

Esparcia, J. (2001): "La evaluación de programas de desarrollo rural", en Pillet, F., y Plaza, J. (coords.): *Lecciones de Desarrollo Rural. Una aproximación formativa desde y para Castilla-La Mancha*. Universidad de Castilla-La Mancha, CEDERCAM-Asociación para el Desarrollo Rural de Castilla-La Mancha, pp. 159-166.

## **LA EVALUACIÓN DE PROGRAMAS DE DESARROLLO RURAL**

*Javier Esparcia Pérez  
Universidad de Valencia*

Los programas de desarrollo rural, como muchas otras intervenciones de carácter estructural o financiadas parcialmente con los Fondos Estructurales, están sujetos a una evaluación reglamentaria. Es así como los Programas LEADER, y en breve también ocurrirá con PRODER, han debido pasar estas evaluaciones, realizadas habitualmente por equipos externos. La evaluación ha empezado a estar muy presente en el quehacer diario de los Grupos LEADER, y en menor medida PRODER, y parece que la tendencia es que sea algo cada vez más habitual: LEADER I tuvo una evaluación al finalizar el Programa; LEADER II va a tener o está teniendo esta misma evaluación final (el límite para realizarla es 2003), pero a ella se añade la evaluación intermedia que hubo de realizarse entre 1998 y 1999. Por último, para LEADER + la Comisión ha añadido a estas dos evaluaciones una tercera, a realizar antes de la aprobación de los programas regionales de desarrollo rural (en los que se diseña la estrategia de los gobiernos regionales para LEADER +), y referida precisamente a estos programas. En definitiva, en pocos años se ha pasado de una evaluación por programa a tres, lo cual pone de relieve la preocupación que la Comisión Europea tiene por evaluar este tipo de intervenciones. Es de suponer que la tendencia del Programa PRODER II va a ser también la de aumentar el número de evaluaciones.

Estas tendencias podrían hacernos pensar que la evaluación de programas de desarrollo rural es tan frecuente como aceptada y asumida por quienes son evaluados. Sin embargo esto no es así. Los Grupos siguen siendo, por lo general, poco receptivos a las evaluaciones reglamentarias, por una serie de razones bien justificadas algunas, y otras en la base de una cultura evaluativa muy poco generalizada en nuestro país. Efectivamente, España es uno de los países de la Unión Europea donde ha habido y hay menor tradición de evaluación de programas, de todo tipo. Esta escasa tradición responde por un lado a unas prácticas administrativas que se han preocupado poco por saber de forma sistemática qué efectos tenían las políticas, y por otro a actitudes personales, muy enraizadas en las administraciones públicas, contrarias o poco favorables a ser evaluados.

En estas actitudes hay un error básico de partida. A la evaluación, especialmente si es externa, se le asocian consideraciones como la de “eso es una intromisión en nuestro trabajo”, o “nadie puede saber tanto de cómo está funcionando un programa como nosotros, que lo diseñamos o ejecutamos”. Ciertamente esta consideración de la evaluación se aleja con mucho de lo que en realidad es: por un lado un instrumento de aprendizaje y búsqueda de la excelencia en nuestro trabajo; por otro, un instrumento de mejora de la toma de decisiones, que nos permite minimizar riesgos y maximizar las oportunidades. Una evaluación, que no ha de confundirse con un estudio científico o académico, ha de cumplir otros requisitos para ser rigurosa, como el hecho de ser independiente, objetiva, sistemática, y con una orientación aplicada que la haga especialmente útil para la toma de decisiones; fuera de este contexto, las evaluaciones pierden su sentido, y podemos tener estudios científicos de mucha calidad, pero que no son evaluaciones, porque no están orientados a la toma de decisiones, o no son útiles para los beneficiarios o los responsables de las políticas.

Y esto último es lo que ha ocurrido en España en los últimos años. Tal y como se han realizado las evaluaciones, especialmente en los programas de desarrollo rural, éstas no han contribuido de manera clara a estos objetivos, el aprendizaje y la mejora de la toma de decisiones; pero no por ello debemos negar la utilidad del instrumento, que nos permite no sólo analizar cómo hemos hecho nuestro trabajo, sino sobre todo cómo podemos hacerlo mejor en el futuro.

