

SE-valoración

Revista de la Sociedad Española de Evaluación | Año 3 | nº 13 | Diciembre 2010



Evaluación de las Políticas de Empleo



Sociedad Española de Evaluación

Inicio Recursos FORO VIRTUAL Acerca de la SEE Contacto con SEE Inscripción a SEE Master Mapa del Web

Calendario de Eventos
Conferencias de la SEE
Revista SEE-Valuación
Boletín SEE
Enlaces de Interés
Foro SEE
Empleo

[CECCOD] III Curso de Evaluación de Intervenciones de Cooperación para el Desarrollo

Comienza: 16/07/2008 - 07:00
Finaliza: 02/12/2008 - 07:59
Temática: EUROMT-1

El Centro de Estudios de Cooperación al Desarrollo (CECCOD), formado por el Instituto Universitario de Estudios Europeos de la Universidad CEU San Pablo y la Fundación CODESPA, recuerda que sigue abierto el plazo de admisión del III Curso en Evaluación de Intervenciones de Cooperación para el Desarrollo, que se desarrollará entre el 23 de septiembre y el 2 de diciembre de 2008 en la Universidad CEU San Pablo de Madrid. Adjunto se remite el programa del Curso.

Leer más Calendario Cursos Cursos y Jornadas

Recursos para Evaluadores
Breve de Datos
Conferencias
Jornadas
Manuales
Revistas
Estudios y Ensayos

Categorías de Noticias
Noticias sobre Políticas Públicas
NOCIAS de AEVAL
Publicaciones
Noticias de SEE
Noticias sobre Evaluación

Inicio de sesión
Número de usuario: *

Sociedad Española de Evaluación

Inicio Recursos FORO VIRTUAL Acerca de la SEE Contacto con SEE Inscripción a SEE Master Mapa del Web

Calendario de Eventos
Conferencias de la SEE
Revista SEE-Valuación
Boletín SEE
Enlaces de Interés
Foro SEE
Empleo

Inscripción a la Sociedad Española de Evaluación de Políticas Públicas

8 Marzo, 2008 - 10:26 — admin

Podrán ser miembros de la SEE:

Todas aquellas personas cuya actividad profesional esté vinculada a la evaluación de políticas públicas y compartan los objetivos de la SEE.

En la Sociedad Española de Evaluación se pueden distinguir tres tipos de socios:

- Socios Individuales, personas físicas con interés por la evaluación de políticas públicas.
- Socios Institucionales, aquellas entidades: Administraciones Públicas, centros de investigación y universidades, con interés en el campo de la evaluación.
- Socios Honoríficos, que serán aquellos que hayan contribuido, bajo diferentes formas, de manera trascendental a los objetivos de la Sociedad. No cuentan con el derecho a voto ni a ser elegidos en las asambleas de la Sociedad.

TIPO DE SOCIO	BO E	CUOTAS
Socio Individual	100 €	
Socio Institucional		Consultar con SEE
Socio Laboral/empresario		Exento

¿Cuáles son los derechos de socios y socias?

- Tomar parte de cuantas actividades organice la Sociedad en cumplimiento de sus fines.

Recursos para Evaluadores
Breve de Datos
Conferencias
Jornadas
Manuales
Revistas
Estudios y Ensayos
Congresos
Informes
Guías
Cursos
Publicaciones

Inicio de sesión
Número de usuario: *

ACERCA DE LA SEE

8 Marzo, 2008 - 13:01 — admin

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

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

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

Recursos para Evaluadores
Breve de Datos
Conferencias
Jornadas
Manuales
Revistas
Estudios y Ensayos
Congresos
Informes
Guías
Cursos
Publicaciones

Inicio de sesión
Número de usuario: *

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

¡¡ Visítanos!!

www.sociedaddeevaluacion.org

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

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

SOCIEDAD ESPAÑOLA DE EVALUACIÓN DE POLÍTICAS PÚBLICAS

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

SEE-valoración

SUMARIO

4 LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN ESPAÑA: ACTORES PÚBLICOS Y PRIVADOS

José L. Osuna

8 EVALUACIÓN DEL PROGRAMA DE LA RENTA MÍNIMA-DE INSERCIÓN DE CATALUÑA

David Casado Marín

25 EVALUACIÓN DE LOS PLANES DE EMPLEO DE CATALUÑA: 2005-2007

Jaume Blasco

37 ANÁLISIS DE LOS OBJETOS EVALUADOS EN ESPAÑA: LA EVALUACIÓN Y LAS LIMITACIONES DE LAS POLÍTICAS SOCIALES

Juan Andrés Ligeró, Belén Rodríguez, Irene Rosales, Isabel Morandeira y Maruxa Fernández

51 LA EVALUACIÓN DE POLÍTICAS PÚBLICAS: EL CASO ESPECÍFICO DE LA COMUNIDAD FORAL DE NAVARRA

Asun Ruesta Ozcoidi, Anabel Salcedo Condón, Marta Zalba Ibero y Graciela Vázquez Rodríguez

64 EVENTOS, FORMACIÓN Y NOTICIAS

EDITORIAL

En las páginas de éste nuestro último número de dos mil diez, los lectores de **SEE-valoración** – en correspondencia con las fechas en que lo reciban – van a encontrarse con el contenido y desarrollo de una fiesta: la entrega del **I Premio de investigación en Evaluación de Políticas Públicas**, que lleva el nombre del primer Presidente y fundador de la SEE: el Prof. Dr. D. **Carlos Román del Río**.

Y con ello, la **SEE** cumple uno de sus objetivos fundacionales: “fomentar la investigación científica y la difusión de las mejores prácticas entre los agentes implicados tanto en la esfera universitaria, de la función pública, del ámbito empresarial y profesional y de los poderes públicos”. Por eso, estamos de fiesta; porque, dando rienda a los sueños hemos cumplido objetivos. Tácitos, que nos comprometían con la figura de Carlos Román y explícitos, por figurar en nuestros estatutos. Y lo estamos también porque la evaluación, en cuanto a la investigación y conocimiento, goza de buena salud.

