

Marielle Berriet-Sollicec¹

Jacques Baudry²

Pierre Labarthe³

*A avaliação de políticas à prova dos fatos:
as contribuições dos métodos “evidence
based policies”**

Na França, no contexto da modernização das políticas públicas, a avaliação aparece como uma ferramenta para aumentar a eficiência da ação política, melhor controlar os gastos públicos e tornar mais transparente, para o cidadão, o uso de fundos públicos. Contudo, as práticas de avaliação refletem realidades muito diferentes tanto do ponto de vista conceitual quanto

* Este artigo se baseia em um trabalho de pesquisa coletivo realizado dentro de um programa intitulado EBP-BIOSOC e coordenado por Catherine Laurent do Instituto Nacional de Pesquisa Agrônômica/INRA-Sadapt, França. Este programa, que há três anos (2007-2010) integra equipes francesas, brasileiras e sul-africanas de pesquisadores, é financiado pela Agência Nacional de Pesquisa (ANR, França) por intermédio do edital “Agricultura e desenvolvimento sustentável” (ADD). Tradução: Tereza da Silva Rosa. Revisão técnica: Sérgio Pereira Leite.

¹ Pesquisadora do ENESAD-CESAER, Dijon. E-mail: berriet@enesad.inra.fr.

² Pesquisador do INRA Rennes. E-mail: jacques.baudry@rennes.inra.fr.

³ Pesquisador do INRA-SADAPT, Paris. E-mail: pierre.labarthe@agroparistech.fr.

metodológico. Às vezes, devido a uma inadequação dos modelos teóricos utilizados, os métodos de avaliação nem sempre se revelam aptos a considerar as transformações técnicas e socioeconômicas. Disso resulta uma incapacidade dos resultados da avaliação em esclarecer o gestor público e orientar suas escolhas.

Partindo dessa constatação, este artigo se propõe, num primeiro momento, a apresentar, a partir de um ponto de vista crítico, as principais etapas avaliativas dos métodos mobilizados nos dias de hoje. Esta análise permite sublinhar a crescente defasagem existente entre as soluções desejadas e a realidade observada nas abordagens econômicas mais usuais, questionando a validade empírica dos modelos econômicos utilizados para a avaliação das políticas públicas.

Essas constatações se apoiam na ilustração de dois estudos de caso: de um lado, aquele da extensão rural⁴ e, de outro, o das medidas agroambientais. O interesse sobre o exemplo da extensão rural se deve ao fato de que, por trás da grande diversidade das políticas públicas de apoio à extensão rural, estão escondidas controvérsias sobre esquemas explicativos relativos à eficácia destas políticas (LABARTHE, 2006,

⁴ A extensão rural (no texto original em francês, *conseil agricole* – NT) pode ser definida como uma atividade de serviço que permite acompanhar as mudanças nas práticas dos agricultores (LAURENT *et al.*, 2006). Esses serviços são considerados como uma alavanca de ação para enquadrar as mudanças técnicas desejadas nos estabelecimentos agropecuários *vis-à-vis* diferentes objetivos: produtividade, proteção do meio ambiente etc. Os investimentos públicos nessa forma de apoio técnico apresentam uma grande diversidade de modelos e foram objeto de numerosas reformas (AMON, 1989; RIVERA e ZIJP, 2001), quer nos países industrializados, quer naqueles em desenvolvimento (RIVERA, 2002). Essas reformas referem-se às modalidades de financiamento e de gestão da oferta de serviços, ao tipo de relação com a pesquisa ou ainda à definição de metas públicas.

CARNEY, 1995). As políticas agroambientais, principalmente aquelas sobre a biodiversidade, foram implantadas como resposta às reduções da fauna e da flora observadas em paisagens agrícolas. Os debates sobre sua eficácia ou mesmo sobre os seus objetivos são numerosos (KLEINJ e SUTHERLAND, 2003).

Nos dois casos, as dificuldades de acesso aos dados para avaliar as intervenções fazem com que seja crucial o desenvolvimento de métodos de avaliação voltados para um melhor conhecimento dos fatos e dos mecanismos subjacentes às intervenções a serem avaliadas.

Relativamente pouco estudada por pesquisadores do desenvolvimento econômico, esta questão é desde já objeto de estudos nas áreas da medicina e da educação (PEERSMAN, *et al.*, 2001; DAVIES e NUTLEY, 2001). Um dos caminhos propostos é o recurso aos métodos de decisão baseados em evidência (EBD), resultando nas políticas baseadas em evidência (EBP) (DAVIES e NUTLEY, 2001) que propomos apresentar na segunda parte deste artigo, antes de discutir, numa terceira seção, as condições de implementação e as dificuldades específicas a este método.

Diversidade de procedimentos avaliativos: qual a validade empírica dos modelos de avaliação?

A literatura sobre a avaliação de políticas públicas faz alusão a uma grande diversidade de posturas metodológicas e de modelos a tal ponto que certos autores propuseram classificações para auxiliar os patrocinadores a se orientarem frente às diferentes práticas avaliativas. Hansen (2005) distingue, então, seis categorias de modelos de avaliação, dentre os quais os modelos econômicos *stricto sensu* (análise custo-vantagem, custo-eficácia ou custo-benefício), os modelos sistêmicos, os modelos de atores ou, ainda, os modelos baseados na teoria do programa

(*programme theory models*). Particular atenção será dada a esses modelos que enfatizam a análise dos mecanismos subjacentes à intervenção pública. Após apresentá-los brevemente, iremos ressaltar os principais métodos que permitem analisar as relações causais associando a implementação de uma medida às suas potenciais repercussões sobre um alvo (indivíduo ou área). Mas, no caso em que tal tipo de análise é feita, raros são os trabalhos que fazem uma confrontação empírica das hipóteses propostas por estes modelos.

