

JAVIER ESPARCIA PÉREZ*
JOAN NOGUERA TUR*

APROXIMACIÓN TEÓRICO-
METODOLÓGICA A LA CULTURA
EVALUATIVA Y LA EVALUACIÓN DE
PROGRAMAS DE DESARROLLO RURAL

RESUMEN

En el ámbito de las políticas públicas muchos países de nuestro entorno han desarrollado prácticas de evaluación que están ya muy asentadas en la conciencia colectiva de los diferentes profesionales. En España la evaluación de políticas públicas y programas ha de afrontar aún muchas resistencias, pero son cada vez más los que admiten que la evaluación ha de enfocarse no tanto como una intromisión sino como un instrumento de aprendizaje y de mejora de la eficacia en la toma de decisiones. La cultura evaluativa en España ha venido de la mano de las autoridades comunitarias, que han mostrado una preocupación creciente por conocer el impacto de las diferentes políticas estructurales y sus correspondientes programas. La evaluación es el instrumento que mejor permite conocer qué resultados se están obteniendo y, en su caso, introducir los mecanismos de corrección oportunos. Desde la perspectiva de la Comisión Europea, de medio-largo plazo, y más preocupada por el diseño de políticas eficaces que permitan alcanzar sus grandes objetivos estratégicos, la evaluación externa ha sido el único mecanismo riguroso y sistemático que permitía obtener información fiable. Sin embargo, las evaluaciones externas no son tan útiles para los propios actores de los programas o políticas; suelen acumularse retrasos importantes y, en algunas ocasiones, los resultados no llegan o llegan incluso filtrados a estos actores. Por ello durante los últimos años se está avanzando en la vía de las evaluaciones internas (participativas o no), que, aunque en principio no son tan rigurosas en cuanto a metodologías, sí permiten obtener respuestas más rápidas y adaptadas a las necesidades de los actores locales.

ABSTRACT. THEORETICAL AND METHODOLOGICAL APPROACH TO THE CULTURE OF EVALUATION AND
THE EVALUATION OF RURAL DEVELOPMENT PROGRAMMES

In many Western European countries, the evaluation of public policies is a long-established practice, well rooted in the consciousness and know-how of all involved actors. In Spain, the evaluation of public policies and programmes is still facing several constraints. However, there is an increasing concern on the need for considering evaluation procedures not as an interference but as a learning tool essential in order to improve effectiveness in decision-making. The culture of evaluation in Spain has been brought about by EU authorities and regulations, showing an increasing

* UDERVAL (Unidad de Desarrollo Rural y Evaluación de Políticas Públicas). Departament de Geografia. Universitat de València.

interest in the real impact of the Structural policies and their corresponding programmes. The evaluation is the best tool to gain knowledge on the actual results of a policy and this knowledge, in turn, allows to introduce the necessary corrections in future policies. From the EU Commission approach to policy making, i.e. medium-long term, effective policies to achieve its big strategic objectives, the external evaluation has been the only feasible procedure in order to obtain confident information. However, external evaluations are not so useful when taking about the own policy actors since it produces important delays and, in some cases, the results of the evaluation process either are filtered or do not reach all actors. In consequence, during last years internal evaluations (participative or not) are being the main way forward because they provide quick and more adapted results.

CULTURA EVALUATIVA Y AUTOEVALUACIÓN

La evaluación: algo habitual en la vida diaria, pero ... ¡grandes resistencias a ser evaluado!

La lógica de la evaluación está presente en la mayor parte de la vida diaria. Constantemente estamos tomando decisiones, y para ello ponemos en práctica, de manera intuitiva, procesos de evaluación que nos permitan valorar mejor los diferentes elementos que intervienen en esa toma de decisiones. Desde este punto de vista esta evaluación es un mecanismo previo a unas decisiones, y por lo tanto está también orientada a mejorar unas acciones futuras.

Sin embargo, tradicionalmente el término evaluación no se ha aplicado a este tipo de situaciones, a modo de valoración previa. Su uso más habitual ha sido, y casi sigue siéndolo, como una valoración posterior a algo. También aquí la vida diaria es un ámbito en el que constantemente se está valorando o "evaluando", como ocurre con las decisiones tomadas recientemente o durante el día, o la actividad que se acaba de finalizar.

Pero si este tipo de "evaluación" está tan presente en la vida diaria, sorprende que en la sociedad actual no hayamos sido capaces de desarrollar una cultura de la evaluación. Varios elementos, íntimamente relacionados con la psicología individual, nos dan las claves de por qué esto es así. En primer lugar, porque todos asumimos evaluar a alguien o algo, e incluso a nosotros mismos; sin embargo, otra cuestión muy diferente es que nos evalúen a nosotros, nuestro trabajo o nuestra organización. En segundo lugar, los mecanismos de toma de decisiones suelen tener, incluso mucho más allá de la vida diaria, un carácter muy intuitivo; si hacemos un pequeño examen comprobaremos que la mayor parte de nuestras decisiones en la vida laboral y profesional se han guiado, en parte, por la intuición.

No es esto necesariamente negativo, si esa intuición está basada en una experiencia de toma de decisiones, o acompañada con un buen asesoramiento, o se dispone para ello de la información adecuada y necesaria (ORGEON y MEDINA, 1997). Desde el punto de vista de psicología individual, estos procesos de toma de decisiones, con un fuerte carácter intuitivo, poco formalizados, restan "receptividad" a cualquier tipo de evaluación formal impuesta desde fuera. En definitiva, la evaluación es vista o sentida, de manera consciente o inconsciente, como una agresión o intromisión en nosotros o en nuestro trabajo, y de ahí que se desarrollen mecanismos en contra de todo intento de ser evaluados desde fuera. Tal vez todo ello explique, en el tema que nos ocupa, que las evaluaciones externas de programas de desarrollo o, en general, de la mayor parte de políticas estructurales, no suelen ser bien recibidas por los protagonistas de las mismas.

Hacia la cultura de la evaluación

No cabe duda a estas alturas que una evaluación sistemática y formalizada es de gran utilidad para mejorar las decisiones referidas a programas o acciones a aplicar tanto en un futuro próximo como a las que ya han tenido lugar (WEISS, 1988; Conseil Scientifique de l'Évaluation, 1996). Si mejoramos estas decisiones, evidentemente esto nos permitirá no sólo conocer mejor cuáles han sido los aspectos positivos y negativos, sino cuáles son los condicionantes y el contexto en el que podemos conseguir un mayor impacto de esas programas o acciones, es decir, podremos satisfacer mejor nuestras necesidades o las de los beneficiarios.

Conscientes de las grandes virtudes de la evaluación, en la mayor parte de sociedades avanzadas se ha desarrollado una verdadera cultura de la evaluación; esta cultura significa, por un lado, que las autoridades públicas, en el ejercicio de su responsabilidad, han desarrollado programas y mecanismos de evaluación que les permiten minimizar riesgos en sus decisiones a la vez que aumentar los beneficios o impactos positivos sobre los beneficiarios, todo ello teniendo en cuenta igualmente unos criterios de eficiencia en el uso de los recursos invertidos en los diferentes programas o acciones. Por otro lado, la cultura de la evaluación significa que los responsables de programas o acciones son cada vez más conscientes de que la evaluación, mucho más que un examen al que se le somete con cierta frecuencia, constituyen un instrumento de aprendizaje y una herramienta de mejora en su trabajo (ALVIRA, 1991; COHEN y FRANCO, 1993).

Desgraciadamente la cultura de la evaluación está aún poco presente en España. La obligatoriedad que impone la Comisión Europea de evaluar muchos de sus programas está contribuyendo cuando menos a que se sepa que es bueno evaluar, aunque aún no se acepte de manera abierta el ser evaluado. Estas evaluaciones permiten –o deberían hacerlo– en primer lugar conocer cómo se están invirtiendo las grandes cantidades de ayudas que proporciona, y qué efecto están teniendo (SOY, 1995). En segundo lugar, constituyen –o deberían hacerlo– el *input* fundamental de cara al diseño de cada nuevo programa. Y así es como efectivamente empieza a trabajarse, al menos desde la perspectiva de la Comisión. También en este ámbito la Comisión Europea va muy por delante de lo que ha venido siendo la práctica en las administraciones públicas españolas. En cualquier caso, lo importante es que la cultura de la evaluación empieza a extenderse, y en el marco de los programas de desarrollo rural probablemente por ello podemos hablar no sólo de evaluación sino también de autoevaluación.

La autoevaluación: un ejercicio de responsabilidad

Desde un punto de vista conceptual es erróneo hablar de autoevaluación como contraposición a la evaluación. La autoevaluación no es más que la evaluación que hacen los propios responsables o los que participan en la aplicación de un programa o el desarrollo de unas acciones (European Commission, 1999). Debe así diferenciarse de manera clara de una evaluación externa, que es puesta en marcha bien por los administradores o gestores de un programa, o bien por los propios técnicos. Desde este punto de vista, se trata de una evaluación interna, que puede ser o no participativa en función de la implicación o no de los usuarios (GABARRÓN y HERNÁNDEZ, 1994).

Aunque tiene determinadas características que la hacen más apropiada para algunas situaciones, en principio la autoevaluación puede adquirir las mismas formas que una

evaluación normal en cuanto al momento de realización (ex-ante, intermedia, ex-post, continua, etc.), y también puede ser global, temática, o centrarse en algún aspecto de interés (funcionamiento interno de la organización y papel de los diferentes miembros o secciones en la gestión e implementación del programa, etc.). Posteriormente entraremos más en detalle en las ventajas e inconvenientes de la autoevaluación para cada objetivo o elemento a evaluar.

La autoevaluación, en relación a otros procesos diferentes de evaluación, es a la vez un punto de llegada y un punto de partida. En primer lugar, es un punto de llegada en la medida en la que cuando se es plenamente consciente de las virtudes de la evaluación, ésta se lleva hasta sus últimas consecuencias con el diseño de unos mecanismos de evaluación propios (complementarios de una probable evaluación externa formal), formalizados y sistematizados, que permitan un conocimiento y un seguimiento más preciso de los diferentes aspectos o elementos evaluables, y que probablemente no pueden ser valorados con suficiente profundidad, o con la continuidad adecuada, por una evaluación externa o coyuntural. En segundo lugar, la autoevaluación puede ser un punto de partida en aquellas situaciones en las que una organización o grupo de personas, responsables de un programa o unas acciones, deciden poner en marcha unos mecanismos de evaluación propios que les guíen de manera más precisa en sus decisiones, en su gestión, o en la implementación del programa o acciones.

