

**A AVALIAÇÃO DO
DESENVOLVIMENTO SOCIOECONÓMICO**

O GUIA

O GUIA

ÍNDICE GERAL

INTRODUÇÃO

- 1. PARTE I: AVALIAÇÃO E DESENVOLVIMENTO SOCIOECONÓMICO**
 - 1.1. As vantagens da avaliação**
 - 1.2. História e objectivo da avaliação**
 - 1.3. Os métodos e respectivas origens**
 - 1.4. A avaliação como meio de reforçar o desenvolvimento socioeconómico**
 - 1.5. Regras de ouro**
- 2. PARTE II: A CONCEPÇÃO E IMPLEMENTAÇÃO DA AVALIAÇÃO**
 - 2.1. A concepção e o planeamento da avaliação**
 - 2.2. A implementação e a gestão da avaliação**
 - 2.3. A utilização da avaliação e a gestão de conhecimentos**
 - 2.4. Regras de ouro**
- 3. PARTE III: A CAPACITAÇÃO ADMINISTRATIVA E INSTITUCIONAL PARA A AVALIAÇÃO SOCIOECONÓMICA**
- 4. PARTE IV: A SELECÇÃO DE MÉTODOS E TÉCNICAS**

ANEXOS

Anexo A ao GUIA – As principais etapas da avaliação

Anexo B ao GUIA – Evolução das regras dos Fundos Estruturais

Bibliografia

INTRODUÇÃO

Índice (da *Introdução*)

- **A avaliação do desenvolvimento socioeconómico**
- **A quem se destina este GUIA?**
- **Porquê outro guia de avaliação?**
- **Actualização da Colecção MEANS**
- **Conteúdo**

A avaliação do desenvolvimento socioeconómico

Este GUIA discute a avaliação do desenvolvimento socioeconómico na Europa, com um enfoque particular nos Fundos Estruturais Europeus. Os fundos encontram-se organizados em programas, e a avaliação decorre nas fases ex-ante, intercalar e ex-post. Os programas constituem um meio para atingir objectivos de política mais alargados, e a avaliação dos programas contribui para a avaliação das políticas. Os programas compreendem diversas intervenções e projectos, e a avaliação ao nível de determinada medida/intervenção/projecto constitui parte integrante da avaliação dos programas. Diversos programas diferentes, e os seus respectivos elementos, contribuem para os objectivos temáticos, e a avaliação temática baseia-se na avaliação dos projectos e dos programas. Os princípios referidos neste GUIA aplicam-se, regra geral, à avaliação temática, dos projectos, das políticas e dos programas socioeconómicos. Deste modo, o GUIA será útil para todos os que tenham que planear, encomendar e gerir avaliações temáticas, de políticas e de projectos, bem como a avaliação de programas.

A quem se destina este GUIA?

Os utilizadores deste GUIA são provenientes das diferentes comunidades activas na avaliação dos programas de desenvolvimento socioeconómico. Entre eles, incluem-se:

- Decisores políticos, interessados em saber o que a avaliação lhes poderá proporcionar, incluindo as forças e as limitações da avaliação, bem como os recursos e as capacidades de que necessitam;
- Administradores públicos e agentes públicos que possam encomendar as avaliações, e que necessitem de uma descrição dos recursos disponíveis, incluindo as opções em termos de abordagem e métodos de trabalho;
- Gestores de programas que pretendam integrar os resultados da avaliação na forma como gerem e planeiam os seus programas;
- Parceiros de programas cada vez mais envolvidos como parte interessada nas avaliações, consultados sobre as agendas de avaliação e que se espera venham a utilizar os resultados da avaliação;
- Avaliadores, muitos dos quais possuem conhecimentos detalhados de áreas de avaliação específicas, mas que possam tirar partido da descrição de um conjunto mais alargado de métodos e abordagens de suporte ao trabalho colaborativo com outros membros de uma equipa de avaliação.

Embora o GUIA esteja mais orientado para utilizadores e o público em geral do que para especialistas, também são abordadas necessidades mais especializadas, com uma série de Manuais Técnicos de suporte ao conteúdo do GUIA.

Porquê outro guia de avaliação?

Actualmente, são muitos os materiais de base, textos e guias de avaliação disponíveis! À medida que as tarefas e a prática de avaliação foram se alargando, foi-se publicando uma biblioteca considerável de estudos sobre a avaliação. Embora esta literatura seja, sobretudo, proveniente da América do Norte, o desenvolvimento da avaliação na Europa - muitas vezes em resposta aos requisitos dos Fundos Estruturais - despoletou uma grande diversidade de novas publicações. A Comissão Europeia publicou Orientações de Metodologia detalhadas – sobre indicadores, a avaliação ex-ante, a avaliação macroeconómica, etc. –, específicas e em estreito paralelismo com a Regulamentação dos Fundos Estruturais. Existe, também, um Regulamento Financeiro que exige uma avaliação ex-ante e ex-post de todos os programas financiados pela UE, de cumprimento obrigatório. As autoridades públicas ao nível dos Estados-membros também publicam orientações destinadas a todos os que avaliam políticas e programas de desenvolvimento socioeconómico nacionais e europeus.

As obrigações de avaliação e as orientações publicadas por todos os que partilham a responsabilidade de programas de desenvolvimento socioeconómico estão sujeitas a alterações. É essencial que a avaliação seja colocada em estreita articulação com as circunstâncias em que decorre o desenvolvimento socioeconómico e com as principais opções políticas que é necessário informar. É fundamental esclarecer que este GUIA não substitui outros recursos; pelo contrário, sempre que necessário, baseia-se e faz referência a esses recursos. Este GUIA destina-se a um público mais vasto - e pretende apresentar práticas e abordagens de avaliação neste tipo de programas e áreas de intervenção política "independentes". Existe naturalmente uma grande diversidade de outros recursos altamente especializados, orientados para áreas da avaliação muito restritas, a um nível particularmente exigente. Este GUIA destina-se a preencher uma lacuna no mercado, no sentido de alargar a compreensão de boas práticas e de métodos sólidos, e apresentá-los de uma forma acessível.

Actualização da Colecção MEANS

Uma das principais fontes para a elaboração deste guia genérico foi evidentemente a Colecção MEANS - um vasto e imprescindível conjunto de manuais publicados pela Comissão Europeia em 1999. A Colecção MEANS tornou-se um texto-padrão para os avaliadores europeus, e alcançou uma reputação justificável, graças ao seu âmbito, ambição e alcance. De facto, uma grande parte desta colecção resistiu à passagem do tempo e foi integrado neste GUIA. Contudo, as condições também evoluíram desde 1999. Especificamente:

- Houve grandes transformações na área da prática de avaliação com a emergência de novas ferramentas de avaliação, um papel mais proeminente da componente teórica, novas abordagens participativas e qualitativas (particularmente relevantes para o desenvolvimento socioeconómico) e uma ênfase particular na institucionalização da avaliação.
- A política europeia evoluiu, sobretudo na sequência da Agenda de Lisboa, a partir da qual se discutiu o papel do capital social e humano, a importância da sociedade da informação e da economia baseada no conhecimento como meios de alcançar uma maior competitividade e ainda as prioridades do desenvolvimento sustentável e da igualdade de oportunidades.
- A adesão de dez novos Estados-membros em 2004 também colocou um conjunto de desafios à avaliação. Os Fundos Estruturais foram introduzidos em sistemas de administração pública com uma experiência de avaliação relativamente reduzida e, conseqüentemente, sem uma cultura de avaliação bem desenvolvida. Os autores deste GUIA tiveram em mente, durante a sua preparação, as necessidades de desenvolvimento de capacidades em muitos dos novos Estados-membros. Em termos práticos, procurou-se maximizar o que é passível de ser

alcançado de uma forma pragmática, com os recursos, capacidades, disposições institucionais e fontes de dados disponíveis.

No seu conjunto, estas alterações são significativas, e este GUIA foi elaborado para dar resposta às suas consequências. Também foi planeado e redigido de modo a poder prever alterações futuras. Não existem motivos para acreditar que o ritmo de alterações nos processos de avaliação e nas políticas socioeconómicas vá abrandar no futuro. Assim, e de modo a permitir a actualização imediata dos materiais, o GUIA, os Manuais Técnicos e o Glossário encontram-se disponíveis para acesso e pesquisa através da Internet. No futuro, espera-se que os materiais sejam ainda mais desenvolvidos, permitindo assim que os utilizadores possam estruturar e configurar o seu próprio GUIA 'pessoal', adaptado às suas necessidades, às funções de avaliação que possuem e aos programas e políticas específicos que estiverem a avaliar ou a gerir.

Conteúdo

O novo GUIA é publicado num único volume. É apoiado por uma série de Manuais Técnicos que proporcionam material mais especializado e aprofundado, ao qual é possível aceder e descarregar da internet.

O GUIA está dividido em quatro secções:

A secção sobre **a avaliação e o desenvolvimento socioeconómico** proporciona uma introdução à avaliação e às suas vantagens. Começa com uma descrição geral sobre a avaliação e a forma como poderá contribuir para a melhoria das políticas e dos programas e, em última instância, para o reforço do desenvolvimento socioeconómico. Segue-se uma introdução a alguns conceitos básicos da avaliação: a sua história, algumas das diferentes tradições nas quais se baseiam os avaliadores, e os diferentes objectivos da avaliação. Finalmente, discute-se as especificidades do desenvolvimento socioeconómico enquanto objecto de avaliação. Este processo inclui a selecção das características específicas da política de desenvolvimento socioeconómico e as suas implicações para a avaliação, bem como as principais teorias e conceitos nos quais assentam as políticas e os programas, e que os avaliadores necessitam de ter em consideração.

A secção dedicada à **concepção e implementação da avaliação** explora várias questões práticas relacionadas com a concepção e implementação das avaliações. Começa por analisar questões relacionadas com a concepção da avaliação, incluindo o seu planeamento, a definição de questões de avaliação e a selecção dos métodos mais adequados, bem como a encomenda e a implementação de um trabalho de avaliação. De seguida, aborda questões relacionadas com a gestão do processo de avaliação, após a fase da sua concepção, o que inclui a selecção de avaliadores e a função das Comissões de Acompanhamento, que deverão gerir as vias de comunicação entre todas as partes envolvidas, bem como assegurar um elevado nível de qualidade da avaliação.

A secção dedicada à **capacitação administrativa e institucional** aborda diversas questões relacionadas com o desenvolvimento das capacidades institucionais de avaliação e as estratégias que se poderão adoptar para o seu reforço. O conceito central desta parte do Guia estrutura-se em torno de um modelo "idealizado", que sugere quatro etapas no desenvolvimento das capacidades, incluindo uma discussão sobre as capacidades internas das administrações, bem como as capacidades externas das parcerias e das redes profissionais.