Os modelos de avaliação baseados na teoria do programa

Os modelos baseados na teoria do programa, quando são mobilizados em procedimentos de avaliação de política, visam à análise e à interpretação das relações de causa e efeito que podem existir entre a implantação de uma medida pelos poderes públicos e os possíveis efeitos que daí resultariam. Os efeitos podem se referir diretamente ao alvo visado pela medida ou, ainda, ao contexto ao qual ela se aplica. O papel do avaliador consiste, mais precisamente, em emitir hipóteses sobre o encadeamento das relações causais entre uma medida e seus potenciais efeitos sucessivos. Trata-se de construir o que é usualmente chamado de uma teoria de ação ou, ainda, um esquema de causalidade (TOULEMONDE, 1997). Quando estabelecido, tal esquema constitui um referencial, uma base do dispositivo avaliativo, a partir do qual o avaliador pode propor indicadores ou fazer um julgamento de valor sobre a base de critérios de avaliação (BERRIET-SOLLIEC, 2007). Embora os modelos baseados na teoria de programa tendam a se desenvolver (HANSEN, 2005), constata-se que eles permanecem, ainda, relativamente raros nas práticas dos avaliadores.

No caso da extensão rural, por exemplo, existem poucos trabalhos de avaliação das políticas propondo tais teorias

de ação das políticas (EVENSON, 1989). A maior parte das avaliações se limita aos recenseamentos de meios colocados em prática, sob a ótica da contabilização dos investimentos realizados juntos aos financiadores (RICHARDSON e MOORE, 2002). Isto talvez esteja relacionado às dificuldades de construção desses esquemas de causalidade. Elas são bem identificadas no caso de serviços de extensão (GADREY, 1996): dificuldades de medida do produto e, portanto, da produtividade dos serviços, complexidade da identificação de seus efeitos diretos (transformação das práticas e dos desempenhos dos estabelecimentos agropecuários) ou indiretos (aprendizagem) etc.

No caso das medidas agroambientais, existem dois tipos de avaliação: uma avaliação do número de agricultores que optaram, voluntariamente, por uma medida; e uma avaliação dos efeitos das medidas sobre a biodiversidade. Mas não há objetivos claramente definidos permitindo dimensionar o sucesso de uma medida. Os esquemas de causalidades aos quais se recorre são sumários e são, frequentemente, limitados a “menos insumos (fertilização, pesticida) = mais biodiversidade”. Devido ao fato de as medidas se aplicarem às unidades de produção, o contexto regional ou da paisagem, que controla a possibilidade de novas espécies colonizarem as áreas produtivas objeto de financiamento público,⁵ não é levado em consideração na implantação. Porém, os trabalhos avaliando os efeitos sobre a biodiversidade tendem a considerá-lo (KLEIJN *et al.*, 2006).

Quando os avaliadores se baseiam em esquemas de causalidade, três constatações apontam para os limites das práticas atuais

⁵ No contexto das políticas agroalimentares na França, os agricultores se engajam, por meio de um contrato, a cumprir, como contrapartida do financiamento recebido, durante um período determinado, algumas tarefas, tais como manutenção de cercas vivas, gestão de riachos, manutenção de espécies ameaçadas de extinção (NT).

(TOULEMONDE, 1997). Inicialmente, a avaliação, na maior parte dos casos, visa a compreender os efeitos, muito raramente a verificá-los e a confrontá-los com os fatos. Em segundo lugar, a avaliação revela os fatores que influenciam a eficácia sem comprovar os mecanismos que demonstram como os objetivos são alcançados. Finalmente, a avaliação reúne, na maioria dos casos, diferentes pontos de vista sobre uma teoria da ação e tenta chegar a uma visão aceitável da política ou do programa estudado, a um esquema de causalidade que reflete, na melhor medida do possível, as opiniões dos envolvidos (*evaluation-theorisation, theory building evaluation*). Muito raramente, a avaliação projeta um esquema de causalidade e o confronta com os fatos (*theory driven evaluation*).

Três métodos para analisar as relações de causa e efeito

Entre os métodos de avaliação, existem três vias para apreender os esquemas de causalidade. A primeira é a análise de política que focaliza a atenção dos avaliadores sobre as redes de atores e as relações dos diferentes segmentos envolvidos, sobre a hierarquização dos objetivos do programa estudado, enfim, sobre as modalidades adotadas para a implantação das ações. A análise da política (ou do programa que a implementa) conduz, a partir do estudo aprofundado dos documentos e das entrevistas com os seus formuladores, ao estabelecimento de um referencial sob a forma de gráfico ou diagrama de objetivos ou, ainda, de diagrama lógico de impactos (BERRIET-SOLLIEC, 2007). A segunda via de investigação científica é a análise microeconômica das relações de causa e efeito. Neste caso, na linha de trabalhos microeconômicos de Heckman (HECKMAN *et al.*, 1999), trata-se de testar, por meio de ferramentas da econometria, a robustez das ligações causais entre variáveis (BONDONIO, 2005). Por fim, as pesquisas experimentais se desenvolvem com a instalação de dispositivos de acompanhamento e de avaliação, desde

a implantação de intervenções públicas, que se traduzem, frequentemente, na identificação de grupos de controle e na distinção de indivíduos com ou sem apoio público. Na França, a implantação do salário mínimo de solidariedade ativa⁶ se beneficiou, a título experimental, deste tipo de dispositivo.

No caso da extensão rural, três métodos de avaliação podem ser encontrados de modo quase similar.

1. Certos trabalhos se fundamentam em teorias sistêmicas para idealizar os métodos de concepção e de avaliação de políticas de extensão rural, baseados na participação de diferentes atores envolvidos. Isto se reflete, notadamente, na utilização da metodologia *Soft System Methodology*⁷ (CHECKLAND, 1981). Por exemplo, este método foi utilizado para a construção de uma avaliação de política de formação de extensionistas rurais de uma universidade nas Filipinas (NAVARRO *et al.*, 2008). Tal método não inclui a análise dos efeitos e do impacto sobre o campo da política estudada.

2. Outros trabalhos de avaliação se baseiam em resultados de modelizações microeconômicas. Por exemplo, a avaliação das políticas públicas de investimento na extensão rural, frequentemente, redundou em reformas que resultaram na privatização dos serviços. Estas avaliações têm por hipótese que um descompromisso total do Estado aumentaria a eficácia

⁶ Em francês, é o chamado “revenu solidarité active” (RSA) que substitui o “revenu minimum d’insertion” (RMI), medida colocada em prática a partir de 1º. de junho de 2009, visando complementar a renda dos trabalhadores que possuem um salário considerado pequeno, limitado (NT).