En el caso de LEADER, las evaluaciones realizadas han cumplido poco estos objetivos. LEADER I sólo tuvo una evaluación final, encargada por el organismo intermediario, el MAPA, que no ha sido difundida, ni siquiera a los Grupos. Cuando se realizó esta evaluación, LEADER II estaba ya diseñado, y su contribución fue escasa o nula, al menos a ese diseño directo, aunque sí es cierto que el MAPA ha tenido en consideración los resultados obtenidos. La evaluación intermedia de LEADER II se ha realizado prácticamente cuando apenas quedaba un año, en el mejor de los casos, para el cierre de la ventanilla (diciembre de 1999, fecha a partir de la cual únicamente se continúan gestionando trámites de los expedientes ya aprobados); por tanto, tampoco aquí los Grupos han tenido capacidad de maniobra o respuesta ante los resultados de esas evaluaciones intermedias. Por otra parte, en algunos casos a este retraso se añade el hecho de la escasa o nula difusión, habiendo alguna Comunidad Autónoma que no ha difundido los resultados, ni pretende hacerlo, ni siquiera a los Grupos LEADER (“para evitar suspicacias y roces entre los Grupos”, se añade por parte de los responsables, es decir, una actitud que es la antítesis de lo que deber ser y para lo que debe servir la evaluación, porque la falta de transparencia es la mejor manera de conseguir lo que se quiere evitar). Por último, la evaluación inicial o ex - ante de LEADER +, que se

está realizando en estos momentos (o prácticamente ya habrá acabado en la mayor parte de Comunidades Autónomas), no ha podido contar con los resultados de la evaluación final de LEADER II, que aún no ha iniciado.

En definitiva, no parece extraño por tanto que los Grupos, que son los que han de ser evaluados, y en teoría, los principales beneficiarios de la evaluación, apenas vean más que incomodidades cuando se anuncia una nueva evaluación reglamentaria-externa. Sin embargo, estos inconvenientes ligados a la organización y práctica de las evaluaciones reglamentarias no debe hacer perder de vista las ventajas y el potencial que tiene una evaluación. Por esta razón, algunos grupos, conscientes de este potencial y de que la evaluación es ante todo un instrumento de aprendizaje y de mejora en su proceso de toma de decisiones, están poniendo en práctica, o piensan hacerlo en breve, de mecanismos de evaluación más flexibles y adaptados a sus necesidades, principalmente en el contexto de la evaluación interna o autoevaluación. Una muestra de este interés son las recientes Jornadas Técnicas de la Unidad Española del Observatorio Europeo LEADER, dedicadas a analizar la "Autoevaluación y Evaluación en LEADER y PRODER: métodos y perspectivas"<sup>1</sup>. Aquí se ha puesto de manifiesto que, ante las deficiencias de las evaluaciones reglamentarias, la autoevaluación es el enfoque más adaptado a las necesidades de los grupos, y a su vez que es posible que los resultados de una autoevaluación bien realizada, puedan constituir el input de las evaluaciones externas reglamentarias en el futuro.

La lógica evaluativa de los programas de desarrollo rural se presta de manera muy clara a establecer unos mecanismos de evaluación útiles para los grupos y para los responsables de las políticas de desarrollo rural. Aunque esa lógica evaluativa está presente en el ámbito de las políticas públicas en general, en el caso de los programas de desarrollo rural esto es mucho más claro: hay unos problemas socioeconómicos, bastante bien delimitados en todos los casos, de los que surgen unas necesidades que, en principio, son bien definibles, aunque no siempre lo están. Estas necesidades se traducen en los programas de desarrollo rural, con unos objetivos que han de estar claros, unas estrategias coherentes y unas orientaciones para las diferentes acciones o tipos de acciones. Según esta estructura parece claro que es necesario realizar una evaluación a la finalización del periodo de implementación de estas acciones, para valorar en qué medida se han conseguido los objetivos previstos. De la misma manera, durante la implementación de la estrategia y de las acciones, parece lógico realizar alguna evaluación que permita a los Grupos tener una cierta capacidad de rectificación. Por último, a la hora de diseñar el programa resultaría de gran utilidad disponer de una evaluación que nos indique los problemas y las necesidades reales.