Nuestra felicitación al premiado **David Casado**, analista del área de evaluación del Instituto Catalán de Evaluación de Políticas Públicas (IVALÚA) y a quienes han gozado de Mención Honorífica: **M^a Asunción Ruesta**, **Ana Isabel Salcedo**, del Instituto Navarro de Administración Pública y **Marta Zalba Ibero** y **Graciela Vázquez Rodríguez** de Consultores Asociados; **Jaume Blasco**, también de IVALÚA y a **Juan Andrés Ligeró** de la Universidad Carlos III, **Belén Rodríguez** de la Fundación Entreculturas, **Inés Rosales** de NODUS Consultores, **Isabel Morandeira** de Cruz Roja y **Maruxa Fernández** de Red2Red Consultores. Este número de **SEE-valoración** va dedicado a ellos; a su rigor con la publicación de las síntesis de sus trabajos, a su esfuerzo con el reconocimiento de su contribución al desarrollo de la cultura de la evaluación en nuestro país y a su amabilidad por habernos acompañado en Sevilla.

Por último, **SEE-valoración** quiere destacar y agradecer la labor de mecenazgo del **Instituto de Estudios de Economía Evaluación y Empleo**, sin cuya benefactora aportación no habríamos podido alcanzar el objetivo de la convocatoria. Tampoco podríamos haberlo dotado de la seriedad y el rigor exigible en aras a la credibilidad de la convocatoria, sin la colaboración altruista de los miembros del Jurado: Dña. **Cristina Narbona**, Dña. **Isabel de Haro**, D. **Antonio Pascual**, D. **Mikel Landabaso**, D. **Fernando González-Laxe** y D. **Javier Rodríguez**.

Una celebración, cuya perpetuación en el tiempo se convierte en misión para la SEE. En efecto, tenemos la intención de continuar y convocar la segunda edición del **Premio Carlos Román**; ciertamente, no sabemos cuándo, no tenemos los recursos ... pero, siendo tiempo de peticiones y dado que la **AEVAL** ha vuelto “a casa”, con el “reno” tirando de la inteligencia y los “camellos” cargando con la férrea voluntad de hacerlo, seguro que encontraremos la oportunidad.

¡Felices Fiestas y buenas evaluaciones !

Edita: SOCIEDAD ESPAÑOLA DE EVALUACIÓN
www.sociedadevaluacion.org

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

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

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



ANÁLISIS DE LOS OBJETOS EVALUADOS EN ESPAÑA: LA EVALUACIÓN Y LAS LIMITACIONES DE LAS POLÍTICAS SOCIALES

Investigador principal: **Juan Andrés Ligeró**; Investigadoras de apoyo: **Belén Rodríguez, Irene Rosales, Isabel Morandeira y Maruxa Fernández**

El presente documento describe una investigación sobre la situación de la evaluación en España y, concretamente, el desarrollo por áreas de actividad pública, identificando qué posibles factores afectan a éste crecimiento.

1. JUSTIFICACIÓN

En España queda mucho trabajo por hacer en el campo de la evaluación. No obstante, la literatura muestra indicios de un fuerte desarrollo del sector en nuestro país (Pazos y Zapico-Goñi, 2002; Bustelo 2006; Fernández- Ramírez y Reboloso, 2006). También hay hechos que así lo sugieren, como son la creación de la Agencia Estatal de Evaluación de las Políticas y Calidad de los Servicios, la consolidación de redes profesionales como la Sociedad Española de Evaluación, la existencia de propuestas metodológicas especializadas (González y Osuna, 2007; Osuna y Márquez, 2000), la presencia y el aumento de la oferta de formación específica y, quizá

lo más significativo de todo, la emergencia de un campo profesional propio dedicado a la evaluación¹.

Las mejoras que parecen provocar la evaluación en la acción pública son beneficios que redundan en el bien de la ciudadanía. Aunque todavía se está en un momento incipiente, cabe preguntarse si el desarrollo de la evaluación se está produciendo correctamente, con el mismo esfuerzo y de manera armónica en los distintos sectores de actividad pública.

Una posible deficiencia en la conformación del campo profesional es que existan sectores de políticas públicas más evaluados que otros. El desigual tratamiento de la evaluación tendría consecuencias sobre el desarrollo de las propias políticas y, por consiguiente, sobre las poblaciones destinatarias.

No se puede conocer el nivel de desarrollo de la evaluación por sectores atendiendo exclusivamente a la literatura existente, no se encuentran estudios que

¹ Esta conclusión se puso de manifiesto en el III Seminario de Experiencias de Evaluación de Programas y Políticas titulado "Desarrollo Profesional: Campos emergentes en evaluación". Realizado en el Centro de Estudios de Gestión de la Universidad Complutense de Madrid el 1 de junio del 2006.

aborden la incidencia de la evaluación por sectores políticos.

Por ello, para contestar a la pregunta es necesario utilizar otros métodos de indagación, como se expone en el apartado de metodología.

2. OBJETIVOS GENERALES Y ESPECÍFICOS

El objetivo general de este estudio es analizar la situación de la evaluación en los distintos sectores públicos españoles y sus causas. Los objetivos específicos que se contemplan son los siguientes:

- ▶ Detectar el grado de desarrollo de la evaluación en los diferentes sectores.
- ▶ Identificar factores que influyen en el encargo de evaluación.
- ▶ Analizar el grado de institucionalización de la evaluación.

3. METODOLOGÍA

En un artículo de Leeuw, Toulemonde y Brouwers (1999) sobre la actividad en el campo profesional de la evaluación en Europa los autores proponen instrumentos para medir el crecimiento (o el decrecimiento) de la actividad evaluadora, como son: la encuesta, la contabilización de informes de evaluación, los análisis bibliométricos y las evidencias

indirectas. Los autores en este caso optan por una encuesta a proveedores de servicios de evaluación pertenecientes tanto al sector privado como al académico.

En otros estudios significativos (aunque ya antiguos) sobre la evaluación por sectores políticos, se ha optado por recopilar información o bien a través de una muestra de evaluaciones, como es el caso del trabajo que Nioche y Poinard (1984) desarrollan en Francia en el que trabajan con una muestra de 269 evaluaciones; o bien a través del análisis de los contratos ofrecidos por la administración, como es el caso del estudio de las Solicitudes de Propuestas de Evaluación² en Estados Unidos de Freeman y Solomon (1981).