⁷ *Soft System Methodology* é uma abordagem capaz de lidar com situações complexas, por articular problemas pouco estruturados, e que pode ser usada na resolução de problemas e na gestão de mudanças (NT).

da oferta de serviços, que se encontraria, então, orientada pela demanda em direção a situações de equilíbrio (CARNEY, 1998). Elas se fundamentam em trabalhos nos quais os esquemas de casualidades se inscrevem na teoria econômica convencional, aplicada aos serviços (DINAR e KEYNAN, 2001). Estes trabalhos raramente são consolidados por confrontações com as realidades locais (DINAR, 1996). Esta distância dos fatos empíricos é tão importante que certos autores discutem a validade de aplicar tais modelos ao caso dos serviços de extensão rural (HANSON e JUST, 2001) porque sua axiomática não permite considerar certas especificidades fundamentais destes serviços (assimetria da informação, caráter não exclusivo e não antagonista dos conhecimentos etc).

3. Existem, enfim, alguns trabalhos fundamentados em métodos experimentais quantitativos. São avaliações baseadas em ferramentas estatísticas para isolar os efeitos de políticas de apoio à oferta ou à demanda na assistência técnica. A escassez das análises empíricas quantitativas sobre a extensão rural é, em parte, devida à dificuldade de medir seu impacto sobre o desempenho dos estabelecimentos agropecuários (EVENSON, 1997, ESPOSITI, 2002). Recentemente, Davis e N’Konya (2008) propuseram o emprego de *métodos quase experimentais* (MQE⁸) para comparar, no tempo, os ganhos de produtividade de estabelecimentos beneficiando-se ou não de apoio público (participação nas escolas de formação para agricultores⁹) e isolar, assim, o efeito de políticas públicas de investimento na extensão rural sobre a redução da pobreza em áreas rurais da África do Leste. Contudo, estes tipos de trabalhos necessitam de protocolos muito densos de coleta de dados e são, ainda, muito raros.

⁸ Sigla em francês para *méthodes quasi-experimentales* – MQE (NT).

⁹ Em inglês, seriam as *farmers field schools*, as quais focalizam a formação em determinadas áreas ou temáticas (NT).

A mensuração dos efeitos das medidas agroambientais sobre a biodiversidade baseia-se numa amostragem aleatória comparando as parcelas submetidas à medida com as parcelas não submetidas. Como não existe avaliação desde o início da implementação da medida, não se pode estimar a sua evolução. As distorções ligadas ao fato de que as parcelas submetidas a medida tenham, geralmente, características específicas (menor uso) dificultam a explicação das diferenças constatadas nas parcelas que não foram objeto de algum contrato com o poder público. Como o contrato se baseia nas práticas de usos, pode ser que os usos estabelecidos no contrato sejam idênticos aos implantados anteriormente.

A quase inexistência de confrontação dos fatos nos modelos de avaliação

Qualquer que seja a via escolhida para proceder à avaliação, a confrontação com os fatos tem importância relativa nos modelos de avaliação, desde a sua inexistência (raciocínio unicamente a partir da modelização teórica) até o empirismo exclusivo (construção da argumentação a partir dos fatos). Na maior parte das situações, o avaliador é confrontado com dois obstáculos principais: a falta de conceituação teórica para apreender os fatos novos e as lacunas estatísticas para testar as hipóteses (LAURENT, 2007).

No campo do agroambientalismo, a passagem de medidas reguladoras restritas às obrigações de meios às medidas com obrigações de resultados vai obrigar, de um lado, os formuladores de medidas a terem acesso aos modelos probabilísticos de mudanças e, de outro, os avaliadores a fazerem as observações de campo. Na verdade, haverá uma confrontação entre modelos e fatos. O debate entre determinismo e probabilismo na dinâmica ecológica ocorrerá, sem dúvida, sempre que o acaso tenha um papel importante.

Então, a questão sobre a validação empírica dos métodos de avaliação se coloca: como as relações de causalidade são testadas e validadas? Como conciliar a construção de modelos teóricos com o conhecimento das realidades observadas, considerando os trabalhos resultados do realismo crítico (LAWSON, 2003)?

Nesse contexto e frente aos desafios metodológicos, os métodos de decisão baseada em evidências – *d'evidence based decision* (EBD) – existem, já há 20 anos, como alternativa visando desenvolver, nas práticas avaliativas, um conteúdo empírico aos conhecimentos produzidos e submetidos aos gestores públicos, tomadores de decisão.

O recurso aos métodos "Evidence Based": uma via a ser explorada na avaliação de políticas?

Esta segunda parte do artigo analisa o interesse no uso do método da EBD para a avaliação das políticas públicas. Ainda muito raramente empregadas na França (LAURENT, 2006), estas práticas são, nos dias de hoje, mais comuns no Reino Unido, em particular desde 1997 com a modernização de políticas públicas.¹⁰ Neste país, a busca pelas "evidências"¹¹ aparece como elemento essencial no auxílio à decisão pública,

¹⁰ Aqui, os autores fazem referência ao vasto programa de modernização empreendido pelo governo de Tony Blair, a partir de 1997, quando o método de avaliação baseada em evidências foi empregado (NT).

¹¹ A palavra "evidência" é de difícil tradução em francês e se refere à noção de evidência empírica, de prova ou, ainda, de dados comprobatórios (LAURENT, 2007) desenvolvidos, inicialmente, no campo da medicina (GUYATT *et. al.* 1992). Estes procedimentos visam promover o uso de conhecimentos científicos disponíveis para os gestores públicos e, isso, do modo mais judicioso e pertinente possível, considerando o campo da validação desses conhecimentos e a dimensão relativa dos resultados.

a tal ponto que o governo preconiza “*the better use of evidence and research in policy making*” (citação de UK-White Paper, 1999, *apud* CARTWRIGHT, 2007).

É por esta razão que os principais instrumentos utilizados nos procedimentos da EBD são, aqui, evocados, enfatizando, de um lado, o interesse pelos estados das artes sistemáticos da literatura científica e, de outro, as técnicas quantitativas considerando o viés de seleção.