<sup>1</sup> Página web de la Unidad Española del Observatorio (<http://redrural.tragsateg.es>).

así como los instrumentos más adecuados para abordarlas. Tenemos así las tres evaluaciones que en estos momentos son reglamentarias en LEADER →: la final, la intermedia y la inicial. Esto nos lleva a hablar prácticamente de evaluación casi continua (aunque en sentido estricto esta no existe, sino únicamente mecanismos de control y seguimiento que permiten generar una parte de la información necesaria en la evaluación). No se trata de trasladar a los Grupos un estado de psicosis evaluativa, pero sí es posible, si se asume la utilidad de la evaluación, tener mecanismos de seguimiento y/o evaluación interna, de bajo coste, que informen con periodicidad sobre la marcha del programa. Evidentemente lo lógico es también que para ello se cuente con el asesoramiento de expertos.

A lo largo de los últimos párrafos ha quedado claro que la evaluación externa está siendo poco útil para los Grupos, y que la evaluación interna o autoevaluación se configura como una alternativa viable. A estos dos tipos de evaluación podría añadirse la evaluación participativa, que puede estar a mitad de camino entre ambas, o asociarse a la evaluación interna. En cualquier caso, a la hora de seleccionar uno u otro tipo es necesario tener en cuenta una serie de criterios-objetivo, para ser plenamente conscientes de las ventajas e inconvenientes de una y otra: credibilidad exterior, legitimidad en el área, aprendizaje del grupo, toma de decisiones por parte del grupo, necesidad urgente de obtener resultados, objetividad, rigor metodológico, o conocimiento del programa, son algunos de estos criterios. En la tabla siguiente tenemos una simplificación de cuál podría ser la situación de partida en cuanto a la contribución de cada tipo de evaluación para cada uno de estos criterios u objetivos. La selección de un tipo u otro de evaluación dependerá en gran parte de qué criterios sean más importantes en cada momento, al margen de que, por supuesto, es posible intentar superar las debilidades que pueda tener el tipo de evaluación que más se adapte a nuestras necesidades.

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

	<b>EXTERNA</b>	<b>INTERNA</b>	<b>PARTICIPAT.</b>
<b>Credibilidad exterior</b>	Alta	Baja	Media
<b>Legitimidad en el área</b>	Alta	Media	Alta
<b>Aprendizaje del Grupo</b>	Baja	Media	Alta
<b>Toma de decisiones Grupo</b>	Baja	Alta	Alta
<b>Necesidad de resultados</b>	Baja	Alta	Media
<b>Objetividad</b>	Alta (1)	Baja	Media
<b>Rigor metodológico</b>	Alta	Baja	Baja
<b>Conocimiento del programa</b>	Baja	Alta	Media

(1) Media o Baja si el equipo evaluador tiene una dependencia clara de contratos de esa administración pública.

Tan importante como saber qué tipo de evaluación es el que más se adecua a nuestras necesidades es saber qué queremos evaluar. Aparentemente parece fácil, pero sin duda es una de las tareas más complejas a la hora de diseñar una evaluación, que ya de por sí es extremadamente compleja en el caso de programas de desarrollo rural en los que tan importante como las inversiones realizadas, o la cantidad de inversión privada movilizada en relación o no a la inversión pública, son aspectos como la participación de la población, la mejora del sentimiento de pertenencia territorial, la creación de una conciencia colectiva territorial, la concienciación de la importancia de actitudes de cooperación frente a actitudes de competencia, el liderazgo, etc. Y a nadie se le escapa que estos últimos son aspectos realmente complejos de evaluar con cierta precisión.