Tanto si responde al deseo de completar la información derivada de otras evaluaciones, como si responde a una iniciativa propia, de querer conocer más y mejor cómo hemos trabajado, qué impacto tiene nuestro trabajo, si responde o no a los objetivos, o si responde o no a las necesidades de los usuarios, etc., lo más importante es que en la autoevaluación está implícita la decisión libre y responsable, de cada uno de los protagonistas o actores respecto de su trabajo y de los que son sus destinatarios. La gran virtud de la autoevaluación está, por tanto, en que es un ejercicio de responsabilidad individual y a la vez colectiva, del grupo humano que está implicado en ella.

LA LÓGICA DE LA EVALUACIÓN Y SUS ELEMENTOS CLAVE

¿Qué es evaluar?

Desde hace algunos años el término evaluación se ha extendido con gran profusión, llegándose incluso a un cierto abuso; el resultado es que se denomina evaluación a procedimientos que no lo son en sentido estricto. Por ello vamos a intentar precisar un poco más en torno a qué es hacer una evaluación. No es ésta una cuestión fácil de responder como a *priori* podría pensarse. Empecemos por indicar qué no es evaluación, y será más fácil definir qué sí lo es.

- En primer lugar, una evaluación no es un estudio científico con carácter académico; la evaluación ha de tener una visión muy práctica, que habitualmente no tienen tales estudios.
- En segundo lugar, una evaluación tampoco es una auditoría, en la que se analiza la legalidad o adecuación de los procedimientos utilizados, las decisiones tomadas, o la gestión realizada en la implementación o en el uso de los recursos financieros disponibles en el programa o proyecto.

No hay una definición única del término *evaluación*, válida para todas las situaciones en las que se hace alguna valoración de algo. Algunas de las definiciones que se asocian con el término *evaluación* son las siguientes:

- “La evaluación es el proceso de determinar los méritos, *valía e importancia de las cosas*” (European Commission, 1999);
- “Una visión crítica de los objetivos [de un programa o intervención pública] y cómo éstos se han alcanzado” (HM Treasury, 1988, Gobierno Británico, citado en European Commission, 1993);
- “Proceso por el que nos formamos un juicio sobre el valor de un programa” (VIVERET, 1989);
- “Juzgar las intervenciones públicas de acuerdo a sus resultados, impactos y las necesidades que éstas pretenden satisfacer” (European Commission, 1993);
- “Un análisis independiente de los objetivos, resultados, actividades y medios empleados [en o por un programa o intervención pública], tendente a extraer conclusiones que sean aplicables de manera más amplia” (European Commission, 1999);
- “Un proceso que trata de determinar de manera tan sistemática y objetiva como sea posible, la relevancia, eficiencia y efectos sobre una actividad, en términos de sus objetivos” (Naciones Unidas, citado en European Commission, 1997);
- “La aplicación sistemática de procedimientos de investigación social para valorar la conceptualización, diseño, implementación y utilidad de los programas públicos”;
- “El examen de si los medios legales, administrativos y financieros puestos en marcha por un programa permiten alcanzar los objetivos que tenía asignados y producir los efectos esperados” (Conseil Scientifique de l'évaluation, 1996).

Podríamos por tanto decir que evaluar es hacer un estudio con unos propósitos muy prácticos, la toma de decisiones en relación a la implementación de unas acciones, proyecto o programa. También podemos decir que evaluar es analizar de manera independiente acciones, programas, o políticas, valorando su conceptualización, diseño o implementación, pero especialmente su importancia en relación a su coste, su impacto, sus objetivos, sus resultados, o las necesidades que pretenden satisfacer. A diferencia de los estudios académicos, la visión práctica con la que se realiza una evaluación le permite contribuir de manera clara a la toma de decisiones, y al menos la información para tomar decisiones debe ser siempre un *output* de la evaluación; otra cuestión es la utilización que se haga de ella por parte de los responsables. Es decir, se trata de que la evaluación ha de ser algo útil, y precisamente la *utilidad* es un concepto que ha de estar presente en la evaluación de esas acciones, programas o políticas, en relación a unos estándares que vienen determinados por los objetivos, las necesidades o los resultados previstos.

Aunque las evaluaciones pueden diferir enormemente en función del objeto de análisis, los objetivos, el cliente, etc., hay una serie de características que deben estar presentes en prácticamente toda evaluación; ésta debe:

- ser analítica, es decir, estar basada en técnicas de investigación reconocidas;
- ser sistemática, por tanto llevada a cabo con una cuidada planificación y una utilización consistente de las técnicas que se han seleccionado;
- ser clara y fácilmente comprensible en su metodología, es decir, reproducible por un evaluador diferente con acceso a los mismos datos y utilizando las mismas técnicas de análisis;

- estar centrada en la aplicación, consecuencias y resultados de las acciones, actividades, proyectos o programas a evaluar;
- estar centrada en los elementos clave, dando una respuesta clara y precisa a las necesidades que tiene el destinatario de los resultados de la misma;
- útil, en tanto que debe aportar información útil para los tomadores de decisiones, dadas unas determinadas circunstancias políticas, restricciones del propio programa y recursos disponibles (HERMAN *et al.*, 1987; PAWSON y TILLEY, 1997).

Evaluación: ¿por qué y para qué?. La “lógica de la evaluación”

En muchos ámbitos de las políticas públicas, y también en el desarrollo rural, la evaluación (en cualquiera de sus procedimientos o tipos) responde a la necesidad de conocer la validez de una programa, proyecto o acciones, pasadas o previstas, para tomar decisiones en consecuencia (STECHEER y DAVIS, 1987; WEISS, 1988; SHULHA y COUSINS, 1997; VERDUNG, 1997; MENÉNDEZ, 1999).

Precisamente por ello no podemos hablar de evaluación de manera aislada, sino insertarla en un contexto en el que quede claro por qué evaluar y qué evaluar.

La secuencia que nos conduce a la evaluación la tenemos en la figura 1:

- En primer lugar tenemos unos problemas socioeconómicos o territoriales como punto de partida habitual. Igualmente estos problemas pueden referirse a aquellos aspectos a los que se pretende dar respuesta desde la institución u organización directamente implicada o responsable, en nuestro caso el Grupo de Acción Local (GAL).
- Esos mismos problemas definen ya de por sí una serie de necesidades, que en el caso de las políticas de desarrollo rural son las que tiene la población que vive y trabaja en esas zonas rurales.
- La consideración de estos problemas socioeconómicos, así como de las necesidades, se traducen, o deben hacerlo, en el diseño y puesta en práctica de un programa, un proyecto o un conjunto de actividades; es decir, acciones encaminadas a dar solución a esos problemas socioeconómicos y a satisfacer esas necesidades. Ese programa debe también diseñarse de manera que se aprovechen las oportunidades del entorno.
- Cualquiere programa, proyecto o acción encaminada a dar respuesta a los problemas y las necesidades, debe fijar necesariamente unos objetivos, en los que se priorice qué es más importante, y cómo ha de conseguirse; es decir, los objetivos van a permitir diseñar un “plan de acción” para dar esa respuesta. Los objetivos pueden ser de tres tipos, generales o estratégicos, operativos y específicos.
- Los objetivos a su vez se traducen en unas acciones concretas, en unas medidas de actuación, que se aplican en función de los objetivos, los medios disponibles, etc.
- Es de suponer que tales acciones tienen unos resultados, a corto plazo o a medio y largo plazo. A su vez, los resultados u *outputs* pueden generar unos impactos, tanto directos como, sobre todo, indirectos (en algunos casos no previstos y a muy largo plazo). En teoría el programa busca como objetivo último tener un impacto positivo y generar un valor añadido sobre el área, organización o grupo humano de referencia. En el caso de las políticas de desarrollo rural este valor añadido suele estar referido al contexto territorial en el que se aplica el programa.
- La fase última en esta secuencia lógica es analizar si esos resultados se corresponden con lo que estaba previsto, con las necesidades de la población, si permiten resol-

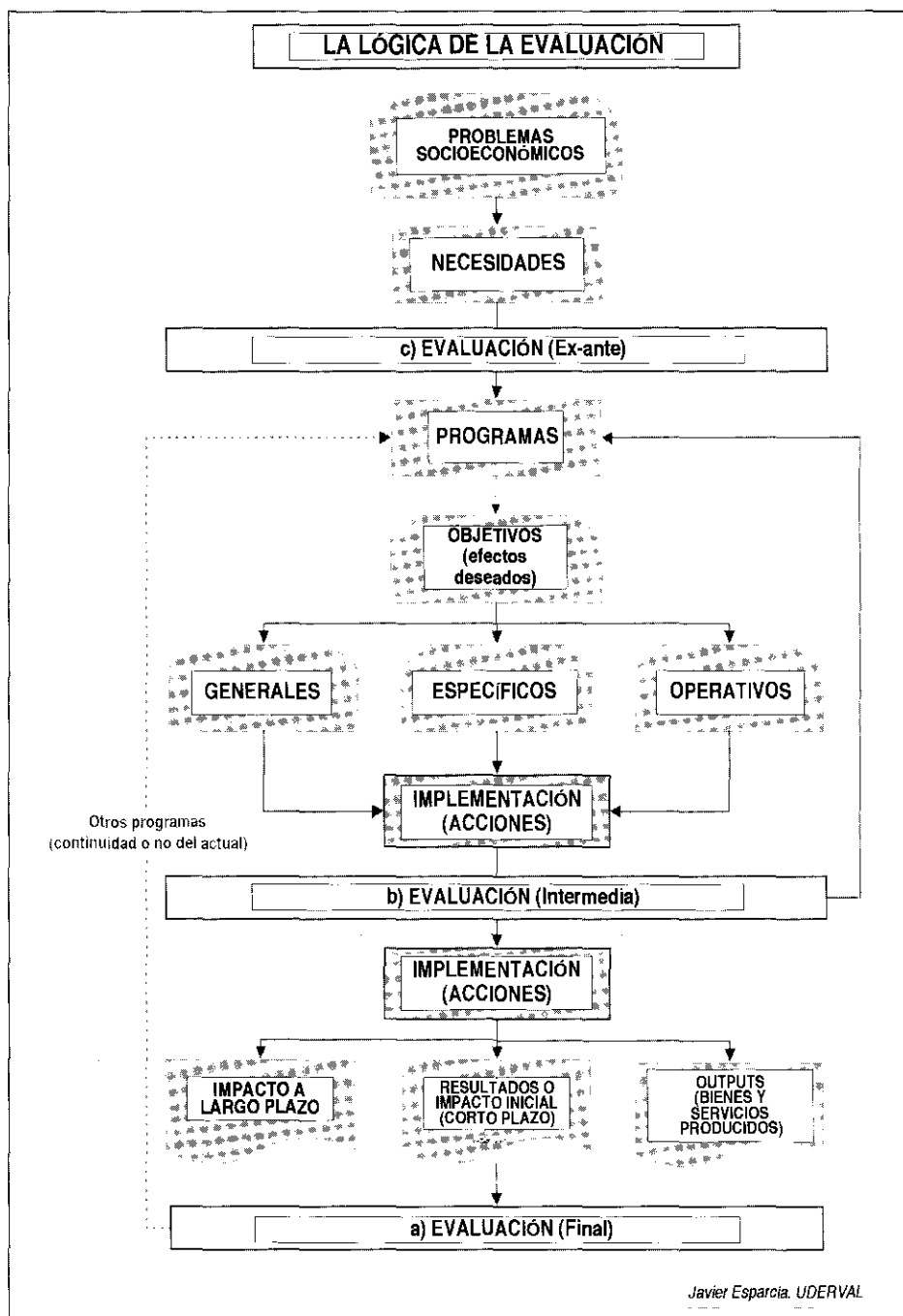


Figura 1. La lógica de la evaluación en los programas de desarrollo.