O interesse dos métodos da EBD é duplo, na medida em que eles permitem, ao mesmo tempo, um melhor conhecimento do processo de tomada de decisão, graças às entrevistas junto aos implementadores e tomadores de decisão, como também uma avaliação baseada na validação empírica de hipóteses sobre a relação de causa e efeito quando da aplicação de uma ação ou de uma intervenção pública. Este trabalho deve conduzir, na prática, à criteriosa utilização dos conhecimentos científicos efetivamente disponíveis.

Essa análise científica se apoia em vários métodos para caracterizar a evidência com relação à sua capacidade de testar uma hipótese. Esses métodos podem ser hierarquizados segundo a qualidade das provas que produzem (CARTWRIGHT, 2007). O primeiro método é a realização da metaanálise, graças às revisões sistemáticas da literatura (SUTHERLAND *et al.*, 2004). O segundo é a utilização de técnicas estatísticas como os ensaios controlados, randomizados, realizados com grupos de controle ou de coorte. Estas técnicas são mobilizadas no campo da avaliação de políticas públicas para testar a certeza das relações de causa e efeito (SCHMITT *et al.*, 2003; ZAHM *et al.*, 2008). Enfim, os estudos de caso e as opiniões de especialistas permitem testar as relações de causa e efeito, mas contribuem pouco para elevar estas relações ao estatuto de evidência. Retomamos, então, a apresentação dos dois primeiros tipos de procedimentos, permitindo elevar uma relação de causa e efeito ao estatuto de evidência.

O método de estados da arte realizados segundo os critérios explícitos (EARC)

O método EARC visa fazer um levantamento, o mais exaustivo possível, de estudos e de provas existentes sobre um dado assunto. Ele se fundamenta em uma metodologia particular que se decompõe em várias etapas, bem definidas na literatura (PULLIN *et al.*, 2004; DAVIES e NUTLEY, 2001; PEERSMAN *et al.*, 2001). Entre estas etapas, uma das fases essenciais consiste na triagem de referências segundo seu grau de validação.

Assim, após selecionados os estudos sobre o assunto pesquisado, trata-se de analisar essas referências com base em critérios ditos de “qualidade”. Estes são contingentes da questão colocada. Por exemplo, para uma questão relativa à eficácia de uma medida de proteção da biodiversidade, Pullin *et al.* (2003) propõem uma classificação de vários níveis de qualidade que é determinada pelo nível de validade empírica das provas produzidas:

I – Provas fortes obtidas a partir de, pelo menos, um ensaio aleatório convenientemente concebido;

II – 1 Provas obtidas graças a experiências (convenientemente concebidas) não aleatórias;

II – 2 Provas obtidas graças à comparação de diferenças entre áreas (lugares) – com ou sem uma espécie ou comunidade cuja presença é desejável para a análise;

II – 3 Provas obtidas a partir de uma série múltipla ou de um resultado catastrófico nas experiências não controladas;

III – Opiniões de autoridades legítimas, apoiando-se em provas qualitativas obtidas no campo ou em relatórios de especialistas;

IV – Provas inadequadas quanto aos problemas de método (tamanho da amostragem), de compreensão ou de conflito de resultados.

Nessa fase da análise, os estudos e os resultados são separados do seguinte modo:

- sobre a base da incoerência de suas premissas (incoerência entre as hipóteses, entre os corpos teóricos selecionados e os estudos empíricos que os aplicariam), ou
- para um certo número de estudos considerados como pertinentes, sobre a base dos problemas importantes de métodos que eles levantam – estudos mal feitos ou métodos não explícitos etc.

Os métodos quantitativos de medida de impactos

Quando a ajuda à decisão pública se refere aos programas ou às políticas que já foram colocadas em prática (caso da situação de avaliação *post factum*), o teste de uma hipótese e a obtenção da evidência sobre a relação de causa e efeito podem, igualmente, emergir do recurso a técnicas quantitativas de avaliação de impactos da política estudada.

O impacto de uma política é definido como a diferença entre a situação observada tendo em vista a existência da política e aquela que teria prevalecido na sua falta. Esta última situação hipotética e não observada é chamada de situação contrafactual. O avaliador, seguindo as demandas dos tomadores de decisão, ou adotando uma ou várias visões dos objetivos da política, seleciona uma ou algumas variáveis-objetivos que a política deveria afetar. É a comparação entre o nível destas variáveis, na existência ou na ausência da política, que constitui o impacto da mesma (ZAHM *et al.*, 2008).

O objetivo é estimar o impacto médio da ação ou, mais amplamente, determinar a situação contrafactual (FORAY e

COWAN, 2002). Assim, ocorre um problema de identificação, o que se traduz, ao nível individual, por uma questão de ausência de dados e, mais globalmente, por um problema qualificado como viés de seleção. Para resolver este problema metodológico, duas soluções são possíveis: a comparação com/sem a intervenção pública; ou a comparação antes/depois da ação.

As principais metodologias utilizáveis poderiam ser (SCHMITT *et al.*, 2003):

- as estimativas sobre dados de amostragem (grupos, coortes), se o esperado é avaliar o desempenho sobre os dados agregados;
- os métodos de “pareamento” (*matching*) com e/ou sem reposição, bem como as estimações em dupla diferença, eventualmente combinadas, que repousam sobre o princípio dos grupos de controle;
- por último, se os dados disponíveis se “limitam” às características iniciais e às atuações finais dos indivíduos e/ou áreas focalizadas (beneficiários e não beneficiários), deveremos nos “contentar” com o uso do método do *matching* com ou sem reposição.

Em todo caso, a avaliação dos efeitos específicos passa por um bom conhecimento do programa e por uma identificação dos efeitos esperados de cada política. Esta etapa é indispensável por diversas razões. De um lado, ela permite escolher os indicadores de desempenho mais pertinentes para a identificação dos efeitos específicos do programa. De outro lado, é na confrontação dos processos de evolução das situações socioeconômicas potenciais ou observadas, identificadas na fase de análise da política com os resultados econométricos da análise dos efeitos específicos, que são procurados os efeitos reais do programa. Na verdade, a grande dificuldade de caracterização das relações de causalidade, em particular nos

casos de programas de desenvolvimento com objetivos amplos como, por exemplo, as políticas territorializadas, não permite às ferramentas estatísticas darem respostas unívocas e seguras quanto aos efeitos específicos da política avaliada. Desse modo, a explicitação dos resultados por meio dos efeitos potenciais ou observados com base na teoria da ação permite uma conclusão mais segura.