Para la presente investigación, entre las diferentes estrategias, se ha optado por analizar dos muestras de evaluaciones, una construida a partir de la información expuesta en congresos, jornadas específicas y tesinas y otra proveniente de una encuesta autocumplimentada por evaluadores/as.

Los pasos que se han seguido son los siguientes:

- 1) Revisión bibliográfica: se ha revisado la literatura existente sobre el estado de la evaluación en España y se han analizado los distintos instrumentos empleados por otros investigadores para analizar la evaluación por sectores³.

² El término en inglés para este concepto es Request for Proposal.

³ Se parte del estudio de Furubo y Sandahl (International Atlas of Evaluation, 2002) que aplican un conjunto de indicadores de "evaluative culture" a un total de 21 países, situando a España en las posiciones más bajas.

Estudios realizados por Pazos y Zapico-Goñi (2002), Bustelo (2006) y Fernández-Ramírez y Reboloso (2006) muestran indicios de un fuerte desarrollo de la evaluación en el último decenio. Se detecta que una posible deficiencia en la conformación del campo profesional es que estén existiendo sectores de las políticas públicas más evaluados que otros. Este es el caso de lo detectado por Toulemonde y Brouwers (1999) para el caso de EEUU y Nioche y Poinard (1984) para el caso francés.

2) Desarrollo de la primera muestra de evaluaciones a partir de la creación y depuración de una base de datos de evaluaciones recopilada a partir de las siguientes fuentes:

- ▶ I, II, III, IV y V Seminario de Experiencias de Evaluación, UCM (2004, 2005, 2006, 2007 y 2008).
- ▶ IV y V Congreso de la Sociedad Española de Evaluación (2005 y 2009).
- ▶ VIII Congreso Español de Sociología. Evaluación de Programas de Intervención Social. Grupo de Trabajo Metodología (2004).
- ▶ VIII Congreso Español de Ciencia Política: "Política para un mundo en cambio". Grupo de trabajo de Análisis y Evaluación de las Políticas Públicas (2007).
- ▶ Tesinas o memorias de grado del Magíster de Evaluación de Programas y Políticas Públicas, UCM (de los años 2003 al 2008).
- ▶ Diversas Jornadas sobre Planes de Prevención de Drogas.

Al no tratarse ni de un censo ni de un muestreo aleatorio es probable que sobre la base de datos operen diversos sesgos. Se han identificado al menos tres posibles influencias:

- ▶ Un exceso de evaluaciones sobre programas de prevención de drogas producto de la atención prestada a dichos foros especializados. Como medida correctora se han eliminado las evaluaciones referidas a la prevención por la dificultad de discernir cuáles procedían de los foros especializados y cuáles no. El resultado final es una base de datos con 159 evaluaciones.

- ▶ Al tener en cuenta sólo los foros específicos de evaluación, que en España suelen estar vinculados a los ámbitos de la sociología, psicología, economía, ciencia política y administración, es probable que los sectores más evaluados sean los más cercanos a estas disciplinas.

- ▶ Puede haber sectores que sean menos proclives a presentar sus resultados en público por diversas razones, como puede ser la necesidad de preservar la información (seguridad interior, defensa, etc.), existencia de amenazas de continuidad del programa o falta de incentivos para comunicar las informaciones. De este modo, los sectores más influidos por estos posibles factores tendrían menos presencia en la base de datos.

3) Desarrollo de la segunda muestra de evaluaciones a partir de los datos recopilados en la encuesta autocumplimentada a evaluadores: desarrollada por las profesoras María Bustelo (UCM) y Jody Fitzpatrick (Universidad de Denver, Colorado). Esta encuesta consistió en un cuestionario online enviado a contactos procedentes de una base de datos de evaluadores y personas afines a la evaluación del Magíster en Evaluación de Programas, una invitación a todos los miembros de la Sociedad Española de Evaluación, a los contactos facilitados por la Agencia Estatal de Evaluación y Calidad (AEVAL) e IVALUA (Institut Català d'Avaluació de Polítiques Públiques). Los cuestionarios que se recibieron fueron un total de 150.

En esta ocasión, se han detectado varios sesgos en la muestra:

- ▶ Se desconoce el universo total de evaluadores de España, habrá gente que trabaja en evaluación que no estén incorporados en las

bases de datos citadas a los que no les habrá llegado el cuestionario.

- ▶ Las encuestas autocumplimentadas están sesgadas por el mayor grado de motivación de quién responde al cuestionario.
 - ▶ Es el criterio del entrevistado quien decide cuál es la evaluación que se va a describir.
- 4) Análisis de ambas bases de datos, pudiendo distinguirse tres fases: clasificación por sectores, clasificación por “universal o selectiva” y análisis de la relación entre ambas variables.

Primer nivel de análisis: Clasificación por sectores, pasos seguidos:

- ▶ Se observaba el título de la evaluación para ver si incluía información clara sobre el objeto. Si era así se clasificaba directamente en alguna de las categorías.
- ▶ En el caso de que existieran dificultades para asignar alguna evaluación a un sector ministerial específico, se consultaba con la persona o entidad que figuraba como responsable para asegurar la correcta clasificación.
- ▶ Si a pesar de todo no se resolvía la duda o no se podía contactar, se eliminaba estrictamente la evaluación de la base de datos.

A continuación, ambas bases de datos fueron incluidas en una sola matriz de SPSS, manteniendo información sobre su origen, y se realizaron una serie de análisis:

- ▶ Para indagar sobre si la distribución era homogénea por sectores se solicitaron distribuciones de frecuencias.
- ▶ Para conocer si las dos muestras, aún provenientes de fuentes diferentes, eran similares en su distribución de porcentajes, se solicitaron tablas de contingencia pidiendo la prueba de inferencia estadística chi-cuadrado. Se ha asumido el nivel de confianza del 95%. También se solicitaron residuos ajustados tipificados para observar si existían diferencias estadísticamente significativas entre las parejas de categorías de las dos muestras. Para esta prueba también se ha asumido el nivel de confianza del 95%, entendiendo que no eran similares los porcentajes cuando sus residuos superaban en valor absoluto el 1,96.