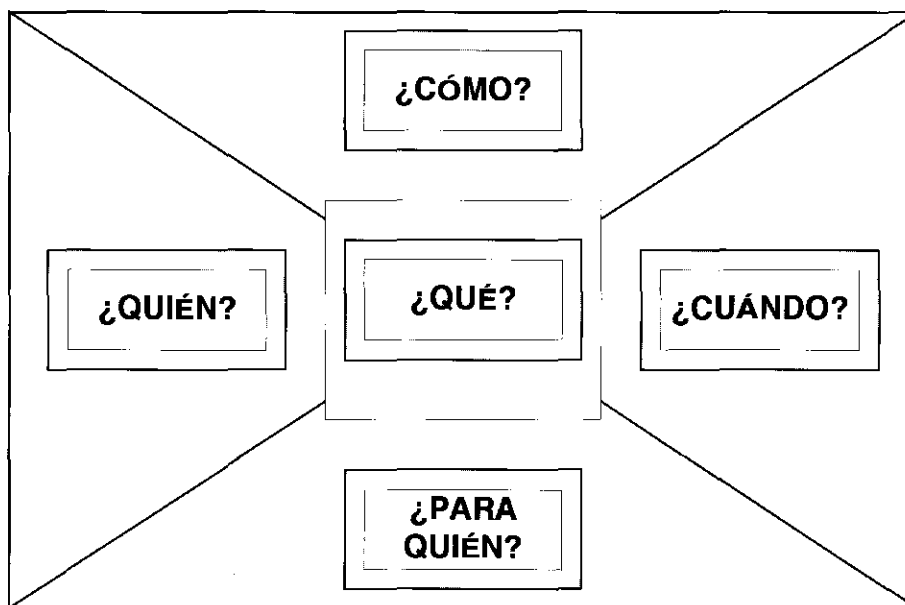


Figura 2. Cinco cuestiones clave en evaluación.

ver o contribuyen a la resolución de los problemas socioeconómicos o de otro tipo (MENÉNDEZ, 1999). Aunque en el mundo de los sistemas físicos es relativamente fácil controlar que todo el proceso se desarrolla según lo previsto, en los sistemas socioeconómicos lo extraño e inusual es que eso ocurra. Es decir, hay toda una serie de desviaciones, por una multitud de factores, que tienen como resultado que la adecuación entre lo previsto y lo conseguido no sea exactamente lo mismo, por no hablar de los efectos o resultados indirectos, que son francamente difíciles de controlar. En definitiva, si queremos conocer cuál es el grado de desviación entre el programa, las necesidades, los propios problemas socioeconómicos, y los resultados obtenidos, hay que acudir necesariamente a la evaluación (European Commission, 1993).

La secuencia anterior responde a una situación en la que cada fase sucede a la anterior en el tiempo. Sin embargo la evaluación puede también tener como objetivo valorar los condicionantes, las posibilidades, o la viabilidad, de aplicar un programa. En este caso no se dispone de unos resultados, pero sí es posible tomar como punto de partida las dos primeras fases, es decir, un buen análisis de los problemas socioeconómicos por un lado, y un buen análisis de las necesidades, por otro. Evidentemente en este tipo de evaluación —ex ante— se incluye el análisis y valoración de todos los elementos que justifican la puesta en marcha del programa o el cumplimiento de los requisitos exigidos por el mismo. De la misma manera, entra dentro de la “lógica de la evaluación” querer disponer de información precisa de cuál es el grado de cumplimiento del programa durante su aplicación, de manera que los responsables tengan capacidad de introducir ajustes en el mismo. Más adelante abordamos estos aspectos.

Diversidad tipológica de evaluaciones

Hasta aquí tenemos clara la lógica de una evaluación, a qué responde y su utilidad. Sin embargo, tomada la decisión de hacer una evaluación, resta una tarea importante y decisiva que es definir qué tipo de evaluación se adecua a nuestras necesidades; es decir, qué le pedimos a una evaluación (ROSSI y FREEMAN, 1979; VIVERET, 1989; VERDUNG, 1997). Para ello podemos plantearnos cinco preguntas básicas, que nos van a ayudar a definir qué tipo de evaluación es más adecuada. La decisión de qué evaluar, aunque parezca la más sencilla, es probablemente la más compleja desde un punto de vista metodológico.

1. La decisión de qué evaluar: objetivos y tipos de criterios básicos

Son muchos los aspectos que pueden evaluarse en un programa o proyecto. Sin embargo es necesario centrarse en primer lugar en un análisis de la naturaleza del programa o proyecto, atendiendo especialmente a *qué* es lo que se esperaba alcanzar con el programa o proyecto. Básicamente esto se traduce en un análisis de los diferentes objetivos, en sus diferentes niveles (estratégicos, operativos y específicos). El por qué es importante profundizar en el *qué* es debido a que constituyen la expresión del problema al que se le pretende hacer frente y de las necesidades que se trata de abordar. Anteriormente se ha señalado la importancia de los tres tipos de objetivos como punto de referencia para definir qué queremos evaluar.

- Objetivos generales o estratégicos. Se expresan en términos de impacto a largo plazo del programa (ej. creación de empleo en un área);
- Objetivos operativos. Se expresan en términos de outputs, es decir, aquello que directamente produce el programa (ej. cursos de formación profesional para los desempleados de larga duración);
- Objetivos específicos. Se expresan en términos de resultados, es decir, el impacto inicial o a corto-medio plazo del programa (mejorar las posibilidades de empleo de los desempleados de larga duración aumentando su nivel de preparación).

Habitualmente interesará conocer si han cumplido las expectativas tanto en relación a los objetivos (estratégicos y específicos) como a la manera de conseguirlo (objetivos operativos), o los resultados directos obtenidos (objetivos específicos). En relación a estos tres tipos de objetivos, podemos hablar de dos grandes tipos de criterios, los referidos al propio programa o proyecto (relevancia y eficiencia), y los referidos a la relación entre el programa y los problemas socioeconómicos que se pretenden abordar (eficacia o impacto, utilidad y sostenibilidad). La evaluación puede estar centrada en cualquiera de estos cinco tipos de criterios, pero tendremos una visión muy incompleta del programa o proyecto si sólo se evalúan uno o dos de estos aspectos. Por lo general, se ha de intentar incorporar a la evaluación tantos como sea posible y, como mínimo y para evaluar programas y proyectos de desarrollo local y rural, ha de tenerse una visión de la eficacia y de la eficiencia de los mismos. Veamos a continuación en qué consisten cada uno de ellos.

- Relevancia: responde a la cuestión de en qué medida los objetivos del programa o proyecto están en relación y se adecuan a las necesidades del área, sector o grupo de población sobre el que se va a aplicar. Este criterio puede ser importante para tomar

decisiones sobre la continuidad del programa, la necesidad o no de introducir cambios y, en su caso, la profundidad de tales cambios.

- **Eficiencia:** es la relación entre el esfuerzo invertido en términos de recursos físicos, humanos o financieros, y los resultados obtenidos, es decir, está referido al coste del programa en relación a cada unidad de output generada (¿podrían haberse obtenido los mismos resultados empleando menos recursos?, ¿se podrían haber obtenido más y/o mejores resultados con los mismos recursos?). *A priori* este es un criterio relativamente fácil de medir; sin embargo, la principal dificultad estriba en disponer de un nivel de referencia con el que poder comparar para saber cuál es el significado real del nivel de eficiencia obtenido en el programa o proyecto evaluado. Cuando se evalúan diferentes proyectos en el contexto de un programa (ej. en un área LEADER), o diferentes áreas con el mismo programa (ej. diversos grupos o áreas LEADER de una misma Comunidad Autónoma o del mismo país), como mínimo se podrá tener un valor medio de eficiencia en relación al cuál podremos poner cada uno de los proyectos o cada una de las áreas.

- **Eficacia:** responde a la cuestión de en qué medida el impacto-s o resultados del programa o proyecto han permitido alcanzar los objetivos generales y específicos previstos. Un programa o proyecto puede ser eficiente en términos de coste, pero con logros o un impacto muy pobre en relación a los objetivos previstos. Habitualmente, cuando se diseña un sistema de indicadores de eficacia antes de la evaluación, e incluso para hacer un seguimiento del programa o proyecto, se piensa principalmente en los efectos positivos que se esperan del mismo. Sin embargo, pueden haber también efectos positivos no esperados, e igualmente efectos negativos, esperados o no. Por tanto, se ha de intentar que en el sistema de indicadores todos estos aspectos estén contemplados para poder analizar una valoración del impacto global, incluyendo aquellos efectos que pueden deberse a factores causales. De cualquier manera esto último no es más que una complicación añadida a un criterio, la eficacia, que es de por sí muy complejo de medir. Los problemas pueden ir desde una definición vaga o genérica de los objetivos, hasta, sobre todo, dificultades técnicas para medir el impacto real con precisión (ej. ¿cómo podemos medir con cierta precisión el grado de innovación que se deriva de la incorporación de un proceso productivo en un área?; ¿cómo podemos medir la diferencia en el grado de innovación de un proceso productivo en ambas áreas?; ¿cómo podemos medir el impacto que la introducción de ese proceso ha tenido en ambas áreas?; ¿cómo podemos tener la certeza de que el empleo que se ha creado depende exclusivamente de una ayuda pública?, ¿qué porcentaje de explicación tiene esa ayuda a la hora de retener o crear ese empleo? ...).