A secção a **selecção de métodos e técnicas** apresenta um vasto conjunto de métodos e técnicas de avaliação, realçando as suas forças e limitações, bem como a sua adequabilidade. Os métodos e técnicas são descritos de acordo com uma diversidade de

enquadramentos: diferentes tipos de programas socioeconómicos, diferentes etapas dos programas, diferentes etapas no processo de avaliação e diferentes objectivos de avaliação. Finalmente, são apresentados os diferentes tipos de dados (quantitativos e qualitativos), sistemas de indicadores e fontes de dados aos quais se recorrem durante um processo de avaliação.

Cada uma das secções do GUIA termina com uma série de "regras de ouro" que destacam, quer as boas práticas, quer alguns princípios empíricos, recomendados a todos os que gerem, requisitam, executam e utilizam avaliações. Contudo, regra geral, este GUIA procura não ser demasiado prescritivo. Isto deve-se em parte ao facto de, em muitos casos, não haver uma única forma correcta de avaliação, e cada uma das diferentes abordagens possuir as suas forças e as suas limitações em diferentes contextos. Também em termos pragmáticos, as condições ideais prévias de avaliação não existem na maioria dos casos – seja devido à falta de dados, problemas de calendarização ou disponibilidade de capacidades. Fazer o melhor que se pode enquanto se procura melhorar as capacidades de avaliação no futuro é um tema transversal a este GUIA.

Como forma de apoio ao GUIA, também foi preparada uma série de Manuais Técnicos, de particular interesse para especialistas e profissionais da área.

O Manual Técnico I intitula-se "**Temas e Áreas de Política**". Considera prioridades temáticas importantes, tais como o desenvolvimento sustentável, a sociedade da informação, a inclusão social e a igualdade de oportunidades, bem como o conjunto das áreas de política no âmbito das quais são implementadas as intervenções com vista à promoção do desenvolvimento socioeconómico. São apresentados e descritos os tipos de métodos de avaliação, dados e indicadores adequados a esses temas e áreas de política, e fornece-se alguns exemplos da sua aplicação.

O Manual Técnico II, "**Métodos e Técnicas de Avaliação**", inclui a análise pormenorizada de um vasto conjunto de ferramentas e técnicas, tanto quantitativas como qualitativas, que se consideram adequadas em diferentes fases do processo de avaliação.

O Manual Técnico III, "**Capacitação Administrativa e Institucional para a Avaliação**", inclui estudos de caso sobre o desenvolvimento das capacidades de avaliação na UE, nomeadamente na Itália, Holanda e Irlanda; e referências a outras experiências regionais, nacionais e internacionais, incluindo os países aderentes. Servem para ilustrar as orientações apresentadas no GUIA, e têm como objectivo incentivar o desenvolvimento das capacidades de avaliação.

Existe um Glossário que contém definições dos termos utilizados no GUIA e nos Manuais Técnicos.

Todos estes materiais encontram-se disponíveis neste website, e podem ser descarregados. Por outro lado, para otimizar a interacção com todos os que se encontram envolvidos em processos de avaliação, este site também inclui: um fórum, uma biblioteca onde se armazena documentos, e onde se poderá aceder aos mesmos, uma secção de novidades e uma listagem de eventos futuros. Pretende-se que o site seja um foco de discussão e debate, e que os materiais sejam permanentemente enriquecidos com referências às práticas mais actuais.

Índice (da Parte I)¹

1. PARTE I: A AVALIAÇÃO E O DESENVOLVIMENTO SOCIOECONÓMICO

- 1.1. As vantagens da avaliação
- 1.2. História e objectivos da avaliação
- 1.3. Os métodos e respectivas origens
- 1.4. A avaliação como meio de reforçar o desenvolvimento socioeconómico
- 1.5. Regras de ouro

1.1. As vantagens da avaliação

- Avaliações que fazem a diferença
- Melhorar os programas ao longo do tempo
- Conceber programas
- Seleccionar os instrumentos mais adequados
- Optimizar a gestão e a aplicação
- Identificar as realizações, os resultados e os impactos
- Identificar consequências não intencionais e efeitos perversos
- Níveis de avaliação: políticas, temas, programas e projectos

Avaliações que fazem a diferença

Investir tempo, dinheiro e esforços na avaliação tem que se justificar em termos da diferença que irá induzir no sucesso de políticas e programas - a avaliação não constitui um fim em si mesmo. No desenvolvimento socioeconómico, o que se pretende em termos de intervenção política é optimizar as condições económicas e sociais de pessoas, regiões ou sectores. Evidentemente, cada programa socioeconómico possui a sua própria filosofia específica, em que, por exemplo, alguns darão destaque à reabilitação dos centros urbanos, outros à modernização de sectores industriais obsoletos ou em declínio, outros ainda à integração de grupos desfavorecidos, e outros à diversificação em zonas rurais. Todas estas prioridades, entre muitas outras, encontram-se amplamente representadas nos programas dos Fundos Estruturais europeus. No entanto, a justificação da avaliação, em todos estes casos, é a mesma: poder-se-á aplicar métodos e processos de avaliação de uma forma que venham efectivamente a optimizar a qualidade de vida, a prosperidade e as oportunidades à disposição dos cidadãos? De modo a que possa efectuar alterações efectivas a esse nível, é necessário que a avaliação coloque questões e forneça respostas relevantes para as partes interessadas nos programas, quer sejam administradores, decisores políticos ou beneficiários.

A contribuição da avaliação de programas é potencialmente superior em áreas de política inovadoras, onde o êxito não se pode tomar por garantido e onde a implementação nem sempre é linear, para além de exigir uma gestão e um planeamento sofisticado. Quando aplicada de forma adequada, a avaliação pode contribuir para a gestão de algumas das inevitáveis incertezas inerentes a situações complexas. O desenvolvimento socioeconómico, como se discutirá mais à frente neste GUIA, é certamente complexo, e debate-se frequentemente com diversas incertezas: não é uma ciência exacta. Seleccionar metas e medidas, conceber programas, implementar e manter uma dinâmica de desenvolvimento - todas elas requerem a análise, previsão, estabelecimento de

¹ As partes e os capítulos não estão numerados na versão on-line.

sistemas de *feedback* e mobilização de diferentes instituições, agências e grupos demográficos.

A avaliação tornou-se uma componente essencial de uma infinidade de iniciativas de desenvolvimento socioeconómico, ao constatar que a sua prática e os conhecimentos adquiridos têm contribuído efectivamente para esses processos.

Existem duas implicações importantes, se a avaliação for justificada nesses termos. Em primeiro lugar, se a avaliação for útil e aplicável, é necessário que seja vista como uma parte integrante da gestão e da tomada de decisão e, efectivamente, de todo o processo de responsabilidade democrática. Assim, um sistema de avaliação que funcione correctamente deverá ser integrado no ciclo de programação/político, sendo por essa razão que este GUIA presta tanta atenção à concepção dos sistemas de avaliação e ao desenvolvimento das capacidades de avaliação nas agências públicas e nas redes de conhecimento e profissionais.

Em segundo lugar, os avaliadores, bem como os requerentes e utilizadores dos resultados da avaliação terão obrigatoriamente de atingir um equilíbrio entre os melhores meios disponíveis e as exigências de um determinado nível de pragmatismo. No mundo real do desenvolvimento socioeconómico, muito raramente existe o tempo ou os recursos - ou mesmo os dados - necessários para implementar uma avaliação exaustiva do estado da arte. É por isso que este GUIA coloca particular ênfase nos tipos de escolha estratégica que é necessário fazer relativamente à avaliação, como por exemplo: quando é que se justificam investimentos mais avultados na avaliação? Em que circunstâncias é que são necessários métodos sofisticados? De que forma é que a avaliação pode preencher algumas lacunas no conhecimento que, num mundo ideal, seriam preenchidas inclusivamente antes de se planear uma intervenção?

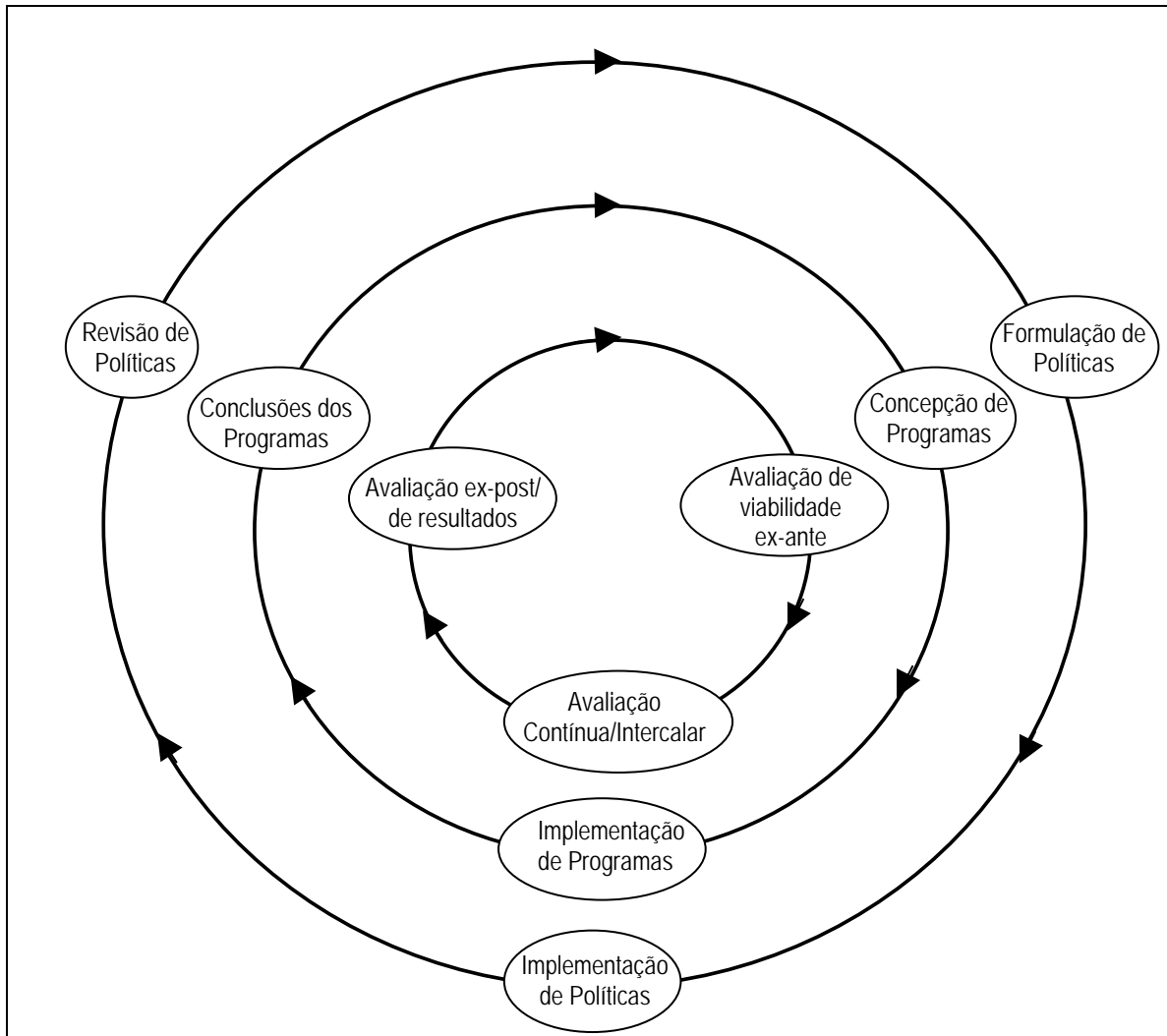
Melhorar os programas ao longo do tempo

Um importante princípio estruturante, transversal a todo o GUIA, tem a ver com o horizonte temporal e o prazo das políticas. Normalmente, fala-se de um ciclo de intervenção política que se inicia quando se formulam as políticas (e os programas com elas relacionados), e que continua com o planeamento e atribuição de recursos, a concepção de programas, a sua implementação e a concretização das suas realizações e resultados. A linguagem da avaliação segue em muitos casos este ciclo, como se pode verificar através dos termos de avaliação como *ex-ante*, *intercalar* e *ex-post*, utilizados normalmente nos Fundos Estruturais europeus. Estes termos encontram-se descritos no Anexo A. Utiliza-se uma lógica semelhante na distinção feita geralmente entre realizações, efeitos, resultados e impactos.