Segundo nivel de análisis: Clasificación por “universal o selectiva”

En este momento, se procedió a clasificar nuevamente las evaluaciones en función de que su objeto fuera universal o selectivo, siendo este proceso similar al descrito en el caso de los sectores políticos.

Una vez clasificadas las evaluandas, se solicitaron distribuciones de frecuencias y se calcularon las diferencias de proporciones “d”⁴ para las dos muestras. La diferencia obtenida no se puede asimilar como el efecto neto de “universal/selectiva” en hacer o no evaluación. Para ello hubiera sido necesario haberse aproximado a modelos experimentales, cuasiexperimentales o de control estadístico de terceras o múltiples variables intervinientes. Ninguna de estas opciones era posible con la información que se tenía, salvo

4 Sánchez Carrión, J.J. (1999).

la asociación con el presupuesto por sectores públicos. No se encontró relación significativa⁵.

Al obtener en dos muestras diferentes datos similares, se reduce la probabilidad de error tipo I y se incrementa la confianza sociológica en los datos.

Tercer nivel de análisis: Relación entre ambas variables

Un tercer análisis ha sido cruzar la variable “universal/selectiva” por “sectores políticos”. Se ha establecido una clasificación donde aparecen aquellos sectores que se apoyan más en estrategias universales, en estrategias selectivas o en ninguna de ellas especialmente. Para lograr dicha clasificación, se solicitaban cruces de variables, estadístico chi-cuadrado y

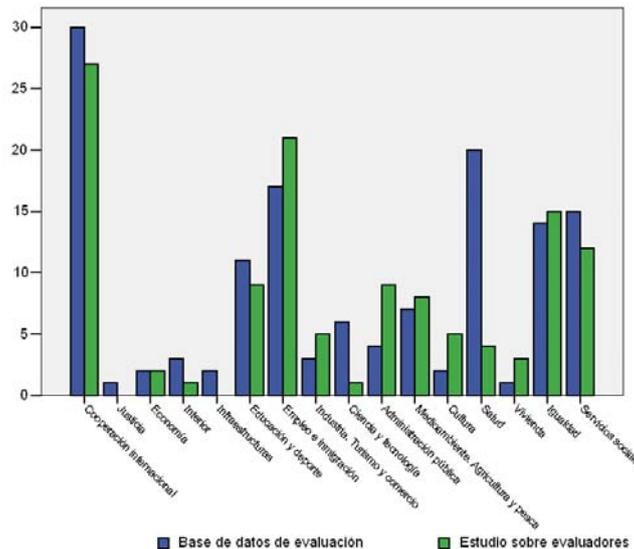
residuos tipificados corregidos. Se clasificaban las evaluandas en cada categoría si el residuo tenía un valor superior a 1,96. En los casos en los que los valores se encontraban entre -1,96 y +1,96 se han clasificado en las categorías “Iguales” y se ha considerado que no hay diferencias significativas entre las estrategias.

De estos tres niveles de análisis se han extraído los resultados que se exponen a continuación.

4. TRABAJO DE INVESTIGACIÓN

4.1. ANÁLISIS

Como se ha recogido en el apartado metodológico, el primer análisis realizado sobre las muestras fue comparar la distribución de frecuencias que se muestra en el **gráfico 1**.



5 Uno de los factores que se ha analizado es el tamaño del presupuesto por sectores públicos. Los Presupuestos Generales del Estado muestran que las frecuencias del Gráfico 1 no guardan relación con el presupuesto asignado a cada área. Por ejemplo, las áreas que más presupuesto reciben en proporciones mayores son Empleo e Inmigración, Economía, Infraestructuras, Defensa e Interior que, con la excepción de Empleo, son de las áreas en las que apenas se registran evaluaciones en las muestras. No obstante, este análisis no deja de ser un mero indicio ya que las muestras de evaluaciones responden a diferentes administraciones y periodos y los presupuestos públicos son los del Estado para 2009.

La primera conclusión que se puede extraer del gráfico es que básicamente las dos muestras se distribuyen de forma similar⁶ entre los diferentes sectores salvo en la categoría de salud donde se evidencia un comportamiento diferenciado.

La segunda conclusión es que no hay el mismo volumen de evaluaciones en los diferentes sectores⁷. Hay algunos sectores que son muy evaluados, como por ejemplo, asuntos exteriores y cooperación internacional, empleo e inmigración, igualdad de género y servicios sociales. Otros sectores como defensa, justicia, fomento, interior, vivienda o economía son menos evaluados o incluso nada evaluados, como es el caso del sector de defensa.

La ausencia de evaluaciones en algunos sectores puede ser atribuible a la necesidad de confidencialidad de sus políticas como podría ser en el caso de defensa, interior o justicia. Pero esto no sería extensible a otras áreas como vivienda, economía o infraestructura, lo que invita a buscar otras posibles explicaciones.

Por otro lado, la dirección que toman los datos suena familiar. Haciendo uso de la invitación de Leeuw y otros (1999) a utilizar evidencias indirectas, en la bibliografía española suelen encontrarse más ejemplos y referencias de evaluaciones sobre

intervención social, drogas, salud, educación, género y empleo que de otros sectores. Quizás esto también se pueda decir de la bibliografía en otros países. En el caso de EEUU, es más frecuente encontrar referencias de programas sociales o educacionales que de otras áreas. Carol Weiss llega a citar que incluso algunos de tales campos, como educación, pobreza o prevención del crimen han sido claves para el desarrollo de la disciplina de evaluación (1998:12).

En definitiva, los datos de las dos muestras arrojan una consistente y significativa desigualdad en el volumen de evaluaciones por sectores, la cuestión ahora es saber por qué se produce esta desproporción y si efectivamente caracteriza la evaluación en España.