- **Utilidad:** valorar los resultados en relación a los objetivos es importante, pero también hay que hacerlo en relación a las necesidades del área. Por esta razón el criterio de utilidad responde a la pregunta de en qué medida el impacto-s del programa o proyecto ha contribuido a las necesidades del área, sector o grupo de población. En desarrollo local y rural, los programas o proyectos son útiles si introducen cambios que satisfacen las necesidades que tenía el área, sector económico, empresa, grupo de población, etc. Uno de los problemas a tener en cuenta con el criterio de utilidad es que, cuando se trabaja en programas con múltiples proyectos o dirigidos a un área con diferentes grupos, no siempre hay una única necesidad, claramente definida y aceptada por todos; lo habitual es que tengamos diferentes intereses, en ocasiones incluso enfrentados: lo que es útil para un sector de población puede ser poco o nada útil para otro sector.

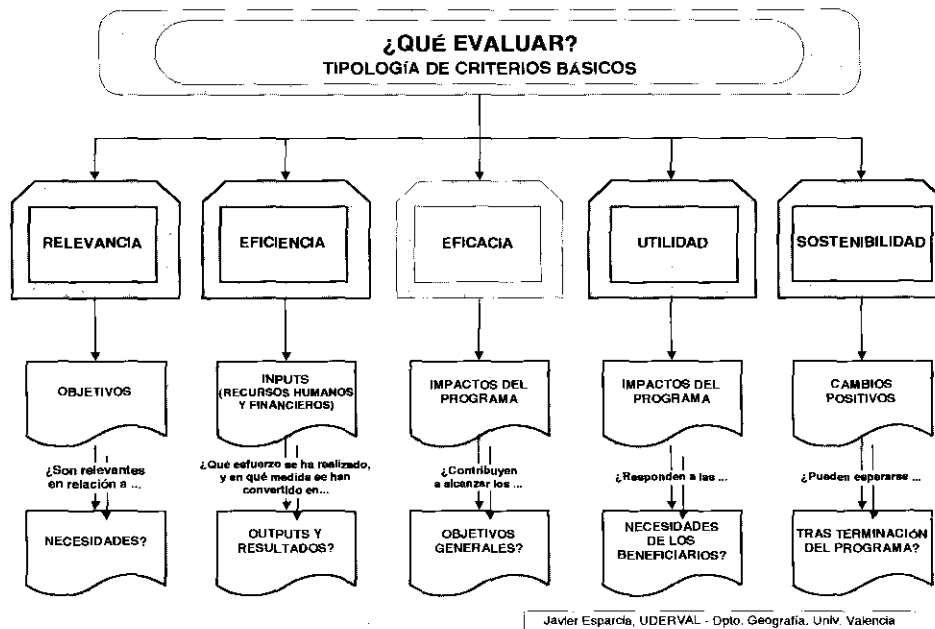


Figura 3. ¿Qué evaluar?: tipología de criterios básicos.

- **Sostenibilidad:** aquí se trata de valorar la continuidad de los efectos después de la finalización del programa, es decir, en qué medida podemos esperar cambios positivos después de que el programa o proyecto haya finalizado. Este criterio está muy próximo al de utilidad. Incluso en las situaciones en las que un programa contribuya a las necesidades del área, tendrá un valor reducido si esos beneficios o efectos desaparecen con la finalización del programa. Es de esperar que un programa verdaderamente útil sea aquel que es capaz de generar efectos directos e indirectos que permanecen después del programa, o que afectan positivamente a sectores o áreas no implicados de manera directa en el mismo.

En el Borrador de anteproyecto de Real Decreto por el que se regula la aplicación de LEADER Plus y PRODER II se insiste precisamente en que “la evaluación permitirá conocer el funcionamiento de los Grupos de Acción Local, su grado de eficacia y eficiencia en aplicación de su programa, procediendo, en caso necesario y a mediados del período de programación, a una revisión de los Grupos, a fin de reajustar su ejecución en orden al mejor cumplimiento del programa regional”.

2. *¿Cómo evaluar?: transformar conceptos en indicadores*

Una de las fases más importantes tanto en el caso del control y seguimiento como en la propia evaluación, es el diseño del sistema de indicadores. La naturaleza, la orientación y los mismos indicadores serán parcialmente diferentes en uno y otro caso; los indicadores relacionados con los efectos o resultados directos constituirán el núcleo del sistema de control y seguimiento. De cara a la evaluación, estos indicadores también estarán presentes, pero serán sólo una parte del sistema de indicadores diseñado. En este último

caso se incluirán también toda una serie de indicadores referidos a los efectos a medio y largo plazo, directos e indirectos.

En cualquiera de los casos, el sistema de indicadores tiene en líneas generales un objetivo: constituirse en unas medidas operativas con las que poder medir de manera clara y precisa cada uno de los criterios que estamos utilizando o que se consideran relevantes. Los indicadores no son sino un conjunto de referentes empíricos claros en relación a aquellos conceptos que en principio se han definido únicamente a un nivel teórico, es decir, el conjunto de objetivos estratégicos, operativos y específicos (European Commission, 1997).

El indicador es la unidad que permite medir esos conceptos que se han enunciado como objetivos del proyecto de desarrollo. Conseguir esos referentes empíricos o indicadores, manteniendo un elevado grado de fiabilidad, es uno de los principales problemas que se aborda en la evaluación de proyectos (o, en su caso, el seguimiento). El indicador hace referencia a características o atributos que pueden medirse y permiten valorar un proyecto o programa en términos de sus *outputs* y de su impacto (tanto a corto-medio plazo, es decir, objetivos específicos, como a largo plazo, es decir, objetivos generales o estratégicos). Los indicadores de *outputs* del proyecto o programa son relativamente fáciles de obtener, debido a que los responsables del mismo tienen información de qué es lo que se está produciendo o haciendo.

Sin embargo, los indicadores de impacto conllevan muchas más dificultades, debido por un lado a que es mucho más complejo definir qué efectos han sido provocados directamente por el proyecto o programa, y por otro a que puede resultar muy costoso en términos de tiempo o de recursos financieros medir directamente estos impactos. Por estas razones con frecuencia se acude a indicadores indirectos (por ejemplo, el número de vehículos de turismo por cada 100 habitantes en un área, en relación al conjunto regional, podría utilizarse como indicador indirecto del nivel de vida del área).

Pero el diseño del sistema de indicadores no es una fase sencilla, ni puede hacerse sólo pensando en la información disponible. Es necesario tomar como punto de partida los objetivos del proyecto, que pueden constituir en su formulación inicial únicamente un conjunto de conceptos, definidos de manera más o menos simple (European Commission, 1997, 1999). En ocasiones podemos tener conceptos simples que a su vez se trasladan a objetivos simples; por la propia formulación de tales objetivos éstos pueden implicar a la vez un indicador directo (ej. aumentar la tasa de inversión privada, aumentar el porcentaje de visitantes a un paraje...); en este caso estamos ante indicadores directos (tasa de inversión privada, porcentaje de visitantes...).

Sin embargo, es frecuente que el objetivo más genérico de un proyecto sea un concepto que esté referido a un conjunto complejo de fenómenos, y que éstos no puedan ser observados directamente. En estos casos es necesario separar las diferentes dimensiones de ese concepto, que puede ser un objetivo general: "elevar el nivel de desarrollo del área". Este es un concepto complejo y a la vez un objetivo muy genérico.

A efectos de la evaluación del proyecto sería necesario descomponer este concepto-objetivo del proyecto en objetivos específicos, y para cada uno de ellos elaborar una batería de indicadores. Así, "elevar el nivel de desarrollo del área" puede descomponerse en objetivos específicos como "mejorar y aumentar el nivel educativo de la población", "mejorar la cualificación profesional de los jóvenes y adultos", "aumentar las actividades económicas", "aumentar los ingresos medios", "aumentar la inserción de las mujeres en el mercado laboral", etc. A su vez, sería posible continuar descomponiendo cada uno de estos objetivos específicos.

Este trabajo de descomponer un objetivo general o estratégico de un proyecto o programa en objetivos operativos y específicos, forma parte de los contenidos necesarios para la definición del propio proyecto o programa. Sin embargo, de cara a la evaluación, también es posible empezar a diseñar una serie de indicadores en el momento en el que se elabora el proyecto o programa. Es conveniente que para cada objetivo específico, de cada programa –si los hay-, se elaboren o planteen varios indicadores, de manera que tengamos la certeza de que efectivamente se va a medir el concepto que se pretende, y en sus diferentes dimensiones.

El objetivo “aumentar las actividades económicas” puede medirse a través de indicadores como “número de empresas creadas durante el proyecto”, “número de empleos fijos generados, retenidos...”, “inversión privada generada”, etc. Pero es posible avanzar aún más, y transformar esos indicadores generales en indicadores específicos e índices, más o menos simples: “porcentaje de empresas creadas respecto a empresas existentes”, “empleo generado o retenido en relación al empleo existente”, “empleo generado en relación a la inversión global del proyecto”, “aumento de la facturación en relación a las ayudas públicas”, “inversión privada movilizada en relación al coste de funcionamiento del centro de desarrollo”, etc. Aquí se está evaluando un concepto definido de manera muy general (“aumento de las actividades económicas”), pero también todo el conjunto de acciones en las que ese objetivo se ha traducido durante el proyecto (programa de ayuda a la inversión, programa de creación de empleo, programa de creación de nuevas empresas, ...).

Por tanto, desde un punto de vista metodológico, de manera previa a la puesta en marcha del proyecto o programa se debería: a) definir muy claramente cómo se descompone cada objetivo general o estratégico en objetivos específicos; y b) definir una batería de indicadores para cada objetivo específico, en relación a su vez a los cinco grandes tipos de criterios, si es posible (relevancia, eficiencia, eficacia, utilidad y sostenibilidad). Ha de tenerse muy claro desde el principio qué información va a ser necesaria y cómo puede obtenerse, estableciéndose un sistema de recogida de esa información, incluso durante la ejecución del proyecto si fuese necesario.