Antes de se voltar às características específicas do desenvolvimento socioeconómico, será apresentada a contribuição da avaliação possível em cada uma das fases do ciclo político.

Como se pode ver no diagrama apresentado na Caixa *Ciclos das Políticas, dos Programas e da Avaliação*, existem três ciclos temporais diferentes, importantes para todos os que estão envolvidos na avaliação. Em primeiro lugar, o ciclo da avaliação que decorre em diferentes momentos e em fases distintas dentro de um segundo ciclo, e depois o ciclo dos programas que gera, ele próprio, uma procura por estes diferentes momentos de avaliação. Existe ainda um terceiro ciclo, o ciclo das políticas, que dá forma e influencia os programas, bem como, inevitavelmente, os requisitos da avaliação. Normalmente, o ciclo das políticas é mais prolongado do que o ciclo dos programas.

Caixa: Ciclos das Políticas, dos Programas e da Avaliação



A Caixa Ciclos das Políticas, dos Programas e da Avaliação é apenas indicativa; obviamente, existem mais fases em cada ciclo, que podem descrever-se de diferentes formas. Mas o diagrama mostra alguns dos principais problemas relacionados com os respectivos horizontes temporais comuns na avaliação. O círculo interno começa com a avaliação ex-ante, que precede o levantamento das necessidades iniciais e a análise da viabilidade dos programas planeados; passa pela avaliação contínua ou intercalar, que documenta o progresso e a implementação; e, finalmente, termina na avaliação ex-post, com um enfoque nos resultados. Contudo, as avaliações ex-ante devem contribuir para a concepção dos programas e formulação das políticas, do mesmo modo que as avaliações intercalares deverão contribuir para a implementação dos programas e das políticas, em termos de aplicação destes e de outros programas semelhantes. No final do ciclo de avaliação, as avaliações ex-post devem contribuir para a revisão das políticas. Colocar estes dois ciclos em paralelo seria o ideal, mas isto nem sempre acontece. As avaliações ex-ante podem decorrer demasiado tarde para contribuir para a concepção dos programas, e menos ainda para a formulação de políticas. Os resultados das avaliações ex-post podem ser obtidos demasiado tardiamente para contribuir para a revisão das políticas. As alterações de política e programação também podem decorrer sempre que estiver a ser executada uma avaliação, comum em programas nacionais e europeus de desenvolvimento socioeconómico. Isto pode, por exemplo, conduzir à

alteração dos objectivos ou prioridades, após a configuração dos sistemas para medir os resultados e, inclusivamente, ao encerramento de determinados projectos ou intervenções objecto de avaliação. Uma das vantagens do envolvimento de decisores políticos e agentes de planeamento na concepção da avaliação consiste em operar uma aproximação entre todas estas actividades interligadas.

Em geral, este GUIA está particularmente interessado na avaliação de programas – mais do que na avaliação de políticas – embora não seja em muitos casos tarefa fácil separá-los. Esta questão torna-se particularmente relevante numa era que se designa frequentemente a da política baseada em provas. Existe um forte imperativo para utilizar recursos de forma racional e para aprender com as experiências anteriores: para que se possa basear as políticas em provas. Esta constitui efectivamente uma característica essencial da forma como se compreende a avaliação ex-ante no âmbito dos Fundos Estruturais europeus, em que a avaliação pode constituir uma ferramenta poderosa para aprender com a experiência de programas anteriores, de modo a otimizar as políticas no futuro.

A avaliação pode proporcionar uma série de contributos específicos ao longo do ciclo dos programas:

Conceber programas

Uma das competências centrais da avaliação consiste na recolha de informações dos diferentes destinatários ou partes interessadas, tratando-se de um aspecto particularmente importante na fase de concepção dos programas. Garantir a relevância dos programas em termos das necessidades dos respectivos utilizadores é essencial nesta fase, para a qual a avaliação pode contribuir de uma forma significativa. Este contributo da relevância não se limita à concepção de programas num determinado momento no tempo. Em muitos contextos de desenvolvimento socioeconómico, os programas mantêm-se activos através de um processo contínuo de *feedback* de utilizadores (potenciais e reais) e de outras partes interessadas, durante o qual, aliás, surge frequentemente um momento explícito de reprogramação.

Seleccionar os instrumentos mais adequados

Um sistema de avaliação bem estruturado e muito particularmente a avaliação ex-ante contribuirão para a selecção de intervenções ou instrumentos específicos, no contexto geral de um programa. A avaliação pode contribuir de diversas formas para a selecção dos instrumentos, podendo assumir a forma de uma análise económica que avalia os benefícios e os custos mais prováveis de um conjunto de instrumentos alternativos, ou talvez de grandes projectos. Também pode implicar uma avaliação da elegibilidade que procura alinhar intervenções específicas com determinados critérios, de modo a verificar a sua relevância para um determinado programa na sua globalidade ou em termos regulamentares. Em alternativa, poderá existir uma avaliação da clareza e credibilidade da intervenção proposta para analisar a sua probabilidade de sucesso.

Otimizar a gestão e a aplicação

Um processo de avaliação plenamente integrado poderá representar uma contribuição essencial para a gestão e para a implementação dos programas. Analisando os dados de monitorização e investigando as causas subjacentes às dificuldades encontradas, a avaliação permite oferecer *feedback* às entidades gestoras dos programas e apoiar a "correção intercalar". Mesmo numa primeira fase, existem normalmente resultados iniciais, sobretudo se existir uma cadeia de implementação e um modelo lógico bem

estruturados. Assim, a avaliação da implementação permite destacar alguns resultados desde o início. Porém, muitas das questões emergentes nestas fases iniciais de implementação estão relacionadas com os processos: como é que os diferentes parceiros interagem, como são tomadas as decisões e como se elaboram os planos, e de que modo se procura sedimentar novas parcerias e disposições organizacionais. A avaliação destes processos - incluindo descrições lineares - podem ser úteis para todos os parceiros, bem como para os gestores e os principais financiadores.

Identificar as realizações, os resultados e os impactos

Os programas de desenvolvimento socioeconómico têm que demonstrar resultados desde muito cedo. Nas fases mais preliminares, é provável que estes assumam a forma de realizações (*outputs*), como por exemplo, o número de empresas que recebem financiamento para actualizar os seus equipamentos ou os números relativos a pessoas desempregadas a receber formação. No entanto, os decisores políticos procuram rapidamente resultados mais substanciais: tornar as empresas mais competitivas ou criar emprego para os desempregados. Espera-se que estes resultados (*outcomes*) possuam um maior impacto relacionado com os objectivos das políticas e dos programas, pelo menos a longo prazo. Isto permite à avaliação colocar questões como: o crescimento das empresas regionais foi sustentado? Ou as expectativas de emprego das pessoas em situação de desemprego prolongado melhoraram de forma sustentável?

Para muitos decisores políticos, identificar, descrever e quantificar estas realizações, resultados e impactos é uma das grandes vantagens da avaliação. No entanto, para que este processo possa ser realmente proveitoso, os decisores políticos necessitam de assegurar que existem objectivos claros e uma relação ponderada entre as intervenções e os objectivos dos programas. Os gestores de programas, por seu turno – se necessário, em coordenação com os avaliadores –, têm de assegurar a existência de sistemas de monitorização e de indicadores, bem como ligações claras entre os indicadores seleccionados e as metas e objectivos subjacentes.

Identificar consequências não intencionais e efeitos perversos

Mesmo quando os programas e os instrumentos cumprem os objectivos propostos, ocorrerão frequentemente consequências ou efeitos inesperados, que podem ser positivos ou negativos. Por exemplo, o apoio a empreendedores rurais pode ter vantagens tecnológicas inesperadas para empreendedores urbanos de uma localidade próxima, que operem no mesmo sector ou mercado. Por vezes, estas consequências inesperadas podem ter um efeito negativo. Por exemplo, um instrumento concebido para aumentar as expectativas de emprego de determinado grupo no mercado de trabalho pode ter consequências negativas para outro grupo. *In extremis*, as intervenções podem, inclusivamente, ter um efeito perverso, conduzindo a intervenção numa direcção diametralmente oposta à pretendida. Por exemplo, uma iniciativa destinada a promover o turismo pode, caso a base da procura turística na região não seja correctamente compreendida, minar a indústria turística existente, sem criar um mercado adicional.

Para captar os resultados das intervenções socioeconómicas, é essencial incluir as consequências inesperadas e os efeitos perversos. Esta é também uma das formas em que a avaliação pode contribuir para aprender com a experiência, sobretudo ao fornecer indicações sobre a execução das intervenções, sobre como conceber melhor os programas e como evitar desaproveitamentos nas intervenções e efeitos perversos.

Os níveis de avaliação: políticas, temas, programas e projectos

A questão das interligações entre políticas, programas e intervenções ou projectos específicos é essencial na avaliação. Muitos programas bons nem sempre representam uma boa política, documentos de programação sólidos não representam necessariamente bons projectos, e bons projectos nem sempre asseguram o sucesso de um programa. Contudo, a avaliação do programa representa necessariamente um recurso ou contributo para a avaliação de políticas, tal como a avaliação de projectos representa um recurso ou contributo para a avaliação de programas.

As avaliações temáticas e os critérios decorrentes das políticas, quando aplicados aos materiais dos programas, constituem formas habituais de introduzir uma dimensão política na avaliação.

Existe, agora, uma tendência da avaliação no sentido ascendente, bem como para prestar cada vez mais atenção ao nível político, reflectindo uma intenção por parte dos decisores políticos em assumir os resultados da avaliação. Simultaneamente, coloca desafios aos avaliadores, que necessitam de ver os resultados do seu trabalho num contexto mais alargado. Considerar o nível político também permite reforçar a avaliação do programa, identificando por exemplo os critérios orientados para os resultados, para o sucesso do programa.