4.2. EXPLICACIÓN

Entre los diversos factores⁸ que se han barajado como posibles explicaciones se ha encontrado una variable que explica en parte la desproporción del número de evaluaciones y que incide sobre el comportamiento de los decisores políticos a la hora de encargar evaluaciones. Dicha variable se ha denominado “carácter universal o selectivo de las intervenciones” como se explica a continuación:

6 Como ya se menciona en el apartado de metodología, para ver si existían diferencias significativas entre la distribución se analizaron los residuos ajustados tipificados. Este análisis indica que no hay diferencias entre las parejas de categorías con la única excepción de la categoría de salud donde sí se encuentran diferencias.

7 Existen discrepancias entre ambas muestras para la categoría de salud.

8 Uno de los factores que se ha analizado es el tamaño del presupuesto por sectores públicos. El presupuesto público asignado a nivel de Estado Central a cada sector de actividad no parece estar relacionado con la distribución de frecuencias que aparece en el Gráfico 1. Por ejemplo, aquellas áreas que reciben una mayor proporción de presupuesto son: empleo e inmigración, economía, fomento, defensa y seguridad interior, sectores que, con la excepción de las políticas de empleo, son las áreas en las que apenas se registran evaluaciones en ambas muestras. No obstante, este análisis no deja de ser un mero indicio ya que las muestras de evaluaciones responden a diferentes administraciones y periodos y los presupuestos públicos son los del Estado para el 2009.

Carácter universal o selectivo de las intervenciones

Se entiende por políticas o programas universales aquellas intervenciones que buscan beneficiar por igual al conjunto de la población existente en un determinado ámbito. Se podrían citar como ejemplos los sistemas de sanidad pública, las infraestructuras públicas, la gestión del tráfico, los servicios policiales o las políticas de defensa nacional.

Las políticas selectivas o de discriminación positiva favorecen a un determinado segmento social o grupo que se diferencia del resto por compartir alguna característica o condición demográfica, social o cultural. Se fundamentan en que no toda la ciudadanía tiene las mismas opciones ni comparte iguales dificultades. Se requiere de una acción posterior para dotar de más oportunidades a quienes no las tenían.

Entre los ejemplos se pueden citar los programas de alfabetización de adultos, de inserción de minorías, realojo de población chabolista o los ingresos de rentas mínimas.

Observando los datos del gráfico 1, la hipótesis planteada en este estudio de investigación es que se evalúan más las políticas selectivas que las universales indistintamente del sector político donde se enmarcaran las intervenciones.

Para comprobar esta hipótesis, se llevó a cabo el segundo análisis de las muestras clasificándolas según el criterio “universal/selectivo”.

Se analizó una distribución de frecuencias de las muestras y se extrajeron las diferencias de proporciones “d” tal y como se muestra en la tabla 1.

Tabla 1

	Universal % evaluaciones sobre evaluandas universales	Selectiva % evaluaciones sobre evaluandas selectivas	”d” diferencia proporciones	N
Base de evaluaciones	31,4%	68,6%	-37,2	138
Encuesta a evaluadores	28,7%	64,8%	-36,1	122

Las dos muestras mantienen un comportamiento muy parecido. Los datos revelan que entre el 68,6% y el 64,8% de las evaluaciones estudiadas son de programas de carácter selectivo; y entre el 31,4% y el 28,7% restante responde a intervenciones universales. Las diferencias entre unas y otras están en torno al 36% o al 37%. Dicho de otra forma, por cada evaluanda de carácter universal se encuentran

aproximadamente 2,2 evaluaciones de políticas selectivas.

Por último, se cruzaron las dos variables, “universal-selectiva” y “sector político”, encontrándose una alta asociación entre las dos variables en las dos muestras⁹. Es decir, la variable “carácter universal o selectivo de las intervenciones” explica en parte la

⁹ Se ha propuesto un análisis de contingencia con residuos tipificados estandarizados. El análisis de Chi – cuadrado en la base de datos de evaluaciones y en la encuesta a evaluadores arroja un nivel de significación de 0.00. Se han clasificado las categorías de las variables en función de que obtuvieran un residuo por encima de 1.96. En los casos que los valores se encontraban entre - 1.96 y + 1.96 se han clasificado en la categoría “iguales” y se ha considerado que no hay diferencias significativas.

desproporción de evaluaciones entre los diversos sectores. En los análisis de contingencia realizados se han solicitado también residuos tipificados

estandarizados para ver la asociación en cada una de las categorías. Los resultados se han sintetizado en las tablas que aparecen a continuación.

Tabla 2

Base de datos de evaluaciones		
Selectiva Aquellos sectores que presentan más estrategias selectivas de lo esperado	Iguales No hay diferencias significativas	Universal Aquellos sectores que presentan más estrategias universales de lo esperado
<ul style="list-style-type: none"> - Asuntos Exteriores y Cooperación Internacional al Desarrollo - Empleo e Inmigración - Igualdad de género - Servicios Sociales 	<ul style="list-style-type: none"> - Industria, Turismo y Comercio - Ciencia y Tecnología - Salud 	<ul style="list-style-type: none"> - Economía - Educación y Deporte - Administración Pública - Medioambiente, Agricultura y Pesca - Cultura - Vivienda

Tabla 3

Encuesta a evaluadores		
Selectiva Aquellos sectores que presentan más estrategias selectivas de lo esperado	Iguales No hay diferencias significativas entre las estrategias	Universal Aquellos sectores que presentan más estrategias universales de lo esperado
<ul style="list-style-type: none"> - Igualdad de género - Servicios sociales 	<ul style="list-style-type: none"> - Empleo e inmigración* - Asuntos exteriores y cooperación internacional - Justicia - Educación y deporte - Industria, Turismo y Comercio - Administración pública - Salud - Vivienda 	<ul style="list-style-type: none"> - Economía - Seguridad interior - Fomento - Ciencia y Tecnología - Cultura

*El valor del residuo es 1,9, muy próximo del 1,96.

Los sectores que aparecían en el Gráfico 1 como los más evaluados son clasificados en las dos muestras o en alguna de ellas como aquellos donde se encuentran más estrategias selectivas de intervención

de lo esperado. Hay dos sectores que no se ajustan a este patrón. En el caso de salud, que ya presentaba diferencias en los datos de las dos muestras, no se encuentran diferencias significativas entre una u otra

estrategia. El otro caso es Educación que siendo el quinto sector más evaluado, se encuentra clasificado en una de las muestras como universales¹⁰.