3. Momento de realización, alcance, destinatario y protagonista de la evaluación

En cuanto al momento de realizar la evaluación, tradicionalmente las posibilidades se han reducido principalmente a tres, antes (evaluación ex – ante), durante (evaluación intermedia), o después (evaluación ex – post). No obstante, aunque en sentido estricto no puede hablarse de evaluación continua, la incorporación de mecanismos de seguimiento prácticamente continuos hace que se hable también de evaluación continua. El seguimiento por tanto es el que aporta y genera la información necesaria para una correcta evaluación. Cuando se dispone de un seguimiento preciso y una fuente de información suficientemente actualizada, la evaluación puede realizarse en cualquier momento; de ahí que se hable de evaluación continua. Por otra parte, la lectura entre líneas de la Iniciativa LEADER Plus pone de relieve que se está pensando prácticamente en una evaluación de este tipo al plantear una de carácter previo, otra intermedia y otra final (European Commission, 1997). En esta misma línea, el anteproyecto de Real Decreto por el que se regula la aplicación tanto de la Iniciativa LEADER Plus como de los programas de diversificación económica y desarrollo rural (PRODER II), se contempla de manera explícita que “las Comunidades Autónomas adoptarán mecanismos internos y/o externos que les permita una evaluación continua de los distintos proyectos y medidas de los programas territoriales y del conjunto del programa regional”.

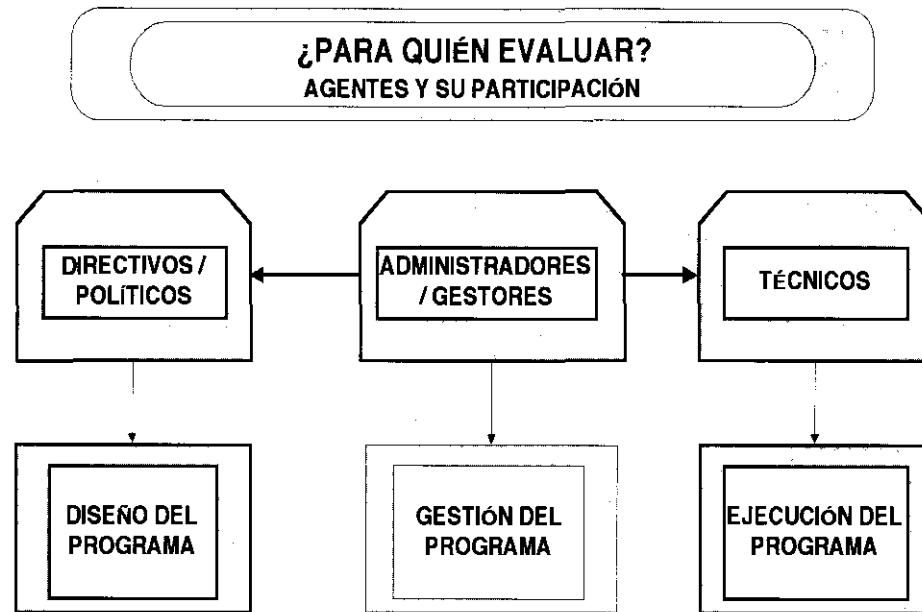


Figura 4. ¿Para quién evaluar?. Agentes y su participación.

En relación al alcance las posibilidades son también reducidas. Definir el alcance significa decidir qué partes del programa, proyecto o acciones queremos evaluar. Habitualmente se piensa en una evaluación referida al conjunto del programa, en cuyo caso estaríamos hablando de una evaluación global; sin embargo, cuando la evaluación se realiza a pequeña escala, o tiene un carácter interno (autoevaluación), es probable que lo que más interese es centrar la evaluación en un tema o aspecto concreto. En este caso estamos ante una evaluación temática o sectorial. Hay una tercera alternativa, que en realidad se deriva de esta última, que consiste en analizar en profundidad varios aspectos o relaciones (CEDER – beneficiarios, GAL – población local, etc.).

Básicamente aquí hay tres grandes tipos de destinatarios, ligados a tres diferentes funciones en relación al programa: directivos y políticos por un lado, administradores y gestores por otro, y técnicos en últimos lugar. Las respuestas que cada uno de estos niveles buscan en una evaluación pueden ser parcialmente diferentes, porque también lo son las funciones: diseño de los programas o políticas en el primer caso, gestión del programa a nivel político-social en el segundo, y aplicación o implementación en el tercero. Aunque este esquema es una simplificación de los muchos niveles y funciones, es posible buscar los grandes paralelismos en el caso de LEADER o PRODER: la Comisión Europea concibe la Iniciativa (o el Ministerio de Agricultura en el caso de PRODER), y por tanto tienen la función de diseñar el programa o la política de desarrollo rural; el GAL es el responsable de la gestión del programa (concepto que va mucho más allá de la pura gestión administrativa), especialmente de cara a la población, a las instituciones locales o regionales, o a los agentes económicos y sociales. Por último, el equipo técnico del CEDER es el responsable directo de la implementación del programa, y éstos buscarán en la evaluación unas respuestas concretas a aspectos técnicos.

Es cierto no obstante que esta separación no existe como tal en la realidad, en gran parte porque los niveles de gestión y de implementación trabajan de manera conjunta, no existe una división tan clara de funciones, y las preocupaciones de unos son también la de los otros: los técnicos están igualmente interesados en qué impacto tiene el programa sobre la sociedad local, y no sólo sobre cuanto dinero se invierte y si el proyecto se pone en marcha o no; a la inversa ocurre algo parecido: los miembros del GAL tienen –o pueden tener– interés en conocer la eficiencia del programa (porque también es su función y responsabilidad) y aspectos concretos de la implementación del mismo.

Por último, en cuanto a quién va a realizar la evaluación, básicamente ésta puede ser externa o interna. La diferencia es que en el primer caso el equipo que realiza la evaluación forma parte o no de la organización responsable de la implementación del programa o acciones. La evaluación externa se identifica habitualmente con procedimientos reglamentarios, es decir, con evaluaciones que se realizan por decisión de un órgano superior (la Comisión Europea es la que decide que éstas han de realizarse, y cuando un GAL al que se le ha aprobado su propuesta se acoge a la Iniciativa LEADER acepta que tal evaluación se realice). Sin embargo, las evaluaciones internas suelen ser decisiones de alguno de los órganos responsables de la gestión o implementación del programa. Junto a la evaluación interna y externa podríamos incluir un tercer tipo, la evaluación participativa (PATTON, 1987, 1990; COUSINS y EARL, 1992). Aunque las conexiones con una evaluación interna son evidentes, no es tal, al menos en sentido estricto: una evaluación participativa puede ser conducida por un equipo externo, no implicado en la implementación del programa, aunque también por un equipo interno. Es aquí donde surge la autoevaluación, como una forma de evaluación interna, realizada por y para los responsables de la gestión o implementación del programa, proyecto o acciones.

LEADER: DE LA NULA CULTURA EVALUATIVA A LA AUTOEVALUACIÓN RESPONSABLE

La lógica evaluativa en LEADER

La Iniciativa Comunitaria LEADER, al igual que el resto de iniciativas y de muchas de las intervenciones comunitarias, está sujeta a mecanismos de evaluación reglamentarios desde sus inicios. La importancia e interés de la evaluación de la Iniciativa ha llevado a la Comisión a pasar de una evaluación final en LEADER I (ESPARCIA y NOGUERA, 1997; SARACENO, 1998; ISLA y SOY, 1998), a una intermedia y otra final en LEADER II, y a plantear a una ex – ante, otra intermedia, y otra final, en LEADER Plus.

LEADER tiene, probablemente de forma más clara que otras iniciativas u otros programas comunitarios, una lógica muy clara en lo que son los mecanismos que conducen a una evaluación en tanto que instrumento útil de toma de decisiones; a todo ello ayuda sin duda la delimitación territorial clara que acompaña a un área LEADER.

- En primer lugar, los problemas socioeconómicos están claramente identificados, y han sido tratados desde hace ya muchos años, desde la despoblación y el envejecimiento hasta la pérdida de tejido empresarial, o los escasos esfuerzos de diversificación productiva (ESPARCIA, 2000a).
- En segundo lugar, las necesidades están igualmente identificadas: consolidar o aumentar el empleo y con éste unas fuentes de ingresos, que son absolutamente

imprescindibles para su mantenimiento en las áreas rurales, y por lo tanto para cumplir todas las funciones productivas y de preservación en relación al entorno rural que se les encomienda en el marco de estrategias territoriales de escala europea. Es una población inmersa en unas estructuras muy frágiles, que ha perdido su tradicional vía de ingresos, y que no puede, no encuentra vías o no sabe cómo valorizar algunos recursos, tradicionalmente poco explotados pero que con los cambios sociales y económicos en las áreas urbanas son susceptibles de ser valorizados. Por tanto, fortalecer esas estructuras productivas, sea a través de la especialización o de cualquier otra estrategia, es una necesidad imperiosa para la población rural.

- En tercer lugar, el programa está claro: la Comisión Europea, con mayor o menor participación de los estados, asume esos problemas y esas necesidades y pone en marcha mecanismos que contribuyan a paliar los problemas y a satisfacer las necesidades de la población rural. Resultado: la Iniciativa LEADER.

- En cuarto lugar, los objetivos de la Iniciativa también son evidentes, especialmente los de carácter general, pero también apunta vías claras para objetivos específicos e incluso operativos. El significado de estos objetivos para cada área LEADER puede sufrir algunas variaciones, sobre todo cuando se trata de objetivos específicos u operativos. Parece claro que los objetivos generales son comunes a todas las áreas LEADER.

- En quinto lugar, la Iniciativa se aplica en cada área seleccionada, y se obtienen unos resultados. Las personas que están más ligadas a la gestión y/o implementación del programa tienen una idea aproximada de cuáles son los resultados que se están obteniendo y, en algunos casos, también del impacto del programa.

- No obstante lo anterior, con frecuencia esta idea aproximada no deja de tener un carácter intuitivo, a tenor de las percepciones que va obteniendo o de la información sobre la marcha de los expedientes. Al final cada uno se forma una idea de si las cosas van más o menos bien, pero raras veces se tiene un conocimiento sistemático y preciso de aspectos que vayan más allá de la eficiencia económica. Se dispone de una cierta "valoración", pero nada preciso; por tanto, no estamos ante una evaluación (recorde-mos las diferentes definiciones). Llega el momento de pensar en una evaluación.

