(I)/2005. **La gestión de la demanda de electricidad Vol. I.** José Ignacio Pérez Arriaga, Luis Jesús Sánchez de Tembleque, Mercedes Pardo.

- 65(II)/2005. **La gestión de la demanda de electricidad Vol. II (Anexos)**. José Ignacio Pérez Arriaga, Luis Jesús Sánchez de Tembleque, Mercedes Pardo.
- 66/2005. **Responsabilidad patrimonial por daño ambiental: propuestas de reforma legal**. Ángel Manuel Moreno Molina.
- 67/2005. **La regeneración de barrios desfavorecidos**. María Bruquetas Callejo, Fco. Javier Moreno Fuentes, Andrés Walliser Martínez.
- 68/2005. **El aborto en la legislación española: una reforma necesaria**. Patricia Laurenzo Copello.
- 69/2005. **El problema de los incendios forestales en España**. Fernando Estirado Gómez, Pedro Molina Vicente.
- 70/2005. **Estatuto de laicidad y Acuerdos con la Santa Sede: dos cuestiones a debate**. José M.<sup>a</sup> Contreras Mazarío, Óscar Celador Angón.
- 71/2005. **Posibilidades de regulación de la eutanasia solicitada**. Carmen Tomás-Valiente Lanuza.
- 72/2005. **Tiempo de trabajo y flexibilidad laboral**. Gregorio Tudela Cambroner, Yolanda Valdeolivas García.
- 73/2005. **Capital social y gobierno democrático**. Francisco Herreros Vázquez.
- 74/2005. **Situación actual y perspectivas de desarrollo del mundo rural en España**. Carlos Tió Saralegui.
- 75/2005. **Reformas para revitalizar el Parlamento español**. Enrique Guerrero Salom.
- 76/2005. **Rivalidad y competencia en los mercados de energía en España**. Miguel A. Lasheras.
- 77/2005. **Los partidos políticos como instrumentos de democracia**. Henar Criado Olmos.
- 78/2005. **Hacia una deslocalización textil responsable**. Isabel Kreisler.
- 79/2005. **Conciliar las responsabilidades familiares y laborales: políticas y prácticas sociales**. Juan Antonio Fernández Cordón y Constanza Tobío Soler.
- 80/2005. **La inmigración en España: características y efectos sobre la situación laboral de los trabajadores nativos**. Raquel Carrasco y Carolina Ortega.
- 81/2005. **Productividad y nuevas formas de organización del trabajo en la sociedad de la información**. Rocío Sánchez Mangas.
- 82/2006. **La propiedad intelectual en el entorno digital**. Celeste Gay Fuentes.
- 83/2006. **Desigualdad tras la educación obligatoria: nuevas evidencias**. Jorge Calero.
- 84/2006. **I+D+i: selección de experiencias con (relativo) éxito**. José Antonio Bueno Oliveros.
- 85/2006. **La incapacidad laboral en su contacto médico: problemas clínicos y de gestión**. Juan Gervas, Ángel Ruiz Téllez y Mercedes Pérez Fernández.
- 86/2006. **La universalización de la atención sanitaria. Sistema Nacional de Salud y Seguridad Social**. Francisco Sevilla.
- 87/2006. **El sistema de servicios sociales español y las necesidades derivadas de la atención a la dependencia**. Pilar Rodríguez Rodríguez.
- 88/2006. **La desalinización de agua de mar mediante el empleo de energías renovables**. Carlos de la Cruz.
- 89/2006. **Bases constitucionales de una posible política sanitaria en el Estado autonómico**. Juan José Solozábal Echavarría.
- 90/2006. **Desigualdades territoriales en el Sistema Nacional de Salud (SNS) de España**. Beatriz González López-Valcárcel y Patricia Barber Pérez.