En el otro extremo, las áreas menos evaluadas como fomento, seguridad interior, vivienda, o economía aparecen en las dos muestras o en alguna de ellas como sectores que se apoyan más en estrategias universales. Justicia solamente aparece en una muestra sin indicar una tendencia clara y sobre defensa no hay información.

Teniendo en cuenta las excepciones mencionadas, se detecta cierta tendencia donde los sectores más evaluados son aquellos que ejecutan sus políticas apoyándose más en estrategias selectivas y, al contrario, los menos evaluados se apoyan más en estrategias universales.

En definitiva, las conclusiones que se pueden extraer de este el análisis es que el hecho de que haya más evaluaciones de las políticas de carácter selectivo que de las universales incide en la desigual distribución de evaluaciones por sectores, ya que ciertas áreas tienden a apoyarse más en un tipo de estrategia de intervención que en otro.

Por lo tanto, debe haber algo relacionado con la estrategia selectiva de intervención que invita a hacer evaluación o, al contrario, que ante estrategias universales se dé un proceso que desincentive la evaluación. Esta es la nueva cuestión que abren los datos.

4.3. DEBATE

La existencia de tendencias que hacen que se evalúe más un tipo de programas que otros es algo que

ya ha sido descrito en la bibliografía de evaluación. Carol Weiss explica que durante los años 60 y 70, los programas que en mayor medida fueron sometidos a evaluación y escrutinio fueron aquellos que iban destinados a grupos marginales, como niños o niñas de clase baja, delincuentes juveniles y personas con problemas psíquicos; mientras que otros programas, los que ella denomina tradicionales, siguieron en marcha sin que cuestionara si estaban funcionando correctamente o no (1991:222).

Otro ejemplo son los programas de gastos fiscales indirectos en EEUU (Tax Expenditure Programs). Según los autores Datto y Grasso (1998), este tipo de gastos fiscales indirectos, que tienden a beneficiar a los más ricos, comparten una serie de características: sus importes se fijan con una gran flexibilidad, su distribución no está sujeta a la competencia por criterios de selección sino que son gastos que siempre se aprueban y, hasta hace poco tiempo, este tipo de programas no habían sido evaluados más que en raras ocasiones. Las características contrarias se aplicarían entonces a los programas de gastos directos que van dirigidos a grupos sociales más pobres (Datto y Grasso, 1998:1).

En ambos casos, las intervenciones más evaluadas son aquellas dirigidas a poblaciones en situaciones de dificultad. Weiss considera que esto se debe en parte a que los programas dirigidos a estas poblaciones no están respaldados por un grupo de presión que los defienda. En estos casos, según la autora, probablemente no existan grupos de interés, asociaciones profesionales, grupos de ciudadanos organizados, ni grupos de funcionarios que presionen por eludir las consecuencias que se derivarían de la evaluación (1991:222).

¹⁰ En esta categoría se generaron dudas sobre dónde clasificar las evaluaciones del sistema escolar y finalmente se atribuyeron a Universales.

Hay ciertas similitudes entre lo que plantean estos autores y lo que se expone en esta investigación para el caso español. Según las referencias citadas, los programas que se evalúan más son los dirigidos a estratos de población con menos recursos, a grupos marginales o que no tienen capacidad de influencia política. Por definición son programas que sólo intervienen con una parte de la sociedad, lo que permitiría clasificarlos como políticas selectivas.

La explicación que se propone en esta investigación es que las políticas universales y selectivas provocan reacciones diferentes y, por consiguiente, demandas diferenciadas de evaluación en el conjunto de la sociedad, pero también en los decisores políticos que encargan evaluación.

El hecho de que no todo el mundo se pueda beneficiar de un programa o servicio público puede predisponer hacia una actitud de recelo – si tengo los mismos derechos que los demás ¿por qué no puedo beneficiarme de este ingreso de integración, de ese piso de realojo o de estos cursos de formación?-. La inquietud que se manifiesta es más profunda que una simple duda teórica; tiene que ver con la sospecha de que se pueda estar ante una política injusta y que el criterio técnico que decide la selección resulte arbitrario. Se genera la suspicacia de que quien se beneficia del dinero público pueda no reunir las condiciones establecidas.

A tanto llega la desconfianza que, en ocasiones, la preocupación por el “fraude” desplaza a los objetivos propios de las políticas y se convierte en su primera finalidad. Se da el caso, por ejemplo, de algunos programas de rentas mínimas que establecen un

férreo y burocrático sistema de control que consume más recursos de lo que sería aconsejable, dificultando el logro de sus metas¹¹.

La duda sobre la veracidad de la condición del beneficiario no es algo nuevo ni responde a una mayor tecnificación política. En el caso de las políticas para combatir la pobreza se ha encontrado la misma preocupación al menos desde el siglo XVIII. Existía ya una inquietud por distinguir entre “pobres reales” y “pobres falsos” indistintamente de que ambos vivieran en la calle, mendigaran y compartieran las mismas condiciones de vida. Antes de dar una limosna o de otorgar ciertos bienes era necesario desentrañar si realmente las personas se habían visto empujadas a la precariedad o si, por el contrario, estaban en condiciones de trabajar y no querían hacerlo.

La reforma de la asistencia social exigía, como paso previo, que fuesen desenmascarados los falsos mendigos, sólo de esta forma podría cada ciudad sostener a sus propios pobres: si existía algún modo fehaciente de establecer en cada caso si se trataba de pobreza real o fingida. En consecuencia, el objetivo de muchas de las obras que se ocupan del tema consistirá en poner en evidencia a los falsos mendigos – esto es, a todos aquellos que podrían trabajar si quisieran – para poder de ese modo dedicar los recursos existentes a atender y asistir a los “verdaderos” pobres, es decir “aquellos que no pueden ganarse por sí la subsistencia” y “verdaderamente deben (sic) mendigar” (Cabrera, 1998: 30).