Los grupos de interés en la evaluación de LEADER

La evaluación de los programas LEADER tiene un interés evidente para muchos.

En primer lugar para los técnicos, que pueden y deben querer conocer si los mecanismos de implementación han sido los adecuados o si se han producido desviaciones, en qué sentido y qué efectos han tenido o están teniendo (ESPARCIA, 1999).

En segundo lugar, el propio GAL, que es el responsable máximo no sólo de la implementación que hagan los técnicos, sino de la forma específica que el programa ha tomado en el área, desde su concepción hasta los mecanismos previstos de participación de la población local, o de dinamización del tejido empresarial. En definitiva, el GAL es, como hemos dicho en otros trabajos, el responsable de si hace de sí mismo un instrumento de liderazgo del proceso de consolidación o de cambio económico y social en el área (que es el objetivo general que plantea la propia Comisión en las comunicaciones de la Iniciativa). Y el primer y máximo interesado en saber si todo esto se está alcanzado es el GAL.

En tercer lugar, una evaluación de los programas LEADER a nivel local interesa también a los gobiernos nacionales y, sobre todo, la Comisión, que son los responsables prin-

cipales de la concepción del programa y de su diseño. Su interés radica en confirmar si la aproximación LEADER es o no adecuada para abordar los problemas socioeconómicos de las áreas rurales desfavorecidas, por un lado, y para dar respuesta a las necesidades de la población que vive y trabaja en estas zonas, por otro (ESPARCIA y NOGUERA, 1997; SARACENO, 1998). En función de cuáles sean los resultados de las evaluaciones, en combinación con otro tipo de factores condicionantes, estos organismos pueden dar continuidad al programa (como ocurrió entre LEADER I y II), rediseñarlo (como ha ocurrido entre LEADER II y LEADER Plus), o proceder a su cancelación (como en algún momento se llegó a especular que podría ocurrir con LEADER, y que sí ha ocurrido con otras iniciativas o programas comunitarios).

Cada uno de estos tres grupos puede tener interés en la evaluación de aspectos diferentes, pero no sería lógico, en principio, realizar tres evaluaciones diferentes. Por eso en las evaluaciones globales se trata de incorporar los aspectos que puedan ser más relevantes para cada uno de ellos. En cualquier caso, y a nivel general, es evidente que los implicados en estos tres niveles pueden y deben tener interés en conocer la evaluación global.

La cultura evaluativa en torno a LEADER: muchas sombras... ¡y algunas luces!

Con frecuencia las evaluaciones quedan restringidas a estos tres grupos, o alguno de ellos. Hay también un cuarto grupo, olvidado casi siempre de una mera comunicación de resultados así como, por supuesto, de una discusión abierta y participativa sobre los mismos (ESPARCIA, 2000b,c). También con frecuencia las evaluaciones se utilizan como un instrumento de poder de facto, de control, en tanto que alguien, en alguna institución, conoce algo de otros que aquellos mismos no conocen. Esta visión, un poco maquiavélica, de la evaluación, ha estado y aún está presente en algunas instancias de nuestro país: ¿por qué no se han hecho públicas las evaluaciones correspondientes a los Grupos de LEADER I?; ¿tiene o ha tenido alguien reservas ante los resultados de las mismas?; ¿a quienes perjudican, si es que esto ocurre?; ¿han aportado información relevante?; ¿han servido para algo?; ¿conocen todos aquellos grupos, o los de LEADER II, los resultados?... Son muchos los interrogantes que aún podrían añadirse aquí.

En la época en la que nos encontramos, tras 25 años de democracia y casi los mismos de supuesta modernización del estado, no parece lógico que esto siga ocurriendo, por una razón muy sencilla: porque la falta de transparencia pone de manifiesto que no se concibe la evaluación como un instrumento de mejora en la toma de decisiones o, si esto es así, es para las decisiones de unos pocos (ESPARCIA, 2000b). No se entiende por tanto que los gerentes digan, en relación a la evaluación de LEADER I, "aunque no me lo han comunicado de manera oficial, sabemos que hemos quedado en la posición x" (siendo "x", por supuesto, una de las primeras posiciones, y curiosamente incluso con varios grupos supuestamente coincidiendo en esa misma posición "x"). Evidentemente la falta de transparencia conduce a despropósitos de este tipo, que no tendrían mayor importancia si no fuese por el flaco favor que se le ha hecho a una concepción mucho más abierta y operativa de la evaluación como instrumento de aprendizaje (a partir de los errores y aciertos de cada uno), de mejora en la toma de decisiones y, sobre todo, de corrección de las posibles desviaciones (ESPARCIA, 2000b). Ciertamente no obstante que los momentos en los que se han realizado las evaluaciones no han sido los mejores para contribuir de manera eficaz a todo ello. Pero incluso en

los casos en los que los resultados contribuyen menos, deben ser más accesibles. Este contexto pertenece a unas fases muy primigenias en la cultura evaluativa, aunque mejor habría que decir que las actitudes ocultistas y de falta de transparencia por lo que se caracterizan es por una total falta de cultura, cuando menos, evaluativa. Se trata de actitudes que no tienen –o deberían tener– cabida en el marco de una estructura moderna del estado.

Ahora bien, es justo señalar que estas actitudes están empezando a dejar paso a una situación más abierta, de mayor permeabilidad y de un mayor convencimiento, sincero, en torno a las virtudes de la evaluación. En esta línea, en España, la accesibilidad a las evaluaciones está empezando a dejar de ser una asignatura pendiente, y en esto han tenido un papel muy importante los gobiernos regionales, que en su mayor parte han facilitado el acceso a las mismas a los propios GAL. Pero no acaba aquí el “debe” en la cultura evaluativa. Desgraciadamente son aún numerosos los GAL que reproducen aquellas actitudes tan negativas, con una clara falta de transparencia hacia el cuarto grupo de interés, que a la postre es el que más interés tiene: la población local. ¿Cuántos grupos en España, de los que disponen de las evaluaciones intermedias, han publicado o difundido los principales resultados de las mismas?. Algunos lo han hecho, han trasladado el análisis de esos resultados a la población, y han intentado corregir algunas de las desviaciones o aspectos más negativos. Sin embargo, y aunque es cierto que la capacidad de respuesta de los Grupos ante los resultados de las evaluaciones intermedias ha sido reducida (por el retraso con el que se han realizado), también se constatan algunas –o muchas– actitudes de falta de convencimiento de que, en caso de haber llegado a tiempo, se hubiesen analizado con la suficiente seriedad. Desgraciadamente, la falta de cultura evaluativa está instalada no sólo en algunos directivos o responsables políticos; lo está en los niveles medios e inferiores, y evidentemente así no podemos avanzar. El primer paso es tener la valentía suficiente de aceptar públicamente los errores, y sólo la falta de ésta explica las actitudes más ocultistas e incluso despreciativas con los resultados de las evaluaciones.

A pesar de las dificultades señaladas en los párrafos anteriores, la cultura evaluativa se está abriendo camino en el complejo entramado institucional, orgánico y funcional ligado al desarrollo rural en España. Los GAL por un lado, y los equipos técnicos por otro, quieren saber cada vez más los resultados de las evaluaciones, y sobre todo quieren que les sea útil en su trabajo. Y surge aquí otro desajuste porque, justo es reconocerlo, las evaluaciones “reglamentarias” invocadas por las comunicaciones de la Iniciativa LEADER han tenido una utilización reducida por parte de los GAL y los equipos técnicos (ESPARCIA *et al.*, 1999a). En unas ocasiones porque no se ha podido disponer de tales resultados, en otras porque las metodologías utilizadas para realizar la evaluación han impedido que los resultados fuesen relevantes, y en otras porque las evaluaciones se han realizado cuando ya no había capacidad de respuesta. De todo ello se deduce que los responsables políticos y técnicos tendrán que afinar mucho más en el futuro si realmente quieren que el enorme esfuerzo que se invierte en una evaluación sirva para algo, y que éstas dejen de verse como algo impuesto y, en definitiva, como un puro trámite que necesariamente hay que pasar (ESPARCIA, NOGUERA y BUCIEGA, 1999a).

La autoevaluación: ¿por qué? ¿para qué? ¿cómo?

Dada la situación de los grupos LEADER parece deducirse hasta aquí que la evaluación no siempre permite cubrir todas las expectativas, y esta es precisamente una de las

justificaciones del interés que puede tener la autoevaluación (evaluación interna), o la evaluación participativa. En la figura 5 se sistematizan algunos de los aspectos de por qué estas últimas pueden ser más adecuadas, a partir de una serie de posibles objetivos inherentes a la evaluación. La importancia que otorgue a cada objetivo el responsable de tomar la decisión sobre el tipo de evaluación, será la que determinará si ésta va a ser externa, interna o participativa. Como se recoge en la tabla, las evaluaciones internas y/o participativas son las que permiten satisfacer en mayor medida este abanico de objetivos, aunque dependerá siempre de las prioridades en cuanto a función que ha de tener la evaluación.

Figura 5: Contribución de cada tipo de evaluación a cada criterio-objetivo.
Situación de partida

CRITERIO – OBJETIVO	TIPO DE EVALUACIÓN		
	EXTERNA	INTERNA	PARTICIPATIVA
Credibilidad exterior	Alta	Baja	Media
Legitimidad en el área	Alta	Media	Alta
Aprendizaje del G.A.L.	Baja	Media	Alta
Toma de decisiones G.A.L.	Baja	Alta	Alta
Necesidad de resultados	Baja	Alta	Media
Objetividad	Alta (1)	Baja	Media
Rigor metodológico	Alta	Baja	Baja
Conocimiento del programa	Baja	Alta	Media

(1) Media o Baja si el equipo evaluador tiene una dependencia clara y directa de contratos con esa administración pública. Fuente: elaboración propia.