Consequentemente, os resultados estatísticos obtidos permitem um teste de solidez das relações de causa e efeito testadas. Mas, mesmo se os resultados obtidos são provados estatisticamente, eles não podem provar nada sobre o sentido próprio da relação (SCHMITT *et al.*, 2003). Por isso, é essencial relativizar o recurso a tais procedimentos ou, pelo menos, conjugá-los com as análises qualitativas complementares (BERRIET-SOLLIEC, 2007).

Ao final, o emprego dos métodos da EBD permite a identificação de mecanismos socioeconômicos recorrentes, testados estatisticamente. Neste sentido, esses métodos constituem uma ferramenta essencial para consolidar, desde o início de um programa de intervenção pública, as teorias de ação das diferentes medidas deste programa e os incentivos das ações que estão na base de cada uma dessas medidas, graças à mobilização dos dados quantitativos e qualitativos. Apesar de tudo, a persistência de lacunas metodológicas leva a ultrapassar certo “efeito de moda” e a relativizar o recurso a estes métodos.

Implantação e dificuldade de aplicação do método EBP

O emprego de métodos baseados em evidências subsidiando a decisão pública exige dedicar financiamentos específicos no incentivo e desenvolvimento das pontes entre institutos de pesquisa, universidades e instâncias governamentais. Este tipo de estrutura já existe desde 2000 no Reino Unido: o Centro de Coordenação e Informação sobre a Evidência para a Política e

para a Prática¹² (OLIVER *et al.*, 2005). Tais práticas não existem na França. Uma das razões é, sem dúvida, um desconhecimento das abordagens *evidence-based* (LAURENT, 2006, RICOCH e BONNAFOUS, 2008).

A necessidade de estruturas de transferência

O estado das artes realizados segundo critérios sistemáticos é ferramenta importante nos métodos EBD na medida em que permite confrontar as conclusões de numerosos estudos e evitar manter proposições que conduzem a resultados pouco convincentes. Bem conhecida nesses 20 anos, essas técnicas estão hoje sujeitas a críticas em razão da relativa falta de rigor na utilização desses métodos e da falta de consideração do contexto no qual os estudos foram feitos.

Frente a tais dificuldades, o estabelecimento de organizações específicas para implantar os métodos de EBD pode se revelar, sob diferentes pontos de vista, útil e necessário. De um lado, os métodos quantitativos de avaliação de impacto das políticas criam necessidades de coleta de dados que tornam indispensável um trabalho coerente e centralizado. Assim, o estabelecimento do estado das artes, segundo critérios explícitos (EARC), necessita da colaboração e das trocas diretas entre pesquisadores aptos a avaliarem os trabalhos de pesquisa em diferentes disciplinas ou vários campos teóricos. De outro lado, essas organizações podem favorecer as relações e o acesso mais direto dos formuladores das políticas aos resultados de procedimentos científicos segundo a metodologia EBP.

Por exemplo, a Unidade de Pesquisa em Ciências Sociais (SSRU¹³), da Universidade de Londres, junto com o Centro

¹² *The Evidence for Policy and Practice Information and Co-ordination Center-EPPICenter* (NT).

¹³ Em inglês, Social Science Research Unit-SSRU (NT).

de Coordenação e Informação sobre as Práticas e as Políticas Baseadas em Evidência (EPPI *center*), propõe novos modelos para o emprego de revisões sistemáticas de literatura a partir de procedimentos conjugando dados quantitativos e qualitativos (OLIVER *et al.*, 2005). A originalidade é abordar o estado da artes combinando dois objetivos: (i) obter as informações para os gestores públicos sobre a implantação das intervenções governamentais; (ii) implantar um dispositivo de avaliação destas intervenções. No protocolo para a realização do estado das artes, os autores distinguem, então, os estudos relativos aos pontos de vista expressos pelas partes envolvidas e outros relativos aos impactos da política estudada. O valor agregado deste tipo de procedimento dos pesquisadores do EPPI *center* é, mais globalmente, o de conjugar este interesse metodológico com a dimensão teórica (neste caso, tratar-se-ia dos avanços teóricos em matéria de saúde pública dos jovens).

No caso da agricultura e dos serviços de extensão, organizações específicas foram implantadas para favorecer o acesso dos gestores públicos aos conhecimentos científicos para a concepção e avaliação de políticas agrícolas. Por exemplo, este é o caso do Grupo Consultivo em Pesquisa Agrícola Internacional¹⁴ (CGIAR) ou do Instituto Internacional de Pesquisa em Políticas Alimentares¹⁵ (IFPRI) que propõem pesquisas para orientar políticas de redução da fome no mundo. No interior deste instituto ocorrem debates sobre a possibilidade de melhor conectar pesquisa e avaliação das políticas agrícolas através dos métodos de EBD. No caso da extensão rural, este instituto propõe, por um lado, combinar as abordagens institucionais e teóricas dos planos de causalidade de reformas da extensão rural

¹⁴ Em inglês, Consultative Group on International Agricultural Research-CGIAR (NT).

¹⁵ Em inglês, International Food Policy Research Institute-IFPRI (NT).

já colocadas em prática e de fazer uma síntese de sua eficácia nos diversos contextos ("*from best practices to best fit*", cf. BIRNER *et al.* 2006); e, por outro, uma avaliação quantitativa ex-post dos efeitos destas reformas (DAVIS e N'KONYA, 2008). Todavia, este procedimento não se baseia, no momento, explicitamente nos métodos EDB e nas metodologias validadas dos estados da arte sistemáticos.