Na sua maioria, a avaliação de projectos, pelo menos quando os projectos têm uma escala relativamente pequena, volta aos promotores dos projectos e a outros intermediários. A principal excepção são os projectos de larga escala (por exemplo, projectos de infra-estruturas), que possuem muitas das características dos programas em termos de complexidade, bem como em termos de dimensão.

Um requisito recorrente para promotores e gestores de projecto é que se execute algum tipo de auto-avaliação. Embora este tipo de avaliação possa não ter a independência considerada importante para assegurar a credibilidade externa, pode mesmo assim constituir um contributo importante no contexto dos programas. Por exemplo, se um quadro bem estruturado for concebido para a auto-avaliação de projectos, poderão existir algumas garantias de que os resultados serão sistemáticos e merecem uma análise detalhada a nível do programa. Além disso, os requisitos de auto-avaliação podem incentivar uma cultura de feedback e de aprendizagem inter- e intra-projectos, com vantagens para o programa como um todo.

Todos os que desempenhem tarefas de planeamento e execução de trabalho de avaliação devem clarificar as ligações entre a avaliação de políticas, de programas, de projectos e a avaliação temática. Os princípios descritos neste GUIA são geralmente aplicáveis a cada tipo de avaliação.

1.2. HISTÓRIA E OBJECTIVOS DA AVALIAÇÃO

- Uma breve história da avaliação
- Diferentes fontes e tradições
- Os principais objectivos da avaliação
- A aprendizagem como principal objectivo de avaliação

Uma breve história da avaliação

A avaliação surgiu como uma área distinta do exercício profissional na América do Norte, no período pós-guerra. As três correntes mais importantes nessa fase inicial eram: a avaliação da inovação na educação (por exemplo, a eficácia de novos currícula nas

escolas); a ligação da avaliação à alocação de recursos (por exemplo, através de um sistema de Planeamento, Programação e Orçamentação); e a avaliação de programas de combate à pobreza (por exemplo, as experiências da Great Society nos anos 60). Estas diferentes abordagens definiram algumas das principais convenções de avaliação que continuam a ser utilizadas hoje em dia, incluindo estudos quantitativos e experimentais que recorrem a grupos de controlo como base de diversas experiências de testes educativos, métodos de avaliação económica e de rentabilidade, e métodos participativos e qualitativos envolvendo os beneficiários dos programas no processo de avaliação.

Subjacentes a estas diferentes tradições encontram-se quatro grupos principais cujos interesses, por vezes, concorrem na definição de prioridades de avaliação. Estes incluem:

- decisores políticos, por exemplo políticos e dirigentes eleitos;
- interesses profissionais e especializados, tais como professores, na área da educação, e cientistas, na área da investigação;
- directores e administradores, tais como funcionários públicos e dirigentes de agências públicas locais;
- cidadãos e todos os que são influenciados pela acção pública, por exemplo, os potenciais beneficiários das intervenções planeadas.

Cada um destes grupos chega a determinadas conclusões relativamente à forma como a avaliação poderá ajudá-los. Por exemplo, os decisores políticos têm tendência para ver a avaliação como uma ferramenta que garante a responsabilidade e a justificação de decisões políticas; os cidadãos têm uma maior probabilidade de ver a avaliação como uma ferramenta de responsabilidade democrática e como uma oportunidade de adequar as intervenções públicas às suas necessidades; os directores e os administradores estão, em muitos casos, preocupados com a implementação de políticas e programas, e com a forma como estes são geridos e se encontram organizados; os profissionais, por sua vez, vêem frequentemente a avaliação como uma oportunidade de aumentar a qualidade do seu trabalho ou mesmo a autonomia do seu próprio grupo profissional.

Isto não significa que a avaliação, num sentido mais lato, enquanto aplicação da investigação social e económica sistemática, fosse totalmente inexistente na Europa ou noutras partes do mundo. Contudo, era provavelmente mais forte no norte da Europa e em zonas da Europa com ligações mais fortes aos Estados Unidos e ao Canadá. A partir dos anos 70, a avaliação começou a enraizar-se nos diferentes países europeus, mas geralmente com tradições e ênfases distintos. Na Escandinávia, por exemplo, onde existe um forte empenho na governação democrática, a avaliação seguiu no sentido dessa tradição. Em França, a avaliação reflectiu, até há muito pouco tempo, as características de um Estado francês com uma abordagem formal estruturada ao nível do governo central e uma prática mais diversa e dinâmica aos níveis regional e local. Contudo, a avaliação nunca foi estática em nenhum destes países. Por exemplo, a experiência de avaliação francesa evoluiu consideravelmente com os requisitos da reforma orçamental que surgiram depois de 2000. Em muitos países, o enfoque e a escala da actividade de avaliação reflectiram as políticas de mudança dos diferentes governos. Por exemplo, no Reino Unido a avaliação aumentou consideravelmente com a mudança de governo em 1997.

Os Fundos Estruturais europeus constituíram um dos principais indutores de disseminação da prática de avaliação em toda a UE. Em cada uma das etapas do ciclo de programação (ex-ante, intercalar, ex-post), existem objectivos e responsabilidades claros. Reconhece-se, muitas vezes, que a introdução da avaliação em diversos países do sul da Europa foi resultante dos requisitos das regulamentações dos Fundos Estruturais. Apesar de um início reservado em 1988, existe agora uma abordagem à avaliação dos Fundos Estruturais sofisticada.

Esta abordagem inclui:

- uma obrigação legal das entidades financiadoras e dos gestores de programas para com a avaliação;
- responsabilidade partilhada entre os diferentes níveis hierárquicos do governo no processo de avaliação global;
- um processo de avaliação inter-relacionado, que decorre em diversas etapas (ex-ante, intercalar, ex-post);
- envolvimento de diversos parceiros nos programas e na sua avaliação;
- ligações claras entre a avaliação, por um lado, e a programação e atribuição de recursos, por outro.

Ao longo dos últimos anos, registou-se uma evolução nas regulamentações dos Fundos Estruturais relativamente à avaliação (ver a Caixa *A Evolução das Regulamentações dos Fundos Estruturais*).

Caixa: A Evolução das Regulamentações dos Fundos Estruturais

1994 - 1999	2000 - 2006
AVALIAÇÃO EX-ANTE	
Qualquer avaliação ex-ante deve ser executada em parceria pela Comissão e pelo Estado-membro, e deverá incluir o impacto ambiental. A Comissão avalia os planos, tendo em consideração a avaliação ex-ante.	O Estado-membro é o primeiro responsável pela avaliação ex-ante. Ao definir o objectivo da avaliação ex-ante, deverá prestar-se especial atenção ao impacto sobre o ambiente, sobre o mercado de trabalho e sobre a igualdade entre homens e mulheres. A avaliação ex-ante encontra-se integrada nos planos.
AVALIAÇÃO INTERCALAR	
Não estão estipulados quaisquer princípios que exijam a recolha e comunicação de informações. Não existem quaisquer provisões relativamente à definição do papel das Comissões de Acompanhamento no processo de avaliação (na prática, o papel principal das Comissões de Acompanhamento tem consistido em nomear os avaliadores).	O Estado-membro é responsável pela avaliação intercalar, em parceria com a Comissão; a Comissão avalia o grau de relevância da avaliação. Define-se o objectivo da avaliação intercalar. A avaliação intercalar é organizada pela entidade que coordena o programa correspondente, e é executada por um avaliador independente. Deve estar concluída até 31 de Dezembro de 2003. Para facilitar as diversas avaliações, a entidade que a coordena deve criar um sistema de recolha de dados financeiros e estatísticos, necessário para as avaliações intercalares e ex-post, e deve proporcionar as informações necessárias para executar a avaliação ex-post. A Comissão de Acompanhamento analisa os resultados da avaliação intercalar e pode propor alterações ao programa, em função dos mesmos. No final de 2005, a avaliação intercalar é actualizada através da preparação do terreno para as operações que se seguem. (A actualização da avaliação intercalar é, também, designada de avaliação final).
AVALIAÇÃO EX-POST	

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

Deverá ser realizada uma avaliação ex-post, numa parceria entre a Comissão e o Estado-membro, com o intuito de avaliar o impacto das medidas adoptadas face aos objectivos propostos.	A Comissão é o principal responsável pela avaliação ex-post, em parceria com o Estado-membro. Define-se o objectivo da avaliação ex-post. A avaliação ex-post é realizada por um avaliador independente, num prazo de três anos após o final do período de programação.
RESERVA DE PERFORMANCE	
	Até 31 de Março de 2004, uma percentagem de 4% da alocação indicativa de cada Estado-membro, retida no início do período, é alocada como “reserva de performance” dos documentos de programação, cujo desempenho a Comissão considere ser bem-sucedido, com base nas propostas do Estado-membro. O desempenho é avaliado com base em indicadores, eficácia, gestão e implementação financeira, que são definidos pelo Estado-membro, após consulta com a Comissão.

Algumas das principais transições foram:

- de obrigações de avaliação impostas externamente, para uma procura de avaliação induzida internamente, decorrente dos próprios gestores de programas e decisores políticos;
- de um processo de avaliação ligado aos programas no final de um ciclo de programação, a um processo de avaliação totalmente integrado nos programas, desde o início;
- das expectativas de que os resultados da avaliação necessitam de ser disseminados em larga escala por motivos de responsabilização, para uma preocupação com a utilização sistemática da avaliação no decorrer da implementação de um programa;
- de uma perspectiva de acordo com a qual a gestão da avaliação constituía essencialmente uma questão de administração de um contrato, para um interesse na forma como a avaliação poderá contribuir para a gestão de conhecimento.

Estas alterações foram acompanhadas por mudanças de responsabilidade entre os diferentes actores ao nível europeu, nacional e regional, e pelo alargamento ao coordenador da parceria. A gama de potenciais interessados na avaliação alargou-se, assim, no sentido de incluir, por exemplo, autoridades locais, parceiros sociais e grupos da sociedade civil.

A reforma dos Regulamentos dos Fundos Estruturais (ver Anexo B) para a terceira geração de programas (2000-2006), ao mesmo tempo que devolve muitas das obrigações de avaliação às autoridades responsáveis dos Estados-membros, exige que estas avaliações sejam utilizadas, tanto numa fase ex-ante, como novamente numa fase intercalar. A revisão das avaliações intercalares decorre até às avaliações finais. Esta combinação de responsabilidades delegadas com escrutínio externo, por estruturas de governação mais elevadas, é comum em avaliações nacionais do desenvolvimento socioeconómico.

Ao longo dos últimos anos, surgiu igualmente uma forte tendência para a reforma da administração pública e a introdução de conceitos de gestão de eficiência em diversos países europeus, bem como na própria Comissão Europeia (ver "*Sound and Efficient*

Management", conhecido como SEM 2000, e o relatório "*Spending more Wisely: Implementation of the Commissions Evaluation Policy*"). Recentemente, este processo avançou, estabelecendo a relação entre as decisões financeiras e orçamentais e a eficiência, que tem por conseguinte de ser medida e descrita. A gestão da eficiência na Comissão Europeia, bem como nos Estados-membros, reflecte-se agora num elemento adicional das obrigações de avaliação nos actuais regulamentos que se encontrem relacionados com a Reserva de Eficiência. Os indicadores de gestão de eficácia e de implementação financeira têm que ser reunidos, de modo a justificar a disponibilidade de fundos retidos no início do período de programação.