Para evaluar es necesario poner a punto un sistema de indicadores de evaluación. Este sistema de indicadores es la clave de la evaluación. El punto de partida para este sistema es una visión clara de las necesidades del área o grupo así y de los objetivos del programa para responder a esas necesidades. En términos generales podemos tener tres grandes tipos de objetivos:

- 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).

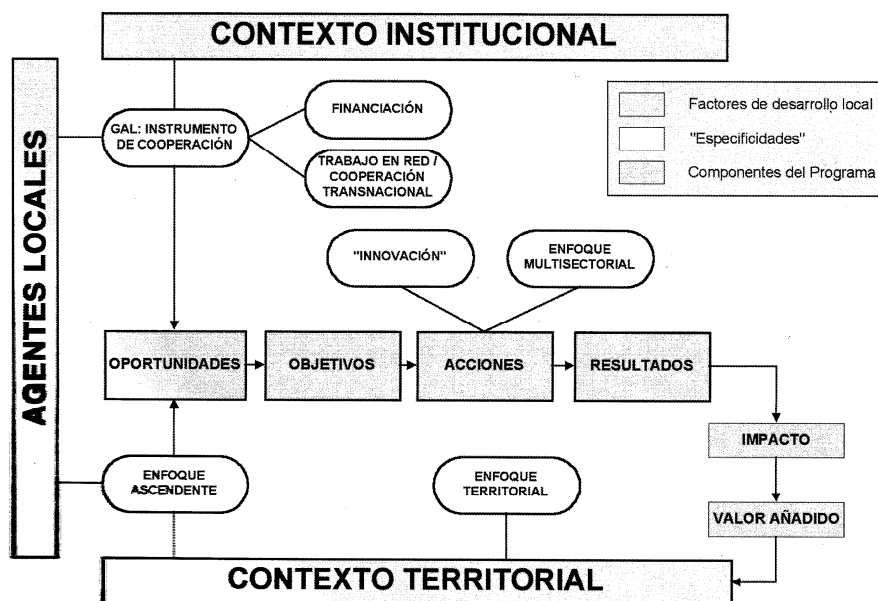
En cuanto a los indicadores, no es posible hacer un listado de los mismos, porque dependerán de lo que se pretenda con la evaluación. Sin embargo, sí es necesario tener en cuenta que los indicadores, prácticamente siempre, responden a cinco grandes tipos:

**1. 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.

**2. 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 resulta-

dos 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), 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.

**3. 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 esperados. 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? ...).

**4. 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.

**5. 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.

A la hora de elaborar la batería de indicadores, tenemos que saber qué es lo que queremos evaluar de los programas de desarrollo rural. Existen evaluaciones muy completas y sistemáticas. Sin embargo, al nivel de los Grupos y de cara a poner en marcha algunos elementos de autoevaluación, vamos a hacer referencia aquí únicamente a un esquema que recoge los aspectos más relevantes de la Iniciativa LEADER, y que pueden y deben ser objeto de evaluación interna por parte de los Grupos. En la

siguiente figura se recogen los tres factores básicos del desarrollo local-rural, el contexto institucional, los agentes locales, y el territorio. En torno a estos factores se diseña un programa que pretende maximizar y aprovechar unas oportunidades, con unos objetivos, unas acciones, unos resultados y un impacto o efectos, principalmente sobre el territorio. Este es el valor añadido más importante de LEADER o PRODER, el canal que permite generar y canalizar sinergias hacia el propio territorio. En este contexto, el enfoque LEADER, y en gran parte también PRODER, tiene una serie de características que explican y determinan que este sistema funcione correctamente y que realmente se generen esas sinergias para el territorio y sus gentes. Estas características o especificidades LEADER son las que en principio son susceptibles de evaluar por parte de los Grupos, porque cómo esté funcionando cada una de ellas nos da las claves del éxito o fracaso del Programa a la hora de alcanzar los objetivos últimos para los que está diseñado<sup>2</sup>.

---

<sup>2</sup> La fuente para la preparación de este material es el cuaderno metodológico nº 4 del Observatorio Europeo LEADER: Evaluar el valor añadido del enfoque LEADER.