No campo da biologia da conservação, no qual se encontram as medidas agroambientais, um grupo trabalha ativamente para produzir os EARC (PULLIN *et al.*, 2004). Os limites do procedimento são consideráveis devido à falta de dados (problema de escala espacial e temporal, principalmente). Isso leva alguns a rejeitar por enquanto tais métodos (FERRARO e PATTANAYAK, 2006) ou a chamar a atenção para os riscos de impasse em razão da estreiteza das situações analisadas (LUDWIG *et al.*, 2001).

Dois obstáculos maiores: falta de teorização e de dados

Se o entusiasmo pelas práticas baseadas em evidências é notório, estes métodos, relativamente recentes, apresentam várias lacunas, tanto nas práticas científicas como nas tomadas de decisão políticas. Identificaremos duas principais. A primeira é a falta de conceito preciso sobre o que é, realmente, uma evidência (CARTWRIGHT, 2007; KIRSCH e LAURENT, 2007). A segunda diz respeito à falta de precisão sobre o uso de diferentes tipos de evidência e sobre as suas áreas de validação para testar uma hipótese emitida pela comunidade científica. Esta última observação alude, indiretamente, à qualidade dos dados disponíveis.

Com efeito, no campo da avaliação das políticas, os métodos mobilizados para estimar os seus impactos estão intimamente ligados à disponibilidade de dados relativos não somente ao grupo de beneficiários, mas também a um

grupo de comparação. Esses dados devem compreender as características descritivas dos indivíduos, notadamente aquelas que, inicialmente, determinaram tanto a seleção do programa, os indicadores de comportamento (cuja escolha deve estar condicionada pelos objetivos esperados do programa), como os seus efeitos potenciais e, na medida do possível, um conjunto de variáveis explicativas dos indicadores de desempenho levados em conta. De modo evidente, quanto mais abundante em relação a recortes temporais e mais disponível a informação, maior será a qualidade das estimativas em termos de correção de distorções potenciais e de precisão dos avaliadores. Pelo menos, três datas de coleta de dados antes da implantação do programa e três datas posteriores são exigidas para obter-se um nível de confiabilidade aceitável. Dispor de um acompanhamento temporal regular das diversas variáveis para o conjunto de indivíduos que comece, ao menos, três anos antes da implantação do programa, corresponde à situação “ideal”.

No caso das medidas agroambientais, faltam informações sobre as relações de causa e efeito entre práticas agrícolas e espécies-alvo. Além disso, as escalas espaciais de intervenção são, geralmente, inadequadas. Neste sentido, o recurso aos métodos da EDB aparece como uma das ferramentas de consolidação das práticas avaliativas e de aperfeiçoamento das recomendações resultantes. Isto requer o estabelecimento da experiência de avaliação considerando o conjunto de fatores que possam explicar as variações da biodiversidade, desde as flutuações naturais das populações animais e vegetais até a história das práticas nas parcelas agrícolas e a história da paisagem que podem explicar as facilidades e impossibilidades da chegada de novas espécies nas parcelas sob contrato e apoio do Estado. Num primeiro momento, é preciso: 1) definir a escala espacial na qual serão avaliados os resultados e 2) construir um modelo, ligando, explicitamente, as práticas agrícolas à biodiversidade.

O primeiro aspecto é crucial. Como conclusão do EARC sobre “a eficácia de medidas baseadas na terra (inclusive agroambientais) para a conservação da densidade de pássaros em terras aráveis no Reino Unido”, Roberts e Pullin (2007) escrevem: “É desconhecido se as espécies são simplesmente redistribuídas entre recursos disponíveis por meio do acréscimo nas áreas sob a gestão agroambiental e abandonando áreas convencionais (nenhuma alteração na tendência populacional), ou conseguindo o aumento do sucesso da reprodução ou a taxa de sobrevivência pós-inverno (tendência populacional positiva), dada a evidência da recuperação das espécies”. Então, observa-se que há tanto um problema de definição das políticas quanto de limites dos EARC.

Na sua síntese sobre as relações entre agricultura e biodiversidade, Le Roux *et al.* (2008) evidenciam a sinergia entre as práticas nas áreas agrícolas e as estruturas da paisagem nas dinâmicas da biodiversidade. Eles chamam também atenção para a falta de pesquisas tratando, explicitamente, da integração da biodiversidade nas propriedades agrícolas.

Discussão: quais são as perspectivas de desenvolvimento dos métodos baseados em evidência nas práticas avaliativas?

Por conseguinte, o desenvolvimento das políticas baseadas em evidência (*evidence-based policies*) para a avaliação de políticas públicas é interessante por diversas razões. Do ponto de vista político, o método obriga os gestores públicos a integrar, mais ainda, os resultados de pesquisas no processo e incentiva-os a praticarem as avaliações anteriores a implantação da medida (*ex-ante*) que se baseiam em teorias de programas estudados e corroborados aos fatos. Do ponto de vista científico, ele encoraja os pesquisadores a atentarem, ainda mais, para os mecanismos e os impactos de decisões públicas e a desenvolverem os

métodos de medição do grau de impactos, pensados a partir da hierarquia dos objetivos inicialmente estabelecidos. Tais desenvolvimentos evidenciam a necessidade de melhor precisar os marcos conceituais e de prosseguir na construção de bases de dados necessárias para mitigar a falta de dados.

Apesar de escutarmos falar de políticas de transporte baseadas em evidência (CARTWRIGHT, 2007), não existem, ainda, políticas de agricultura baseadas em evidências. E, portanto, seria judicioso mobilizar mais tais análises no contexto europeu de reforma profunda da política agrícola e de desenvolvimento rural com vistas a 2013, enfatizando tanto a síntese de conhecimentos existentes como, mais amplamente, o interesse de reunir produções científicas sobre um conhecimento das realidades observadas. Assim, além da pesquisa das “evidências”, os métodos de EDB apontam para a importância de reunirem-se abordagens positivas de auxílio à decisão pública, num contexto onde predomina o recurso às perspectivas normativas e ao emprego de simulações em detrimento de dados de observação.