Actualizando los términos, se puede encontrar la misma preocupación en programas contemporáneos de subsidios de desempleo, realojo o asistencia

¹¹ Hay excepciones a esta lógica y algunos responsables entienden que hay que asumir ciertos porcentajes de “fraude” (que han estimado alrededor del 5%) para evitar una excesiva burocratización e incremento de costos.

domiciliaria, entre otros. En realidad, lo que está detrás de esta distinción es la persistente duda y sentimiento de fraude que parecen llevar aparejados los programas que dotan de bienes y servicios a sólo una parte de la población.

Estas reacciones se extreman cuando las dificultades sociales y personales son difíciles de discernir o de percibir a primera vista. Sirven como ejemplo los problemas de desarrollo social de los “hijos de la inmigración”, de las clases sociales bajas o las desigualdades entre hombres y mujeres provocadas por el sistema de género¹². En estos casos se pueden encontrar reacciones vehementes y argumentos que pueden llegar a negar la existencia de diferencias o, si se reconocen, a defender la no utilización de dinero público para estos fines.

Aunque esta interpretación está centrada sobre el caso español, también en otros países se encuentran reacciones similares ante la discriminación positiva. En EEUU puede llegar a tal grado la inquietud social sobre las acciones positivas que en algunos lugares se han llegado a prohibir por ley. En 2006, en el Estado de Michigan se consultó a la población a través de referéndum sobre la legalidad de este tipo de estrategias: un 58% votó en contra de su uso y como resultado se ha declarado ilegal la utilización por instituciones públicas de programas basados en discriminación positiva “por razón de raza, género, etnia o nacionalidad de origen” (International Herald Tribune 9/11/2006). Unos meses más tarde el Tribunal Supremo de Estados Unidos declaró que la política de discriminación positiva destinada a favorecer la integración racial en las escuelas es anticonstitucional (El País, 29/06/07).

La reflexión sobre la oportunidad y la eficacia de la discriminación positiva parece no venir al caso. Lo que preocupa es la estrategia, el cómo se hace, indistintamente de que los resultados puedan ser altamente satisfactorios y sirvan para reducir desigualdades sociales. En cambio, las políticas universales no despiertan ese celo legislativo aunque ahonden y agranden las dificultades que se pretendían erradicar.

La evaluación es una herramienta útil y legítima para conocer más sobre la intervención pública y rendir cuentas a la ciudadanía de su desempeño (Román, 2006). Pero el sistema de evaluación no es neutro y como manifiesta House la evaluación puede estar dirigida políticamente (2006:121).

El encargo de evaluación también lo es y forma parte del juego de influencias políticas. Patrizi and McMulla (Henry, 1998: 421) encuentran como resultado de su encuesta a líderes de fundaciones, que una de las mayores prioridades para financiar evaluaciones es la posibilidad de poder influenciar una intervención pública. En este sentido Weiss entiende que “existe una intención política implícita a la hora de seleccionar los programas que serán evaluados y aquellos programas que en cambio escapan al examen, (1991:221) y Chelimsky encuentra en algunos casos que las preguntas de evaluación esconden propósitos partidistas o ideológicos (1998: 404).

Por lo tanto, es probable que se encargue más evaluación en los casos en que exista un interés, motivación, conveniencia y necesidad de rebatir o confirmar la eficacia del programa entre los actores sociales con capacidad de decidir. No es extraño

12 Por sistema de género se entiende: “lo mismo que hay un modo de producción en cada sociedad, hay un sistema de género, que es el aspecto de la vida social que organiza las relaciones entre hombres y mujeres” (Britt-Marie Thuren, 1993: 97).

entonces que en este contexto se utilice la evaluación para analizar las políticas que generan más inquietud.

En la práctica, si una determinada intervención no genera dudas ni cuestionamientos es probable que los actores no demanden evaluación alguna, a no ser que sea requerida por otras instancias. Por lo tanto, si las acciones selectivas generan inquietud es probable que ésta se traduzca en un mayor número de evaluaciones en comparación con las políticas de carácter universal, que no provocan las mismas sospechas.

Lo paradójico de esta situación se da en la coexistencia de intervenciones de carácter universal con inversiones millonarias sin ningún sistema de evaluación frente a programas selectivos más humildes reiteradamente evaluados. En España se pueden desarrollar proyectos costosos como autovías en zonas de baja densidad de tráfico, túneles y soterramientos de vías, remodelación frecuente del mobiliario y paisaje urbano, subvenciones estatales para la compra de automóviles, multiplicidad de televisiones y radios públicas, presencia institucional en eventos culturales o construcción de grandes edificios emblemáticos que no sólo no son evaluados (salvo por las auditorías o controles contables) sino que tampoco surge una demanda pública manifiesta que así lo reclame. No se sabe nada sobre sus resultados, adecuación, priorización de las necesidades sociales o sobre el coste de oportunidades que han supuesto dichas inversiones.

Por el contrario, ciertos planes y programas como por ejemplo los de rentas mínimas, de apoyo a dependientes, de mujeres prostituidas, de formación a determinados colectivos, de reinserción social,

de apoyo a los desempleados o subvenciones a la creación artística, indistintamente de su coste, suelen soportar una mayor presión evaluativa.

En cualquier caso, la evaluación no debe ser un castigo: es una buena noticia disponer de evaluaciones. En teoría, sus efectos constructivos se deberán hacer notar y habrá una actuación más sensible y responsable por parte de los profesionales y decisores políticos. Según este planteamiento, las perjudicadas por la ausencia de evaluación serían las políticas de carácter universal por no disponer tanto como las otras de la retroalimentación de la evaluación. Pero, aunque en términos absolutos la evaluación sistemática supone una ganancia, existe un cierto riesgo en el desarrollo desigual de la evaluación se detallan a continuación.

4.4. RIESGOS

Un ejercicio de evaluación debe partir de la posibilidad de que la valoración del programa pueda ser negativa. Por lo tanto, los programas más evaluados están más expuestos a este tipo de enjuiciamiento que los que no se evalúan. Utilizando las palabras de Weiss los programas evaluados airean sus trapos sucios, mientras que otros programas mantienen su información guardada en el armario (1991: 221).