Diferentes fontes e tradições

A avaliação possui diversas raízes: não constitui uma prática unificada, ou decorrente de um único conjunto de tradições. Isto resulta, em parte, da evolução histórica da avaliação, tanto na Europa como na América do Norte. Como já se referiu, destacam-se normalmente três fontes importantes da filosofia da avaliação: as iniciativas nos anos 60 da Great Society nos Estados Unidos; a inovação educativa e sobretudo a inovação dos currículos nas escolas; e o controlo orçamental e os sistemas de eficiência como o PPBS (*Planning, Programming and Budgetary Systems*). Na realidade, estas constituem apenas três fontes específicas, e poder-se-ia acrescentar a gestão por objectivos, a investigação participativa no desenvolvimento rural e comunitário, e a gestão com base nos resultados, entre muitas outras.

Uma forma de distinguir algumas destas diferentes origens consiste em estabelecer algum distanciamento em relação a exemplos específicos e analisar as ideias ou teorias centrais subjacentes a estas diferentes tradições de avaliação.

Pode fazer-se a distinção entre quatro conjuntos principais de ideias:

- *Metodologia e investigação científica*. Muitos dos métodos e ideias de base utilizados na avaliação são partilhados com a comunidade de investigação mais alargada, sobretudo nas ciências sociais e económicas. Considerando a lógica que integra o teste de hipóteses, a observação, a recolha de dados e a análise de dados, procura-se explicações para os fenómenos que se observam. Em programas socioeconómicos complexos, as explicações raramente são lineares. Muito do trabalho dos avaliadores consiste em tentativas de aplicar os recursos conhecidos aos resultados observados e vice-versa.
- *Teoria económica e escolhas públicas*. A reflexão económica está presente na avaliação a diversos níveis. Incluem-se neste caso noções de eficiência e atribuição de recursos para enfrentar a escassez, incentivos e comportamentos institucionais (sobretudo microeconómicos), e estudos macroeconómicos que procuram identificar efeitos agregados (por exemplo, em termos de PIB ou competitividade) das intervenções políticas.
- *Teoria da organização e gestão*. Esta começou por estar incluída mais significativamente na avaliação nos últimos anos, à medida que o enfoque passou de forma crescente para a implementação e aplicação de programas e políticas. Esta linha de reflexão destaca aspectos de concepção organizacional, coordenação inter-organizacional (por exemplo, através de parcerias e consórcios), bem como aspectos de motivação, propriedade e participação.
- *Ciências políticas e da administração*. À medida que os programas públicos e os seus gestores procuram resolver questões relacionadas com o processo político e com a reforma do sector público, baseiam-se cada vez mais em ideias sobre governança, responsabilidade e cidadania. Muitas das ideias centrais da reforma do sector público e da nova gestão pública, tais como a transparência e a responsabilidade, foram influenciadas por estas perspectivas. Além disso, nas

perspectivas políticas contemporâneas destaca-se a importância da construção de consensos como forma de reforçar a legitimidade da acção política.

Consequentemente, os avaliadores são igualmente diversos. Poderão ser economistas preocupados com a eficiência e com os custos, consultores de gestão interessados no funcionamento estável da sua organização, analistas políticos empenhados na reforma e transparência do sector público, ou cientistas (de várias áreas disciplinares) preocupados em estabelecer a verdade, gerar novos conhecimentos e confirmar/rejeitar hipóteses.

Um dos principais problemas enfrentados actualmente pelos gestores ou requerentes de avaliações consiste em reunir uma equipa adequada ou um conjunto de competências que poderá representar uma síntese de todas estas tradições (este assunto é aprofundado na secção Concepção e Implementação do GUIA, em que se discute o perfil da equipa de avaliação e a selecção dos avaliadores mais adequados).

A um nível sistémico (por exemplo, a nível nacional ou europeu em geral), uma das principais funções da construção de capacidades de avaliação é construir pontes entre estas diferentes partes das comunidades de avaliação profissionais. Conferências, redes e organizações profissionais que reúnem os avaliadores constituem uma forma de aumentar a familiaridade entre todos aqueles que provêm de diferentes tradições, bem como uma forma de transferir e partilhar "know-how", conhecimentos e experiências (este aspecto também é desenvolvido em profundidade na secção Concepção e Implementação do GUIA).

Apesar das diferenças existentes nas origens e tradições da avaliação, é possível distinguir alguns dos principais tipos de avaliação. Estes têm tendência para ocorrer em torno de dois eixos principais. O primeiro eixo relaciona-se com os objectivos da avaliação, e o segundo está relacionado com as metodologias de avaliação.

Os principais objectivos da avaliação

As avaliações contribuem sempre para um objectivo mais vasto, que consiste em contribuir em particular para uma área de política pública e para os seus programas. Os objectivos de avaliação reconhecidos com maior frequência são:

- Planeamento/eficiência - garantir a existência de uma justificação para uma política/programa e que os recursos se desenvolvem eficientemente.
- Responsabilidade - demonstrar em que medida é que determinado programa alcançou os seus objectivos, e de que forma utilizou os seus recursos.
- Implementação - aumentar a eficiência dos programas e a eficácia da sua implementação e gestão.
- Capacitação institucional - melhorar e desenvolver as capacidades entre os participantes nos programas e as suas redes e instituições.

Estes diversos objectivos de avaliação são do interesse dos diferentes intervenientes, e também têm tendência a serem associados a diferentes tipos de questões de avaliação. Por exemplo:

- Se o objectivo for planeamento/eficiência, irá satisfazer sobretudo as necessidades dos decisores políticos e dos decisores do planeamento, bem como dos cidadãos. São estas as partes interessadas que estarão preocupadas com a forma como os recursos públicos se distribuem entre objectivos concorrentes, e como se aplicam depois de distribuídos. Estas partes interessadas colocam questões como: será esta a melhor utilização dos fundos públicos? Existem utilizações alternativas dos recursos que proporcionam mais vantagens? Existe uma correspondência entre os custos implicados e as vantagens daí decorrentes?

- Se o objectivo da avaliação for a responsabilidade, irá sobretudo satisfazer as necessidades dos decisores políticos, dos financiadores dos programas e dos parlamentos. São estas partes interessadas que, depois de aprovar um programa ou uma política, pretendem saber o que aconteceu aos recursos atribuídos. Este tipo de avaliação coloca questões como: qual o êxito deste programa? Cumpriu os seus objectivos? Os recursos foram aplicados de uma forma eficaz e eficiente, e com que impacto?
- Se o objectivo da avaliação for a implementação, satisfará sobretudo as necessidades dos gestores do programa e dos principais parceiros do programa. São estas partes interessadas que pretendem otimizar a gestão e a implementação, em relação às quais são os principais responsáveis. Este tipo de avaliação coloca questões como: os mecanismos de gestão funcionam eficientemente? O nível de envolvimento por parte dos parceiros é o desejável? Os programas encontram-se devidamente orientados em termos de elegibilidade? A planificação temporal está a ser cumprida?
- Se o objectivo da avaliação for a produção de conhecimento, irá satisfazer sobretudo as necessidades dos decisores políticos e de planeamento - incluindo aqueles que têm como função planear novos programas. São estas mesmas partes interessadas que pretendem saber se os princípios do programa estão a ser confirmados e quais as lições a tirar para o futuro. Este tipo de avaliação levanta questões como: o que é que já se aprendeu acerca do que melhor funciona? Os mecanismos de intervenção e de mudança compreendem-se melhor? É necessário questionar a lógica do programa e os seus princípios? Será esta uma forma eficiente de alcançar os objectivos propostos, ou existem alternativas? Quais são as provas de sustentabilidade das intervenções?
- Se o objectivo da avaliação for a capacitação institucional, esta deverá sobretudo satisfazer as necessidades dos parceiros do programa e de outras partes interessadas. Estes pretenderão saber como poderão ser mais eficazes, de que forma poderão aumentar as suas próprias capacidades e como é que os beneficiários podem tirar o máximo partido dos pressupostos do programa. Este tipo de avaliação coloca questões como: os beneficiários (e inclusivamente as comunidades locais) estão suficientemente envolvidos na delineação do programa e das suas medidas? O que é que é possível fazer para aumentar a participação e desenvolver o consenso? Os mecanismos do programa apoiam e estão abertos a abordagens ascendentes?

A aprendizagem como principal objectivo de avaliação

Sugere-se, muitas vezes, que a avaliação pode ser vista como tendo um objectivo abrangente, ao qual poderão adequar-se todos os outros objectivos referidos acima. Esta finalidade abrangente refere-se à aprendizagem e, nesta perspectiva, a avaliação possui como finalidade aprender, através de um processo sistemático, a melhorar a concepção, implementação e execução de políticas e programas públicos.

Esta ênfase na aprendizagem realça uma das principais características da avaliação, que é consistente com as necessidades dos programas de desenvolvimento socioeconómico. Como já se constatou, nestes programas o conhecimento é imperfeito, e existe uma necessidade constante de aprender sobre diferentes contextos e sobre a melhor forma de combinar medidas diferentes mais eficientemente.

Um enfoque na aprendizagem também realça um aspecto importante de uma cultura de avaliação - um elemento essencial da capacidade de avaliação discutido mais detalhadamente de seguida. É geralmente aceite que, para que a avaliação seja devidamente integrada na elaboração de políticas, é necessário haver uma cultura que

suporta a aprendizagem e que seja capaz de apreender lições positivas para o futuro, tanto a partir de problemas ou mesmo insucessos, como de casos de sucesso.

1.3. OS MÉTODOS E RESPECTIVAS ORIGENS

- Os eixos metodológicos
- Três tradições filosóficas
- Quatro tipos de teoria relevantes para a avaliação

A secção do GUIA intitulado A Seleção de Métodos e Técnicas incide nos métodos e ferramentas (ou técnicas) de avaliação. Aqui centramo-nos, sobretudo, nas raízes ou alicerces subjacentes a estes métodos e ferramentas. Descrevem-se cinco grandes perspectivas metodológicas, e, de seguida, a forma como estas se encontram ligadas a escolas filosóficas mais gerais, discutidas na maioria das ciências económicas e sociais aplicadas. Muitas destas discussões filosóficas destacam a centralidade da teoria na avaliação, que se tem tornado cada vez mais importante ao longo dos últimos anos. Por isso, procede-se a uma revisão breve dos motivos pelos quais a teoria é importante para a avaliação do desenvolvimento socioeconómico e das diferentes formas que este pode assumir. Estes alicerces metodológicos e filosóficos são a base de algumas das principais categorias ou famílias de métodos que se discutirá na secção do GUIA A Seleção de Métodos e Técnicas e que são aqui apresentados.