Um dos aspectos previstos quanto ao desenvolvimento dos métodos baseados em evidência seria o de tirar proveito dos resultados de inúmeras avaliações feitas, até aqui, pela Comissão Europeia. As práticas avaliativas são, na verdade, bastante conhecidas na Europa, em particular através de intervenções da Comunidade Europeia. A avaliação ocorre em diversos períodos da política (*ex ante*, intermediário, final e *ex post*): as Direções da Comissão Europeia possuem, assim, inúmeros resultados em seus diferentes Estados-membros. As metasanálises são, certamente, executadas, mas são raras as sínteses sistemáticas realizadas graças às técnicas reconhecidas.

Tal perspectiva exige a mobilização de competências específicas e a criação de uma agência especializada se dedicando à

implantação de métodos de avaliação com base em evidência tal como existem no Reino Unido desde a reforma governamental iniciada em 1999. Esta experiência permite colocar como mote o necessário desenvolvimento conjunto de pesquisas neste campo. Tanto a consolidação de equipes de pesquisa sobre as questões baseadas em evidências como a institucionalização deste procedimento, intermediado por estruturas de transferência especializadas nestes métodos, são, assim, dois fatores que, conjugados, se revelam condição *sine qua non* para o desenvolvimento de políticas baseadas em evidências.

Referências bibliográficas

- ARNON, I. *Agricultural research and technology transfer*. Londres: Elsevier Applied Science, 1989.
- BERRIET-SOLLIEC, M. *L'analyse de politique: une démarche essentielle pour l'évaluation. L'exemple des politiques territoriales. Evaluation des politiques publiques*, Ecole-chercheurs INRA, n. 18-20, junho de 2007.
- BIRNER R., DAVIS K., PENDER J., NKONYA, E.; ANANDAJAYASEKERAM P., EKBOIR J., MBABU A., SPIELMAN D., HORNA D., BENIN S.; KISAMBA-MUGERWA W. *From "best practice" to "best fit": a framework for designing and analyzing pluralistic agricultural advisory services*. IFPRI Research Brief, 2006.
- BONDONIO, D., *Evaluating Decentralized Policies: A Method to Compare the Performance of Economic Development Program across Different Regions or States*. *Evaluation* n. 8 (1), 2002.
- BOWEN, W. *An analytical Review of Environmental Justice Research: What Do We really know?* *Environmental Management*, v. 29, n. 1, 2002.
- CARNEY, D. *Changing public and private roles in agricultural services provision*. Laundress: Overseas Development Institute, 1998.
- _____. *The changing public role in services to agriculture: a framework for analysis*. *Food Policy*, n. 20, 1995.
- CARTWRIGHT N. *With help from Goldfinch A e Ho wick J. Evidence-based policy: Where is our theory of evidence?* Working paper, LSE e UCSD, 2007.
- CHECKLAND, P. B. *System thinking, system practice*. New-York: John Wiley, 1981.

- DAVIES H. T. O., NUTLEY S. M. Evidence-based policy and practice: moving from rhetoric to reality. *Proceedings of the Third international Interdisciplinary Evidence-based Policies and Indicator Systems Conference*, 2001.
- DINAR, A. Extension commercialization: how much to charge for extension services. *American Journal of agricultural economics*, 1996.
- DINAR, A., KARAGIANNIS, G., TZOUVELEKAS, V. Evaluating the impact of agricultural extension on farms' performance in Crete: a nonneutral stochastic frontier approach. *Agricultural Economics*, 36: 2, 2007.
- DINAR, A. e KEYNAN, G. Economics of paid extension: lessons from experiences in Nicaragua. *American Journal of agricultural economics*, 83 (3), 2001.
- ESPOSTI, R. Public agricultural R&D design and technological spill-ins: A dynamic model. *Research policy*, 31: 5, 2002.
- EVENSON, R. E., *Economic issues in agricultural extension policy*. New Haven, CT: Yale University, Economic Growth Center, 1998.
- FERRARO, P. J. e PATTANAYAK, S. K. Money for Nothing? A Call for Empirical Evaluation of Biodiversity Conservation Investments *Plos Biology* 4 (4), 2006.
- FORAY, D., COWAN, R., Evolutionary economics and the counterfactual threat: on the nature and role of counterfactual history as an empirical tool in economics. *Journal of Evolutionary Economics*, 12 (5), verão de 2002.
- GADREY, J. *Services: la productivité en questions*. Paris: Descend de Brought, 1996.
- GUYANA et al., (EVIDENCE-BASED MEDICINE WORKING GROUP) Evidence-Based Medicine. A new approach to teaching the practice of Medicine (1992). *JAMA*, v. 268, n. 17, 1996.
- HANSEN, H. F. Choosing Evaluating Models: a discussion on Evaluation Design. *Evaluation*, v. 11 (4), 2005.
- HANSON J. C., JUST R. E. The potential for transition to paid extension: some guiding economic principles. *American Journal of agricultural economics*, 83 (3), 2001.
- HECKMAN J., BLONDER. E SMITH J., The Economics and Econometrics of Active Labor Market Programs. In: O. ASHENFELTER e D. CARD (eds.), *Handbook of Labor Economics*, v. III, North Holland, 1999.