Cuando se airean los resultados negativos en público es más fácil que el siguiente paso puede ser la restricción, denegación de fondos o el cierre del programa. Se pueden encontrar variados ejemplos de este tipo. En la Comisión Europea, la unidad de la Dirección General de presupuestos argumenta la utilidad de los procesos de evaluación poniendo el ejemplo del cierre del programa de desayunos

escolares¹³. En EE.UU. cuando el presidente Ronald Reagan “asumió el cargo, su programa político se basaba en reducir el gobierno, por lo que, con este propósito, se pidió a los evaluadores que encontraran ineficiencias” (House, 2006: 120). Por su parte, el Presidente George W. Bush, anunció su intención de reducir o eliminar más de 150 programas públicos que no estaban consiguiendo resultados. Como Bush advirtió aquellos programas que no pudieran ser obligados a rendir cuentas de su desempeño serían reducidos o eliminados (Renger, 2006).

No se está cuestionando que se puedan cerrar o limitar programas que no funcionan sino que la mayor exposición de las políticas selectivas al análisis público las hace más vulnerables. Mientras tanto, las políticas o intervenciones menos evaluadas, es decir, las de carácter universal, permanecen más alejadas de esta dinámica sin que sean cuestionadas. Volviendo a utilizar ejemplos de EEUU, Monnier (1992) explica como la evaluación fue utilizada por la administración Nixon para retrasar la ejecución de medidas sociales impuestas por el Congreso y reducir el número de programas sociales innovadores.

Ante estas situaciones no son extrañas las reacciones de alejamiento y blindaje ante la evaluación cuando los responsables de intervenciones piensan que la evaluación no les va a recompensar positivamente (Newcomer, 2004). Viene a colación el fenómeno detectado por Chelimsky; entre 1980 y 1994 la autora percibió un aumento de la información secreta y clasificada en un gran número de agencias, “la ironía es que la amenaza que esta información planteaba no tenía tanto que ver con la seguridad nacional sino más bien tenían que ver con la cara oculta de las agencias independientes” (2008:407).

Igual que ciertos sistemas de evaluación pueden producir efectos perversos (Leeuw and Furubo, 2008), lo mismo ocurre con el desigual desarrollo de la evaluación, puede tener un efecto penalizador de determinadas políticas.

Teniendo en cuenta la situación actual de la evaluación en España, el riesgo es que se estén penalizando las acciones públicas de carácter selectivo o basadas en la discriminación positiva por su mayor exposición. La consolidación de esta dinámica podría contribuir a largo plazo a conformar una determinada concepción política que, apoyada en un discurso técnico, se caracterizaría por potenciar las intervenciones lineales en las que todo el mundo recibe lo mismo indistintamente de las dificultades o necesidades sociales que se soporten.

Pero además, la relación existente entre las estrategias (universal y selectiva) y los sectores políticos -el hecho que sectores como justicia, fomento, interior o defensa suelen actuar apoyándose en estrategias universales y asuntos sociales, trabajo, inmigración, e igualdad necesiten dirigirse selectivamente a grupos específicos-, supone indirectamente una merma de estos últimos sectores políticos, los que además se asocian con las políticas de bienestar social.

Tal y como está planteado, el efecto de la evaluación contribuiría a un modelo de Estado tecnócrata. Fuerte en justicia, defensa, interior e infraestructuras pero débil y descapitalizado en lo social. Un Estado en el que si se considera técnicamente imprescindible intervenir con algún colectivo específico se hará, aunque lo probable es que sea una intervención acotada, limitada y “vigilada” a través de diferentes mecanismos entre los que se encuentra la evaluación.

13 Caso citado por Eduardo Zapico en la ponencia sobre la Evaluación en la Comisión Europea en el VI Seminario de Experiencias de Evaluación, Magíster en Evaluación de Programas y Políticas Públicas de la Universidad Complutense de Madrid, 28/09/09.

4.5. RECOMENDACIONES

Aunque las conclusiones finales de esta memoria de investigación hacen vislumbrar que podrían ser generalizables a otras realidades es necesario apoyarse en datos concretos para saber si ocurre lo mismo en otros países. Hay indicios que así lo parecen apuntar. Los estudios de Leeuw y otros, Noche y Poincard y el análisis de las Solicitudes de Propuestas de evaluación, muestran a primera vista tendencias similares al caso español.

1. Por lo tanto, la primera recomendación es una invitación a investigadores a analizar muestras, bases de evaluaciones, censos o documentos como las solicitudes de propuestas de evaluación en diferentes países para saber si se dan procesos parecidos.
2. La segunda recomendación se centra en corregir la situación actual del desarrollo desigual de la evaluación en España. El hecho de identificar algún efecto perverso no debe llevar a pedir menos evaluación; todo lo contrario, lo que se demanda es incrementar el campo de la evaluación para que todas las políticas- universales y selectivas - estén sometidas a las mismas pautas de aprendizaje y retroalimentación.
3. En la fase emergente del desarrollo de la evaluación en la que se encuentra España, es especialmente importante que la evaluación

no sea vista como una enemiga o como una estrategia fiscalizadora, sino que se asocie con una metodología constructiva, de ayuda y no de fiscalización. Se proponen dos opciones:

- ▶ Premiar a las organizaciones que son capaces de realizar el ejercicio de transparencia de la evaluación y aplicar sus recomendaciones indistintamente de que los resultados que se obtengan sean positivos o negativos. Esto es sobre todo importante en ONG u organizaciones que dependen de subvenciones. Organizaciones que encargan evaluación como la Agencia Española de Cooperación Internacional para el Desarrollo pueden incluir en sus pliegos el compromiso de no perjudicar a las organizaciones que confían en la evaluación.
- ▶ Supeditar subvenciones o donaciones a la espera de la realización de evaluaciones. En el caso de las organizaciones que persistentemente no son evaluadas, deberá ponerse en cuestión la otorgación de subvenciones o donaciones. Esta medida es aplicable a todos los sectores de intervención, pero especialmente a aquellos que han desarrollado menos la evaluación como son defensa, justicia, fomento, seguridad interior, vivienda o economía.