En cualquier caso, lo cierto es que hay un desajuste claro entre las necesidades más habituales de los grupos de acción local y lo que les ofrece una evaluación externa: el GAL y el equipo técnico esperan que la evaluación sea un instrumento de aprendizaje, de toma de decisiones, de mejora en la implementación de su programa, de mejora en la aplicación de los proyectos por parte de los beneficiarios, o simplemente para conseguir que el programa no se desvíe de los objetivos marcados. La evaluación externa o reglamentaria no resuelve todas estas necesidades. Precisamente en un contexto caracterizado por este desajuste, mayor o menor según los casos, entre lo que ofrecen las evaluaciones reglamentarias y las necesidades del GAL y el equipo técnico, es donde la autoevaluación surge como exponente máximo de la cultura evaluativa, entendida no como formación en metodologías de evaluación sino como el convencimiento de que la evaluación es un instrumento útil y necesario en la puesta en marcha o aplicación de cualquier programa. Como ejercicio voluntario de evaluación, tiene la gran virtud y el gran potencial de poder diseñarse a medida. Ahora bien, una evaluación realizada por el grupo, por el equipo técnico, e incluso por los beneficiarios, no puede ni debe dejar de ser menos rigurosa en las metodologías y en las técnicas que una evaluación externa, realizada por un equipo de profesionales. Es probable que el GAL o el equipo técnico no sea capaz de diseñar una autoevaluación que cumpla con todos los requisitos metodológicos y técnicos; no es su función ni su cometido. Pero puede y debe asesorarse por profesionales (OWEN, 1998).

En cualquiera de los casos, también aquí, antes de empezar es necesario que el equipo responsable de la autoevaluación responda claramente a las cuatro cuestiones planteadas anteriormente (porque la quinta está clara, es el GAL o el equipo técnico quién la realiza, directa o indirectamente): hay que saber qué aspectos son relevantes y quieren evaluarse (¿global? ¿temática?, ¿parte de un sector? ...); para qué va a utilizarse la información (¿poder corregir las desviaciones durante la implementación del programa? ¿para preparar un programa nuevo? ¿por un simple examen de aciertos y errores del programa finalizado? ¿para estimular la autoestima?, ...). Por último, y en función del objetivo u objetivos de la autoevaluación, qué es lo que vamos o debemos evaluar (¿la eficacia del programa?, ¿la eficiencia en la utilización de los recursos? ¿la relevancia de los objetivos respecto de las necesidades? ¿la utilidad del programa y sus resultados en relación a las necesidades? ¿la sostenibilidad de los procesos que se han puesto en marcha? ¿otros aspectos?...). Hasta aquí la labor ha sido "relativamente" sencilla, porque cuando un equipo técnico o un GAL decide poner en marcha una autoevaluación suele saber lo que quiere.

A continuación viene una fase con un componente metodológico y técnico más importante, de diseño de una autoevaluación correcta y, sobre todo, que pueda ser útil y responder a las necesidades del GAL o del equipo técnico. Esta es la fase de diseño de una batería de indicadores claros, precisos y sistemáticos, que puedan medir, cualitativa o cuantitativamente, los diferentes aspectos (FITZ-GIBBON y MORRIS, 1987). Aunque esta fase debería requerir la participación de profesionales (OWEN, 1998), es cierto que con un bagaje rudimentario sobre metodologías de evaluación se puede empezar a trabajar; por otro lado, muchos avances en autoevaluación responden más a la imaginación y compromiso con obtener unos resultados útiles, que a una aplicación sistemática de las metodologías más adecuadas en cada caso (CARPENTER, 1998; European Commission, 1999). En cualquier caso es importante tener en cuenta que el alcance de una autoevaluación realizada por profesionales ha de tener, en principio, un mayor alcance que cuando no se cuenta con el concurso de éstos; en este último caso no puede ni debe exigírsele a una autoevaluación unos resultados o respuestas muy profundas y perfiladas a las cuestiones planteadas. El reto de la autoevaluación, y su virtud, es poder dar respuestas concretas a problemas concretos, siempre y cuando se efectúe con las suficientes garantías de calidad metodológica y técnica (NORDLANDER y BLOM, 1998).

AUTOEVALUACIÓN Y EVALUACIÓN CONTINUA: HACIA UN MÉTODO OPERATIVO PARA LOS GRUPOS DE ACCIÓN LOCAL

La autoevaluación es por tanto un instrumento de trabajo especialmente adecuado a situaciones en las que se necesitan respuestas concretas. En muchos casos no parecería lógico establecer un amplio programa de autoevaluación global, entre otras razones porque el esfuerzo que habría de invertirse supera con diferencia la capacidad de los grupos. Sin embargo, sí parece viable y hasta incluso necesario que los GAL, por su propia naturaleza, tuviesen algún mecanismo de autoevaluación periódico, al igual que los equipos técnicos. Esta autoevaluación puede tener desde un carácter puramente temático (referida a algunas de las principales medidas o especificidades de LEADER), hasta una valoración global del programa con una amplia participación de la población.

Este último aspecto es de sumo interés por cuanto lo que percibe la población, y especialmente los beneficiarios, es el impacto directo e indirecto que el programa está teniendo.

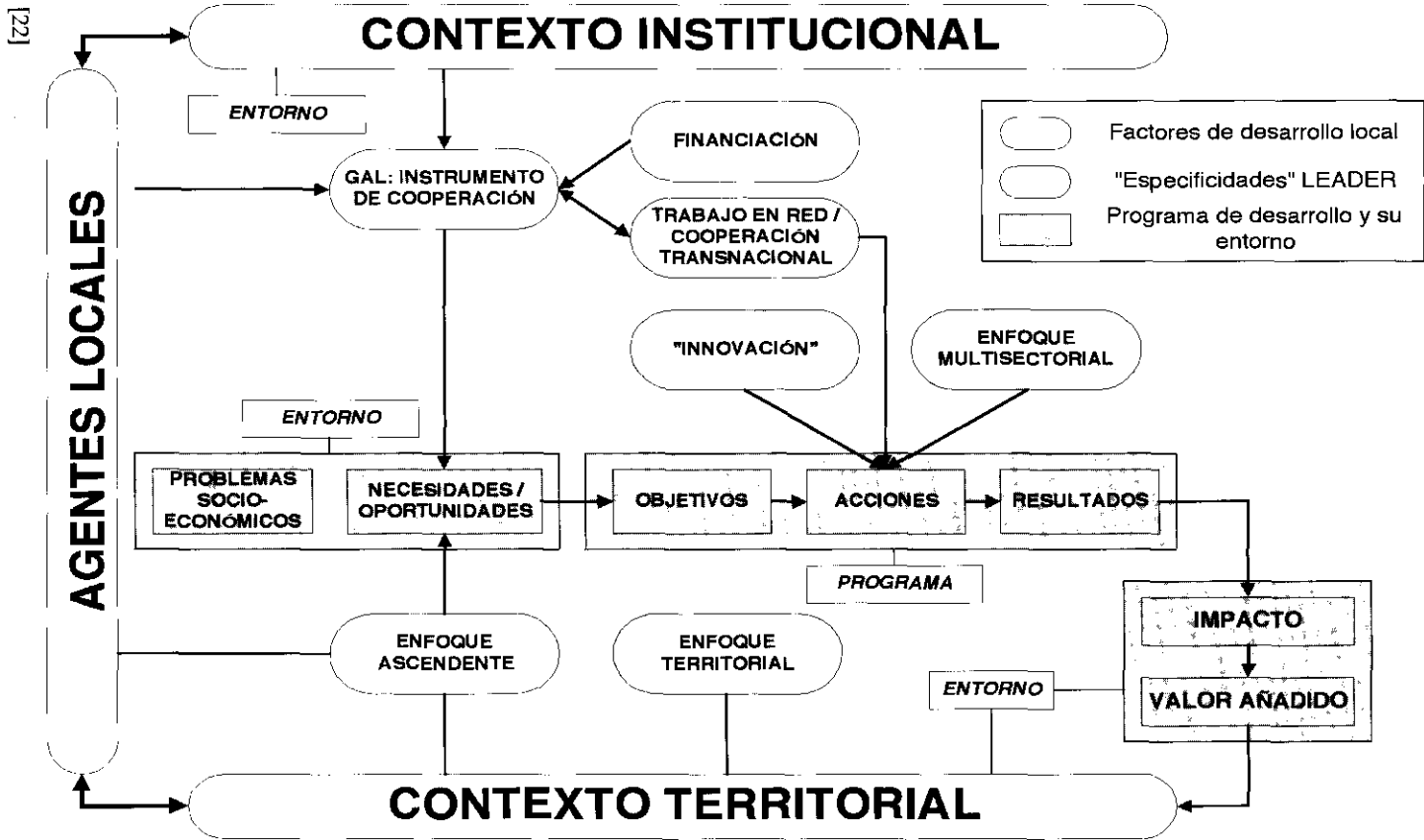
do sobre sus vidas y, en definitiva, sobre la estructura social y económica de las áreas rurales. La participación de la población en ejercicios de autoevaluación suficientemente rigurosos tiene un gran potencial como vía de valoración de la eficacia del programa, sin duda es uno de los aspectos de más difícil análisis, y esta es probablemente la mejor estrategia para conseguirlo. Adicionalmente, con un ejercicio de autoevaluación por parte de la población se contribuye a uno de los objetivos de la propia iniciativa, la participación e implicación de la población local (ESPARCIA *et al.*, 1999b; ESPARCIA *et al.*, 2000). Los GAL y los equipos técnicos tienen por delante un gran reto en el amplísimo campo de posibilidades de autoevaluación, todas ellas útiles para su trabajo. Ahora es necesario actuar, primero, sobre la extensión de la cultura evaluativa; segundo, sobre los medios materiales y humanos que hagan posible emprender la tarea; por último, sobre unos diseños rigurosos desde un punto de vista metodológico y técnico, que aseguren la utilidad y la aplicabilidad de los resultados.

¿En qué medida la autoevaluación puede ser a la vez una evaluación continua?. Como se ha señalado anteriormente la evaluación continua no existe en la práctica. Lo que sí es posible es un sistema de control y seguimiento prácticamente continuo, con el objetivo de corregir de manera inmediata cualquier desviación que pudiera producirse en los objetivos previstos. El seguimiento es un procedimiento de gran importancia para mejorar el funcionamiento del programa y, aunque no es una evaluación, con frecuencia una buena evaluación está basada en un buen sistema de seguimiento (debido a que éste permite generar información que puede utilizarse en la evaluación). Estas son las razones por las que en el propio anteproyecto de Real Decreto para regular la aplicación de LEADER Plus y PRODER II se insiste en avanzar hacia una evaluación continua.

El sistema de seguimiento sin embargo no suele orientarse a obtener todo tipo de información. Por su propia naturaleza, está más orientado a recoger y valorar los efectos directos del programa, es decir, los bienes, servicios o efectos que se derivan directamente de la ejecución del programa. El destinatario de un sistema de seguimiento tampoco tiene por qué coincidir con el destinatario de una evaluación. Al tratarse de información sobre el funcionamiento directo del programa, y no sobre su impacto a medio o largo plazo, son los técnicos y responsables directos de la ejecución del programa los primeros interesados en un buen sistema de control y seguimiento.