Os eixos metodológicos

Em termos de metodologia, observando as diferentes abordagens de avaliação descritas acima, é possível distinguir cinco posições metodológicas:

- A posição de atribuição de recursos, que diz respeito à utilização eficiente dos recursos, tanto prospectivamente, em termos de planeamento, como retrospectivamente, em termos da forma como os recursos são utilizados.
- Os padrões ou alvos, que estão relacionados com a avaliação do sucesso e do desempenho, de acordo com a aplicação de determinados critérios.
- A posição explanatória, que consiste em explicar o sucesso e os impactos do programa, e em fazer afirmações causais relativamente ao que funciona, quando e como funciona.
- A posição formativa ou orientada para a mudança, que proporciona um *feedback* positivo e mais complexo, no sentido de apoiar a auto-correcção da monitorização durante a vida de um determinado programa.
- A posição participativa/de desenvolvimento, que procura desenvolver redes, comunidades e territórios recorrendo a métodos participativos de abordagem ascendente.

Todos estes enfoques metodológicos podem ser úteis para a avaliação, pois permitem ver diferentes aspectos. Estes são familiares para aqueles que têm experiência na avaliação de programas e de políticas. Por exemplo:

- Uma análise custo-benefício utilizada na fase de apreciação do projecto é um exemplo da posição de atribuição de recursos.
- Um estudo de indicadores que procure avaliar se um determinado programa cumpriu os seus objectivos constitui um exemplo da posição baseada em padrões ou alvos.
- Uma avaliação intercalar ou uma avaliação contínua, realizada com o objectivo de proporcionar *feedback* no sentido de permitir aos parceiros e aos gestores do

programa registar – e, se necessário, reorientar – os seus programas constitui um exemplo de posição formativa.

- Uma avaliação temática, com o intuito de analisar as provas existentes em diversas intervenções, e que procura compreender quais os tipos de intervenção concebidos para dar apoio a PME com êxito em determinadas circunstâncias, constitui um exemplo de uma posição explanatória.
- Uma avaliação centrada e executada localmente, com o objectivo de reforçar e construir o consenso entre actores locais, no sentido de apoiar as suas agendas e aumentar as suas capacidades, constitui um exemplo de uma posição de abordagem ascendente ou participativa.

Uma das principais características das intervenções de desenvolvimento socioeconómico é a forma como integram diferentes programas num território ou sector comum. Estes programas resultam de diferentes domínios ou áreas de política: a educação e formação, a investigação e o desenvolvimento tecnológico, o ambiente, o desenvolvimento de infra-estruturas, etc. Este é um dos diversos desafios com que se defrontam os avaliadores em programas de desenvolvimento socioeconómico: como avaliar conjuntos complexos de intervenções inter-relacionadas, e como avaliar, não só a contribuição de cada um dos elementos, mas também as sinergias entre eles, i.e., a sua matriz de impactos cruzados.

Cada um destes domínios ou áreas de política tem os seus próprios pressupostos e tradições de avaliação, que podem ser difíceis de articular. Existem obviamente boas razões para o facto de a avaliação assumir determinados perfis em diferentes áreas de política. Estas diferenças reflectem tendencialmente características, metas e expectativas políticas específicas, em diferentes áreas de política, que afectam os tipos de medidas ou indicadores utilizados, e todo o estilo de avaliação comum. Por exemplo, os processos de avaliação de intervenções nas áreas da ciência e tecnologia tendem para a utilização de indicadores bibliométricos como medida dos resultados de investigação, enquanto a avaliação das intervenções educativas depende, em grande parte, das medidas de desempenho dos estudantes. Em algumas áreas políticas poderá existir uma tradição em que os gestores de programa executam a sua própria avaliação, relativamente às intervenções por si geridas. É o caso, por exemplo, do desenvolvimento internacional, onde a avaliação consiste frequentemente em avaliações administrativas de projectos de campo, por parte de gestores responsáveis por um portefólio de projectos, num país ou sector específico em vias de desenvolvimento. Noutras áreas de intervenção política, o envolvimento directo na avaliação de gestores com responsabilidades nos projectos seria questionado em termos da sua independência e objectividade. É prática comum solicitar uma avaliação económica de projectos infra-estruturais antes de comprometer recursos. Este procedimento será menos habitual no caso de projectos onde os investimentos em infra-estruturas constituem uma parte mais reduzida dos investimentos ou recursos totais aplicados para apoiar uma intervenção em particular.

Três tradições filosóficas

Existem três tradições filosóficas subjacentes às abordagens metodológicas alargadas da avaliação, utilizadas em programas de desenvolvimento socioeconómico.

O positivismo constituiu a base filosófica da ciência dominante a partir do século XVIII. A tradição do positivismo defende, essencialmente, a crença de que é possível obter conhecimentos objectivos através da observação. A aplicação dos mesmos instrumentos de observação por pessoas diferentes deveria permitir chegar às mesmas conclusões sempre que a análise fosse realizada por técnicas objectivas, levando aos mesmos resultados independentemente da pessoa responsável pela aplicação da técnica. As tradições positivistas procuram descobrir regularidades e leis (tal como acontece nas ciências naturais). As explicações assentam na agregação de elementos individuais, nos

seus comportamentos e interações. Isto constitui a base do reducionismo, segundo o qual o todo é compreendido analisando as partes, a base de métodos de inquérito e de modelos econométricos utilizados em avaliação. Na melhor das hipóteses, estes métodos podem proporcionar provas quantificáveis das relações entre os recursos aplicados nas intervenções e os seus resultados.

As limitações da tradição no contexto da avaliação do desenvolvimento socioeconómico decorrem das dificuldades de medição de muitos dos resultados significativos, da complexidade das interações entre as intervenções e outros factores, e a consequente ausência de perspectivas relativamente àquilo que funciona.

As limitações de uma forma pura de positivismo são hoje largamente reconhecidas. Entre estas encontram-se: a dificuldade de observar a realidade quando o que se pode observar se encontra normalmente incompleto e, por conseguinte, necessita de ser interpretado à luz de enquadramentos ou teorias; a inevitabilidade de efeitos instrumentais, segundo a qual o que pode ser observado é sempre mediado, simplificado ou inclusivamente distorcido pelas ferramentas e técnicas utilizados na recolha de dados; a dificuldade, na maioria dos contextos humanos, em esperar encontrar regularidades e leis que não variam de acordo com os contextos locais; problemas de complexidade, onde os próprios fenómenos mudam na sua interação, muitas vezes de forma imprevisível; e os julgamentos subjectivos e centrados no valor das pessoas que constroem a sua própria realidade, particularmente importante em muitos contextos de desenvolvimento social, onde os problemas como a exclusão social se encontram tão sujeitos a juízos como os factos irrefutáveis.

Estas limitações do positivismo levaram ao surgimento de diversas escolas pós-positivistas. A mais radical, rejeitando a maioria dos pressupostos do positivismo, é o construtivismo, que nega a possibilidade de conhecimento objectivo. O realismo, por outro lado, concentra-se na compreensão dos diferentes contextos e das teorias ou enquadramentos que permitem a explicação e a interpretação.

Ou seja, o *construtivismo* preconiza que é apenas através das teorizações do observador que se pode compreender o mundo; as construções existem, mas não podem ser necessariamente medidas; os factos são sempre orientados pelas teorias; e os factos e os valores são interdependentes. Nesta tradição, os avaliadores e as partes interessadas encontram-se no centro do processo de inquérito. O avaliador possui uma grande probabilidade de assumir um papel reactivo, interactivo e orquestrativo, reunindo diferentes grupos ou partes interessadas com perspectivas divergentes para exploração mútua e para gerar consenso. O avaliador desempenha um papel essencial no estabelecimento de prioridades das perspectivas expressas, e na negociação entre as partes interessadas. As partes interessadas constituem, muitas vezes, a fonte de dados mais importante, mas podem executar-se e utilizar-se outros inquéritos específicos e informações geradas externamente para informar diferenças de perspectiva marcadas.

O *realismo* procura abrir a caixa negra das políticas e dos programas, no sentido de desvendar os mecanismos responsáveis pela mudança. Ao fazê-lo, a tradição reconhece que os programas e as políticas se encontram integrados em processos sociais e organizacionais estruturados em diversas camadas, e que deverá ter-se em conta a influência destas diferentes camadas, bem como de diferentes contextos. A ênfase encontra-se no inquérito social, que explica regularidades interessantes em padrões de contextos-mecanismos-resultados. Os sistemas em estudo são vistos como abertos. Nesta tradição, o enfoque dos avaliadores encontra-se nos mecanismos causais subjacentes e na explicação dos motivos pelos quais as coisas funcionam em diferentes contextos. É provável que o avaliador construa relações do tipo professor-aluno com decisores políticos, profissionais e participantes. As avaliações temáticas têm tendência

para funcionar de acordo com esta tradição. Os estudos de caso comparativos aprofundados constituem abordagens características do trabalho de avaliação, na tradição realista.

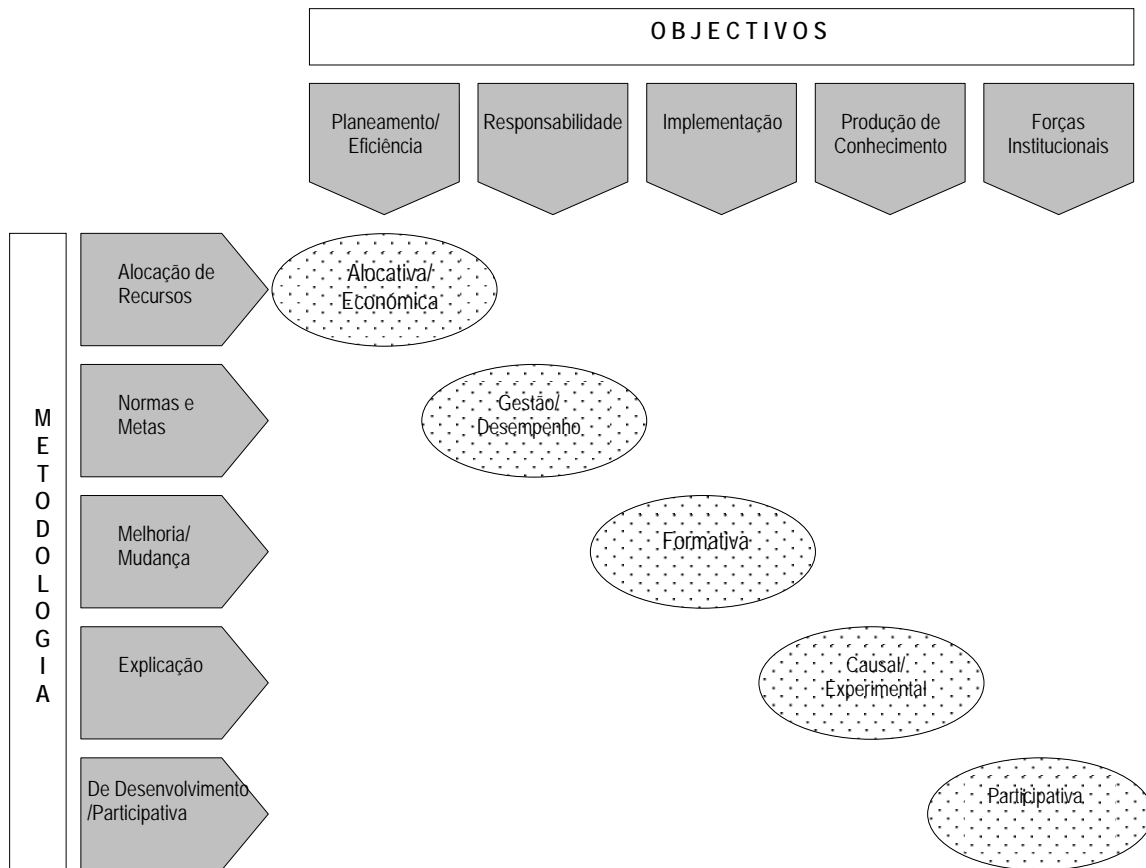
Na prática, é pouco provável que os avaliadores se vejam como operando exclusivamente à luz de qualquer uma destas tradições filosóficas, mas a tendência é para uma ou outra, dependendo das circunstâncias da avaliação. Em termos genéricos, as ferramentas de avaliação utilizadas na tradição do positivismo serão úteis para efeitos de escrutínio. As abordagens realistas apresentam uma maior probabilidade de gerar perspectivas formativas, sobretudo nos casos em que o trabalho de avaliação decorre num contexto onde a política esteja a ser desenvolvida. As abordagens construtivistas poderão ser particularmente úteis para acertar os programas, mas dependem especialmente do grau de confiança e de química entre o avaliador e as partes interessadas.