- KIRSCHM., LAURENT C. *Scientificity, policy making and empirical validity of knowledge: new issues?* Comunicação para o colóquio mundial de filosofia das ciências. Pekin, agosto de 2007.
- KLEIN, D.; SUTHERLAND, W. J. How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Ecology*, 40, 2003.
- KLEIJN, D., R. A. BAQUERO, Y. CLOUGH, M. DÍAZ, J. ESTEBAN, F. FERNÁNDEZ, D. GABRIEL, F. HERZOG, A. HOLZSCHUH, R. JÖHL, E. KNOP, A. KRUESS, E. J. P. MARSHALL, I. STEFFAN-DEWENTER, T. TSCHARNTKE, J. VERHULST, T. M. WEST e J. L. YELA. Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology Letters* 9 (3), 2006.
- LABARTHE, P. *La privatisation du conseil agricole en question. Evolutions institutionnelles et performances des services de conseil dans trois pays européens (Allemagne, France, Pays-Bas)*. Tese de doutorado em Ciências Econômicas, Universidade Paris XII, 2006.
- LAURENT L. (ed.) *Validité des connaissances scientifiques et intervention publique: le cas de l'agriculture dans le développement durable*. Projeto ANR "EBP-BIOSOC. 2007-2010", 2006.
- LAURENT L. *Inégalités sociales, pauvreté et développement durable. Nouveaux enjeux autour des modèles de connaissances qui sous tendent l'action politique*. Comunicação ao colóquio internacional sobre Desenvolvimento Territorial Sustentável. Florianópolis, 22-25 de agosto de 2007.
- LAURENT C., CERF M., LABARTHE P. Agricultural extension and market regulation: learning form a comparison of six EU countries. *Journal of Agricultural Education and Extension*, 2006.
- LAWSON, T. *Reorienting Economics*. Londres: Routledge, 2003.
- LE ROUX, X., R. BARBAULT, J. BAUDRY, F. BUREL, I. DOUSSAN, E. GARNIER, F. HERZOG, S. LAVOREL, R. LIFRAN, J. ROGER-ESTRADE, J. P. SARTHOU e TROMMETTER, M., (eds.). *Agriculture et biodiversité. Valoriser les synergies*. Expertise scientifique collective, synthèse du rapport. Paris, INRA, 2008.
- LUDWIG, D., M. MANGEL *et al.*, HADDAD. Ecology, conservation, and public policy. *Annual Review of Ecology and Systematics* 32 (1), 2001.
- NAVARRO M., ROCHS F., PRUSSIA S. E., KANLAYANARAT S., MCGLOSSON B. Using soft-systems methodology to plan

- advanced academic programs in post-harvest thechnology in thailand. XXIV *annula conference of the assocation for international agricultural and extension education*. Guacimo (Costa Rica), 10-15 de março de 2008.
- OLIVER S.; HARDEN A.; REES R.; SHEPERD J.; BRUNTON G.; GARCIA J.; OAKLEY A. An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation* v. 11 (4), 2005.
- PEERSMAN, G.; OLIVER, S.; OAKLEY, A. Systematic reviews of effectiveness, in S. OLIVIER e G. PEERSMAN (eds.) *Using research for effective Health Promotion*, pp 96-108, Buckingham: Open University Press, 2001.
- PIERSON, P. When effect becomes cause. Political feedback and political change. *World Politics*, 1993.
- PULLIN A.S., KNIGHT T. M., STONE D. A., CHAPMAN K. Do conservation managers use scientific evidence to support their decision-making? *Biological conservation*, 2004.
- RICHARDSON J. G., MOORE C. L. 2002. *Determining extension program economic value* (2002). North Carolina State University, Working paper, AEE, 1-2, 2004.
- RICROCH e BONNAFOUS. *Scientific knowledge used in three French ago-environment schemes related to European agriculture and biodiversity policies*. Comunicação ao seminário, EBP-BIOSOC. Rio de Janeiro, 28-29 de agosto de 2008.
- ROBERTS, P. D. e A. S. PULLIN. The Effectiveness of Land-Based Schemes. Agri-environment. at Conserving Farmland Bird Densities within the U.K. *Systematic Review* n. 11. Collaboration for Environmental Evidence <http://www.environmentalevidence.org>, 2007.
- RIVERA, W. R. Confronting global market: public sector agricultural extension reconsidered. *Journal of extension systems*, 16, 2000.
- RIVERA, W., ZIJP, W. *Contracting for agricultural extension. International case studies and emerging practices*. Cambridge (USA): CABI Publishing, 2002.
- SCHMITT, B.; LÉPICIER, D.; BERRIET-SOLLIEC, M. *Economic impact evaluation of the european program for rural development in Burgundy: allowance for selection bias*. Dijon: INRA-ENESAD UMR CESAER, Working paper 6, 2003.

TOULEMONDE, J. Faut-il libérer l'évaluation de ses liens de causalité? *Annales des mines*, março de 1997.

ZAHM, F., CHABE-FERRET, S. GASSIAT, A., SUBERVIE, J. *Proposition d'un cadre méthodologique pour évaluer les effets propres des mesures agro-environnementales européennes – application au dispositif français*, Communication au colloque de la Société Française d'Evaluation, Strasbourg, 3-5 de julho de 2008.

BERRIET-SOLLIEC, Marielle, Jacques Baudry e Pierre Labarthe. A avaliação de políticas à prova dos fatos: as contribuições dos métodos "*evidence based policies*". *Estudos Sociedade e Agricultura*, Outubro de 2009, vol. 17, n. 2, p. 227-253, ISSN 1413-0580.

Resumo: (*A avaliação de políticas à prova dos fatos: as contribuições dos métodos "evidence based policies"*). Os métodos habitualmente empregados na avaliação de políticas públicas deparam-se com dois obstáculos fundamentais. O primeiro é a análise das relações de causa e efeito somente a partir de opiniões e de pareceres das partes envolvidas (stakeholders). Quando esses tipos de causalidade mobilizam as teorias em ciências sociais (theory driven evaluation), o segundo obstáculo é que eles são raramente confrontados com os fatos. Para atenuar essa insuficiência de validade empírica, os métodos baseados em evidência (evidence-based) propõem mobilizar os conhecimentos testados empiricamente. Contudo, mesmo se esses métodos são difundidos nas áreas de educação e saúde nos países anglo-saxões, eles são ainda muito raramente utilizados nos dois campos de estudos analisados neste artigo, o da extensão rural/agrícola e o agroambiental.

Palavras-chave: agricultura, políticas públicas, agroambiente, biodiversidade.

Abstract: *(The evaluation of public policies based on facts: the contributions of "evidence based policies" (EBP) methods).* The methods for the evaluation of public policies face two major limits. The first occurs when the schemes of causal relations involved in a given public policy are analysed only through the opinions and knowledge of some stakeholders. When the description of these schemes is instead driven by certain social sciences theories, a second limit may occur: the lack of confrontation with field observations. Both methods raise the question of their level of validity. In that respect, the EBP method proposes to ground the evaluation of public policies on scientific knowledge that relies on the best available empirical evidence. These methods are widely used in the cases of health or education policies evaluation in Anglo-Saxon countries. However, they are much less frequently implemented in the two cases studied in this paper: agri-environment and agricultural extension.

Key words: Agriculture, public policies, agri-environment, biodiversity.