La clave de un buen sistema de control y seguimiento es el sistema de indicadores. Estos deben ser precisos y claros, tanto en relación a la obtención de la información como de cara a su interpretación. Por ello una fase fundamental en la puesta en marcha de un sistema de control y seguimiento (esté o no orientado a la evaluación), es definir cuidadosamente el sistema de indicadores, como se ha indicado en el punto 2.

Ni el sistema de control y seguimiento, ni la autoevaluación más o menos ocasional, son únicamente un análisis interno para la toma de decisiones por parte de los protagonistas locales, que deba quedar únicamente en eso. Hay que recordar que la autoevaluación puede ser tanto un punto de llegada como un punto de partida. De llegada porque nos puede permitir completar los resultados obtenidos por otras evaluaciones –reglamentarias o no- externas, pero que no ofrecen todas las respuestas que se necesitan; de partida porque sus resultados pueden integrarse, o al menos alimentar, a esas evaluaciones reglamentarias externas. Integrar unas y otras, tal vez contando con unas autoevaluaciones que además de ser rigurosas estuviesen convenientemente validadas por profesionales, es otra de las vías que han de considerarse por parte de los responsables ministeriales y de la Comisión. Puede también pensarse en unos mínimos para que los



Fuente: Adaptado a partir de AEIDL, 1999

Figura 6. Elementos clave para una autoevaluación en el Programa LEADER.

resultados de la autoevaluación sean útiles tanto en la toma de decisiones internas como a la hora de integrarlos en una posible evaluación externa reglamentaria. En la figura 6 se integran los tres principales factores de desarrollo (el territorio, los agentes y el contexto institucional), a partir de los cuales surgen una serie de elementos clave a analizar, las denominadas “especificidades LEADER”. El diseño del programa, con sus objetivos y estrategias, así como su ejecución, depende en gran parte de cómo se orienten estas siete especificidades (pese a corresponder a LEADER II, el esquema general puede ser igualmente válido en LEADER Plus). Por tanto, tal como señala el Observatorio Europeo LEADER, centrar la autoevaluación en estas especificidades puede dar las claves del funcionamiento del programa y del impacto sobre el territorio y sus agentes, que no en vano son los principales destinatarios de los esfuerzos contenidos en los programas de desarrollo rural.

Agradecimientos: Los temas de evaluación no suelen ser abordados por los geógrafos. Mi introducción en ellos se deriva de la visión práctica, aplicada y de utilidad social que el profesor José María Bernabé tuvo del quehacer del geógrafo. A él agradezco su apoyo e incitación directa, siempre con entusiasmo, a entrar en las metodologías y en el trabajo práctico de la evaluación de políticas públicas.

REFERENCIAS

- ALVIRA, F. (1991): *Metodología de la evaluación de programas*, Centro de Investigaciones Sociológicas, Cuadernos Metodológicos, n. 2. 94 pp.
- CARPENTER, D. (1998): Do-it-yourself Evaluation – the beginner’s guide to reconciling theory and practice, *International Conference on Evaluation: profession, business or politics?*, October 29th-31st, Rome, Italy. 13 pp.
- COHEN, E. y FRANCO, R. (1993): *Evaluación de proyectos sociales*, Siglo XXI, 318 pp.
- CONSEIL SCIENTIFIQUE DE L’ÉVALUATION (1996): *Petit guide de l’évaluation des politiques publiques*, Paris, CSE.
- COUSINS, J.B. y EARL, L.M. (1992): The Case for Participatory Evaluation, *Educational Evaluation and Policy Analysis*, n°14(4), pp. 397-418
- ESPARCIA, J. (1999): Elaboración y evaluación de proyectos de desarrollo local. *Curso sobre Herramientas para el Desarrollo Local*, Asociación para el Desarrollo Rural de la Isla de La Palma (ADER) – Universidad de La Laguna, Villa de Mazo, Abril-Mayo 1999, 14-15 de Mayo, 45 pp.
- ESPARCIA, J. (2000a): The LEADER Programme and the rise of rural development in Spain, *Sociologia Ruralis*, n° 40(2), pp. 200-207
- ESPARCIA, J. (2000b): Evaluación y autoevaluación: límites y posibilidades, *Autoevaluación y Evaluación en LEADER y PRODER: métodos y perspectivas. Jornadas Técnicas de la Unidad Española del Observatorio Europeo LEADER*, Cudillero (Asturias), 25-26 de septiembre, 24 pp. + 54 transparencias.
- ESPARCIA, J. (2000c): Evaluación de programas de desarrollo rural. *Curso de Especialista en Desarrollo Rural*, Septiembre 2000 – Febrero 2001, Universidad de Castilla-La Mancha, Facultad de Letras, Ciudad Real, 16-17 de octubre, 30 pp. + 95 transparencias.
- ESPARCIA, J., MOSELEY, M. y NOGUERA, J. (2000): *Exploring Rural Development Partnerships in Europe. An Analysis of 330 local partnerships across eight EU countries*, Informes UDER-VAL-Dpto. Geografía, Univ. de Valencia, 285 pp.
- ESPARCIA, J. (dir.), NOGUERA, J. y BUCIEGA, A. (1999a): *Evaluación Intermedia de LEADER II*

- en la Comunidad Valenciana, UDERVAL-Dpto. de Geografía, Universidad de Valencia, 376 pp. + anexos.
- ESPARCIA, J. y NOGUERA, J. (1997): El Programa LEADER I y el desarrollo rural en España. Análisis y Evaluación, en ONTIVEROS, A. MOLINERO, F. (eds.) *La investigación Hispano-Británica reciente en Geografía Rural: del campo tradicional a la transición postproductivista*, Murcia, Asociación de Geógrafos Españoles, pp. 269-284.
- ESPARCIA, J., NOGUERA, J. y BUCIEGA, A. (1999b): Emerging community development in Spain, en WESTHOLM, E., MOSELEY, M. y STENLAS, N. (eds.), *Local Partnerships and Rural Development in Europe: a Literature Review of Practice and Theory*, Falun: Dalarna Research Institute (Sweden), Contryside and Community Research Unit, Cheltenham and Gloucester College of Higher Education (UK), pp. 65-77.
- EUROPEAN COMMISSION (1993): *Project cycle management. Integrated approach and logical framework*. Directorate-General for Development.
- EUROPEAN COMMISSION (1997): *Evaluating EU Expenditure Programmes. A guide. Ex post and Intermediate evaluation*, Office for Official Publications of the European Communities, Luxembourg, 114 pp.
- EUROPEAN COMMISSION (1999): *Evaluating socio-economic programmes*. MEANS Collection, 5 vols.
- FITZ-GIBBON, C.T. y MORRIS, L.L. (1987): *How to Design a Program Evaluation*, Sage Publications, Newbury Park, California, 168 pp.
- GABARRÓN, L.R. y HERNÁNDEZ, L. (1994): *Investigación participativa*, Centro de Investigaciones Sociológicas, Cuadernos Metodológicos, n. 10. 82 pp.
- HERMAN, J.L., MORRIS, L.L. y FITZ-GIBBON, C.T. (1987): *Evaluator's Handbook*, Sage Publications, Newbury Park, California, 159 pp.
- ISLA, M.M. y SOY, A. (1998): Approaches on evaluation: the evaluation of LEADER I Initiative in Spain, Indicators to assess the LEADER Program and follow its implementation, *Seminar of the LEADER European Observatory*, Galashields, Scottish Borders, Scotland, 4-8 March, 12 pp.
- MENÉNDEZ, R. (1999): Seguimiento y evaluación de experiencias de desarrollo local, en GARCÍA, J.L., GODENAU, D. Y FEBLES, M.F. (Eds.) *Instrumentos para el Desarrollo Local*, Ayuntamiento de Santa Cruz de Tenerife, pp. 279-303.
- NORDLANDER, L. y BLOM, B. (1998): Participatory Evaluation as a Road to Knowledge Utilisation, *International Conference on Evaluation: profession, business or politics?*, October 29th-31th, Rome, Italy. 8 pp.
- ORGEGON, E. y MEDINA, J. (1997): *Prospectiva: construcción social del futuro*, ILPES-Universidad del Valle, Santiago de Cali, Colombia, 371 pp.
- OWEN, J.M. (1998): Interactive evaluation: Using Professional Evaluators to Encourage Effective Participatory Evaluation, *International Conference on Evaluation: profession, business or politics?*, October 29th-31th, Rome, Italy. 18 pp.
- PATTON, M.Q. (1987): *How to Use Qualitative Methods in Evaluation*, Sage Publications, Newbury Park, California, 176 pp.
- PATTON, M.Q. (1990): *Qualitative Evaluation and Research Methods*, Newbury Park, California, 532 pp.
- PAWSON, R. y TILLEY, N. (1997): *Realistic Evaluation*, Sage Publications, London, 235 pp.
- ROSSI, P.P.H. y FREEMAN, H.E. (1979-93): *Evaluation. A Systematic Approach*. Sage Publications, Newbury Park, California. 488 pp.
- SARACENO, E. (1998): Comparing local development processes across Europe: some met-

- hodological learning from the LEADER I ex post evaluation experience, *International Conference on Evaluation: profession, business or politics?*, October 29th-31th, Rome, Italy. 9 pp.
- SHULHA, L.M. y COUSINS, B. (1997): Evaluation Use: Theory, Research, and Practice Since 1986, *Evaluation Practice*, n° 18(3), pp. 195-208.
- SOY, A. (1995): Estrategias de desarrollo local: algunas propuestas basadas en la evaluación de los programas estructurales e iniciativas comunitarias, 35th *European Congress of the Regional Science Association*, Odense, 22-25 Agosto, 25 pp.
- STECHER, B.M. y DAVIS, W.A. (1987): *How to Focus an Evaluation*, Sage Publications, Newbury Park, California, 94 pp.
- VEDUNG, E. (1997): *Evaluación de políticas públicas y programas*, Ministerio de Trabajo y Asuntos Sociales, Instituto Nacional de Servicios Sociales, Madrid, 409 pp.
- VIVERET, PP. (1989): *L'évaluation des politiques et des actions publiques*, Paris, La Documentation Française.
- WEISS, C.H. (1988): Evaluation for decisions: Is anybody there? Does anybody care?, *Evaluation Practice*, n° 9(1), pp. 5-19.