É importante reconhecer estas diferentes tradições, quanto mais não seja para ajudar a explicar por que é que as abordagens da avaliação podem ser tão diferentes umas das outras. Para os avaliadores (e para todos aqueles que pedem avaliações) é certamente útil ser claro relativamente às suas preferências e tradições filosóficas.

Se os dois eixos da avaliação, i.e., os objectivos da avaliação e a metodologia de avaliação, forem inter-relacionados, é possível começar a identificar alguns dos diferentes tipos de avaliação predominantes na avaliação do desenvolvimento socioeconómico. Os dois eixos e tipos de avaliação encontram-se ilustrados na Caixa *Os dois eixos da avaliação: Objectivos e Metodologia*. Obviamente, as tarefas de avaliação específicas possuem, normalmente, mais do que um objectivo e poderão aplicar mais do que um tipo de metodologia. Por exemplo, uma avaliação formativa destinada a melhorar a implementação de um programa também pode contribuir para a produção de conhecimento, do mesmo modo que os tipos de gestão de eficiência podem ajudar a reforçar as instituições e as redes, quando realizados de forma adequada. Assim, uma avaliação específica não corresponde necessariamente, de uma forma inequívoca, a um dos cinco tipos. Pelo contrário, é provável que se aproxime de um dos tipos. Ao mesmo tempo, é provável que um sistema de avaliação reflecta elementos de diversos tipos de avaliação.

Caixa: Os Dois Eixos da Avaliação: Objectivos e Metodologia

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA



Na intersecção dos dois eixos, dos objectivos e da metodologia, existem cinco tipos de avaliação principais. São eles:

- *Avaliações alocativas/económicas* na intersecção dos objectivos de planeamento e de eficiência com as metodologias de atribuição de recursos;
- *Avaliações orientadas para a gestão/eficiência*, na intersecção da responsabilidade, como finalidade, com o estabelecimento de normas e metas enquanto metodologia;
- *Avaliações formativas*, na intersecção da implementação e realização como objectivo, com as metodologias orientadas para a melhoria e para a mudança;
- *Avaliações causais/experimentais* na intersecção da produção de conhecimento como objectivo, com as metodologias que procuram oferecer explicações;
- *Avaliações participativas* na intersecção do reforço institucional (ou das redes) como objectivo, com as metodologias de desenvolvimento social/organizacional.

Estes tipos não podem ser aplicados de forma rígida, mesmo que isso ajude a posicioná-los de forma aproximada. Por exemplo, a avaliação formativa também pode ajudar a reforçar as instituições e contribuir para que a equipa de gestão cumpra as metas, e as avaliações alocativas/económicas podem contribuir para os fins de responsabilidade e produção de conhecimento, bem como para o planeamento e decisões ex-ante relativamente à eficiência. Porém, estes tipos captam algumas das principais correntes de pensamento sobre a avaliação e o desenvolvimento socioeconómico. Referimos, por exemplo, a importância da gestão do desempenho, da participação e do reforço das capacidades. E todos estes factores se encontram integrados nos tipos de avaliação identificados neste âmbito.

Estes cinco tipos de avaliação oferecem um ponto de partida útil para a discussão de métodos e técnicas na secção *A Selecção de Métodos e Técnicas* do GUIA e dos Manuais Técnicos que o acompanham.

Uma distinção importante destaca-se nesta diversidade de finalidades, métodos e tipos de avaliação. Trata-se da distinção entre a avaliação que se destina a avaliar, medir e demonstrar a eficácia de políticas e programas, e a avaliação que procura aconselhar, desenvolver e otimizar as políticas e os programas. Muitas das avaliações pedidas ao abrigo dos Fundos Estruturais Europeus correspondem à primeira categoria, embora também exista uma expectativa de que as avaliações ex-ante e intercalares contribuam para melhorias no programa.

No entanto, muitas das avaliações em contextos de desenvolvimento socioeconómico consistem integralmente em aconselhamento, desenvolvimento e melhoria. Este GUIA procura adoptar as duas abordagens alargadas, apesar de terem diferentes implicações. Por exemplo, as avaliações externas que procuram avaliar e medir a eficácia estão mais direccionadas para as estruturas de avaliação, incluindo a sua independência. Por outro lado, as avaliações internas estão menos preocupadas com a independência do que em proporcionar informações e um aconselhamento atempado, capazes de consolidar os programas a partir do seu interior. É necessário que os utilizadores deste GUIA tenham esta distinção em mente.

Quatro tipos de teoria relevantes para a avaliação

Já observámos que a teoria se tornou cada vez mais importante na avaliação contemporânea. Em parte, isto deve-se à perda de fé no positivismo puro, onde se assumia que as observações conduziam ao conhecimento e à explicação de forma independente, sem interpretação. Tanto os realistas como os construtivistas destacam, nas suas formas distintas, a necessidade de uma teoria. Mas existem razões mais práticas para reconhecer a importância da teoria, observando a máxima de que não existe nada tão prático como uma boa teoria. É apenas com o auxílio da teoria que a avaliação poderá analisar as intenções do programa e identificar a lógica de intervenção, compreender os processos de implementação e de mudança, ou explicar os efeitos parcialmente compreendidos dos programas em diferentes áreas de política. Estas teorias são de diferentes tipos. As quatro mais comuns são:

- Teoria do programa
- Teorias acerca da avaliação
- Teorias de implementação e mudança
- Teoria específica de políticas

Teoria do programa: a elaboração de Modelos Lógicos, utilizados extensivamente no contexto dos programas de desenvolvimento financiados pelo Banco Mundial e pela UE constitui um tipo de teoria de avaliação simples, que se centra nos recursos, realizações, resultados e impactos do programa. A Teoria da Mudança constitui uma abordagem à teoria do programa de acordo com a qual se pretende abrir “a caixa negra”, indo para além da descrição dos recursos e das realizações, e procurando compreender quais são as teorias dos actores relativamente às intervenções do programa e porque consideram que devem funcionar.

Teorias sobre as práticas de avaliação: Existem cada vez mais documentos sobre as práticas de avaliação, ou seja, sobre o que a avaliação procura fazer e o que parecem ser abordagens eficazes. Estas teorias constituem uma filtragem do que se aprendeu com o estudo de avaliações passadas. Por exemplo, como garantir a utilização das

avaliações, como tirar conclusões a partir de determinados indicadores, e como atribuir um valor a um programa ou intervenção.

Teorias de implementação e de mudança: Entre estas incluem-se entendimentos sobre alterações de política, a disseminação da inovação, o comportamento administrativo e organizacional e a liderança. As teorias resultam, sobretudo, de ciências políticas e de estudos organizacionais. A sua aplicação na avaliação pode condicionar o sucesso das intervenções do programa.

Teorias específicas das políticas: Existe um corpo teórico associado ao desenvolvimento socioeconómico, isto é, à forma como decorre o desenvolvimento, espacial e sectorialmente. Existem diversos corpos teóricos na maioria das áreas de política prioritárias, tais como a educação, saúde, emprego, protecção ambiental, etc. O Manual Técnico I descreve as teorias relacionadas com cada um dos temas e áreas de política.

Uma vez que a concepção de intervenções assenta por norma numa filosofia resultante da teoria das áreas de política, é tanto útil como normal que os avaliadores possuam alguns conhecimentos das teorias relevantes para os temas e áreas de política em causa.

1.4. AVALIAÇÃO COMO FORMA DE REFORÇAR O DESENVOLVIMENTO SOCIOECONÓMICO

- A forma como avaliamos depende do que se vai avaliar!
- O contexto das políticas da UE
- Fundamentos teóricos do desenvolvimento socioeconómico
- Valor acrescentado da comunidade
- Fazer o melhor num mundo imperfeito

A forma como avaliamos depende do que se vai avaliar!

Este não é o local para analisar em toda a sua abrangência a natureza do desenvolvimento socioeconómico. Contudo, são necessárias algumas notas introdutórias, porque a forma como se avalia está intimamente relacionada com o objecto de avaliação. Os avaliadores referem a necessidade de clareza relativamente ao objecto de avaliação, e o desenvolvimento socioeconómico constitui um objecto muito particular.

Obviamente, o desenvolvimento socioeconómico integra diversas intervenções possíveis, incluindo infra-estruturas, educação e formação, ciência e tecnologia, prevenção criminal e programas activos do mercado de trabalho, em diversas combinações. Porém, é possível fazer algumas generalizações, e identificá-las logo à partida prepara os alicerces para pensar e executar avaliações de desenvolvimento socioeconómico.

O mais importante, ainda que óbvio, é que o desenvolvimento socioeconómico se refere ao desenvolvimento. As definições do desenvolvimento socioeconómico nem sempre são consistentes; porém, normalmente contemplam os aspectos que se seguem:

- O desenvolvimento constitui um processo descontínuo, que nem sempre pode ser previsto ou controlado. Mais do que uma questão de implementação, constitui uma iniciação e gestão de um processo de mudança, que representa um corte no que poderia, de outro modo, ter decorrido;
- Uma dimensão espacial - todo o desenvolvimento decorre em algum território, e, certamente, nas intervenções dos Fundos Estruturais, relaciona-se intimamente

com a política regional. Esta dimensão é mais forte em alguns programas e prioridades do que noutros (por exemplo, os programas de desenvolvimento local em zonas rurais e urbanas), mas estão sempre presentes;

- Uma base existente - o desenvolvimento socioeconómico procura assentar nas fundações já existentes, e que são vistas como tendo mais potencial. Isto realça a dimensão temporal: o desenvolvimento decorre sempre num momento futuro, embora algumas áreas de política (como sejam o desenvolvimento sustentável) possam ter um horizonte temporal maior do que outras).
- Existe uma dimensão quantitativa e qualitativa - trata-se, quer do crescimento em números (de rendimento, emprego, empresas, etc.), como do aumento da qualidade de trabalho, ambiente, oportunidades de educação, etc.
- Existe uma política e uma dimensão normativa - o desenvolvimento pode prosseguir em direcções distintas, e a política estabelece um quadro de prioridades e valores, ao abrigo do qual se fazem escolhas. Por exemplo, o facto de a Estratégia de Lisboa procurar, quer aumentar o emprego, quer alcançar uma economia de conhecimento dinâmica exprime prioridades de valor.

Isto tem consequências para a avaliação, como por exemplo:

- Os métodos adoptados têm, muitas vezes, que ser capazes de registar a mudança e os processos de mudança (incluindo as decisões tomadas no decorrer da implementação dos programas), bem como de medir os resultados;
- Analisar a integração de diversas medidas num único território constitui um requisito essencial de avaliação do desenvolvimento socioeconómico;
- A avaliação ex-ante, a pré-programação ou as avaliações de planeamento necessitam de identificar recursos sobre os quais poderá assentar o desenvolvimento;
- Juntamente com medidas quantitativas (descrições, indicadores e modelos), também é necessário avaliar dados qualitativos do conteúdo, padrões de estilo e relevância das medidas;
- Os enquadramentos políticos e as declarações de valor ou normativas associadas (por exemplo, relativamente à solidariedade social, ao desenvolvimento sustentável e à igualdade de oportunidades) definem critérios essenciais para uma avaliação, e o que deverá constituir o enfoque da avaliação.

Entre as características mais importantes do desenvolvimento socioeconómico, destacam-se as seguintes. Em geral, a maioria dos programas de desenvolvimento socioeconómico:

- Procuram resolver problemas estruturais persistentes e, aparentemente, insolúveis, ou necessidades de adaptação fundamentais. Muitas vezes as indústrias encontram-se em sectores em declínio, as empresas não são competitivas, a administração pública possui capacidades limitadas, os grupos sociais agora excluídos já foram excluídos há muito tempo, os sistemas de educação e formação possuem uma ligação fraca às necessidades económicas e as infra-estruturas são, regra geral, fracas. Estas circunstâncias podem ser comparadas com programas onde as intervenções são menos ambiciosas, por exemplo, onde implicam a introdução de um novo sistema de qualificações ou a actualização de um sistema rodoviário existente.
- São constituídos por diversas intervenções, que têm como objectivo reforçar-se mutuamente. Por exemplo, podem combinar infra-estruturas com programas de formação, desenvolvimento de pequenas empresas e programas de transferência de tecnologia num único território. Mesmo um programa sectorial ou temático específico pode incluir diversas intervenções ou medidas. Isto decorre de uma perspectiva de que muitos problemas estruturais persistentes são multi-

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

dimensionais, e só podem ser tratados se estas diversas dimensões mudarem simultaneamente.

- São concebidos à medida das necessidades dos seus contextos, comparativamente a muitas intervenções públicas que são normalizadas em diversos contextos. Por isso, os programas de início de negócios no norte da Alemanha ou no sul de Portugal podem ter objectivos semelhantes, mas, muito provavelmente, abordam as suas actividades de formas diferentes, reflectindo as circunstâncias, a história, os recursos e estratégias nacionais ou regionais mais alargadas, nesses diferentes territórios.
- São, no entanto, planeados e financiados ao abrigo de um quadro nacional ou transnacional mais alargado. Assim, apesar de concebidos à medida de contextos específicos, os programas socioeconómicos são, muitas vezes, orientados por conceitos, estratégias ou políticas mais amplas. É isto que acontece com os Fundos Estruturais Europeus, delineados por um compromisso geral para com a coesão socioeconómica através da redução de disparidades de PIB *per capita* e outras condições em toda a UE. O mesmo acontece com muitos programas socioeconómicos nacionais e regionais.
- Têm uma abordagem ascendente e descendente. São concebidos para responder às necessidades e prioridades de partes interessadas e actores específicos, que poderão basear-se num território ou sector ou, então, envolvidos em prioridades como a protecção ambiental ou a igualdade de oportunidades. Estes actores são vistos como parceiros no sector do desenvolvimento socioeconómico. De facto, em muitos programas socioeconómicos que adoptam uma abordagem de desenvolvimento local, estes parceiros assumem a liderança na definição de prioridades, concebendo programas e implementando - e, eventualmente, gerindo - as realizações do programa.
- Estando comprometidos com a mudança estrutural e sistémica, os programas de desenvolvimento socioeconómico nem sempre conseguem alcançar os seus objectivos de longo prazo num único período de programação. Além disso, uma vez que possuem ambições a longo prazo, estas políticas e estes programas preocupam-se, normalmente, com a sustentabilidade e a viabilidade dos resultados do desenvolvimento. Será por essa razão que, geralmente, os programas de desenvolvimento socioeconómico não só incluem resultados ou realizações convencionais, como sejam sistemas de transportes optimizados ou novos cursos de formação, mas incluem igualmente objectivos relacionados com a reforma das disposições institucionais e das capacidades administrativas que garantam a sustentabilidade destes resultados para as gerações futuras.

É possível relacionar directamente estas características do desenvolvimento socioeconómico com a avaliação. A Caixa *Implicações das características dos programas de desenvolvimento socioeconómico para a avaliação* apresenta uma síntese de algumas das principais relações.

Caixa: Implicações das características dos programas de desenvolvimento socioeconómico para a avaliação		
Características do Programa	Pressupostos decorrentes	Implicações para a avaliação
Necessidades de desenvolvimento persistentes e estruturais	Natureza, a longo prazo, da alteração dos objectivos alcançáveis irá levar tempo e necessitará de uma mudança sistémica	A avaliação deverá captar o início da mudança, a longo prazo, e implementar sistemas de registo da mudança ao longo do tempo. A avaliação deverá considerar o sistema mais vasto, bem como resultados específicos.

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

Natureza multi-dimensional dos programas e das intervenções	Presume-se que as Intervenções e as medidas interagem e reforçam-se mutuamente	A avaliação deverá analisar a interacção entre diversos programas/intervenções. A avaliação deverá considerar a complexidade e as sinergias
Programas adaptados às definições	Os programas e as medidas diferem, mesmo quando os objectivos são os mesmos. Os contextos também diferem.	A avaliação necessita de considerar as intervenções na sua definição. As avaliações deverão avaliar a relevância, e ajudar a identificar aquilo que funciona em diferentes contextos. É difícil estabelecer regras gerais.
Num quadro político mais alargado	Cada um dos programas de desenvolvimento socioeconómico propõe, de certa forma, os objectivos de um quadro mais alargado.	Os avaliadores poderão retirar critérios destes quadros mais alargados, a um nível mais elevado, bem como dos objectivos de programas específicos.
As parcerias da base para o topo são importantes	Os parceiros da base para o topo são sempre importantes. Por vezes, articulam as necessidades e as prioridades através de conhecimentos locais/regionais, e possuem, frequentemente, a voz dominante na concepção e implementação.	A avaliação tem de aplicar métodos participativos e da base para o topo. Sempre que os actores locais/regionais são dominantes (e.g. desenvolvimento territorial/local), a avaliação deve suportar actividades de autogestão/direcção
Sustentabilidade	Os programas deverão incluir alterações sistémicas, no sentido de apoiar a sustentabilidade dos seus resultados	A avaliação deverá centrar-se nas alterações sistémicas, incluindo o desenvolvimento de capacidades - que influenciam a sustentabilidade

A caixa começa por explicar a razão pela qual determinados assuntos aparecem em destaque neste GUIA. A ênfase dada ao registo da mudança a longo prazo; ao desenvolvimento de capacidades; às interacções e sinergias entre as medidas; e aos métodos participativos resulta directamente das características dos programas de desenvolvimento socioeconómico e dos pressupostos que contêm.

Um aspecto deste quadro que deve ser destacado diz respeito às parcerias. É por o desenvolvimento socioeconómico ser tão multi-facetado, agregando diferentes intervenções de diferentes domínios políticos, e envolvendo diferentes agências e actores (formais e informais), que as parcerias se tornaram muito importantes na avaliação. A maioria dos programas resultantes do nível europeu baseia-se em parcerias, tal como acontece com os programas de apoio nacional e regional. De facto, nos programas dos Fundos Estruturais o Princípio do Partenariado integra-se nos regulamentos do programa. As parcerias colocam desafios particulares à avaliação.

Os parceiros partilham sempre alguns compromissos comuns, caso contrário, não se envolveriam em programas de desenvolvimento socioeconómico. No entanto, também possuem os seus próprios interesses. Estes interesses comuns e específicos exigem que as avaliações e os gestores de programa integrem nos seus planos o que, à partida,

possam parecer objectivos, critérios e questões de avaliação contraditórios. Em certa medida, as avaliações podem lidar com este fenómeno integrando cuidadosamente os diferentes contributos dos parceiros. A forma como isto se pode processar irá ser analisada na secção *A Concepção e Implementação da Avaliação* neste GUIA. Também há métodos que podem ser desenvolvidos de modo a integrar os diversos critérios das diferentes partes interessadas (ver a secção *A Seleccção de Métodos e Técnicas* do GUIA). Contudo, trabalhar em contexto de parceria também possui implicações para o papel do avaliador ou da equipa de avaliação. Estes deverão ser mais do que colectores passivos das diferentes prioridades das partes interessadas. Os avaliadores têm de assumir um papel mais activo, de modo a apoiar a obtenção de consenso, tanto no início, quando as avaliações são concebidas, como no decorrer do ciclo de avaliação, incluindo no momento de disseminação de resultados e de discussão de conclusões. Coordenar, negociar e facilitar o consenso também são capacidades necessárias.

O contexto das políticas da UE

Como já se referiu, qualquer programa de desenvolvimento socioeconómico insere-se no âmbito de um contexto político mais complexo. Isto aplica-se tanto aos Fundos Estruturais Europeus, como aos programas a nível de cada Estado-membro.

Os programas dos Fundos Estruturais possuem as suas próprias orientações e prioridades horizontais específicas. As prioridades horizontais para o período 2000-2006 são: emprego e recursos humanos, desenvolvimento sustentável, ambiente, igualdade de oportunidades para homens e mulheres, sociedade da informação e desenvolvimento local. Os Fundos Estruturais também estão integrados numa agenda política europeia mais alargada, que define cada vez mais o conteúdo dos programas.

É importante referir que as principais orientações políticas estão a delinear as abordagens ao desenvolvimento socioeconómico na Europa, uma vez que estas possuem implicações para o enfoque global de uma avaliação e de critérios específicos que poderiam ser aplicados para julgar o êxito dos programas.

Entre os principais factores de delineação da política dos programas de desenvolvimento socioeconómico encontram-se:

- A política de coesão da Comunidade, que tem como objectivo reduzir as disparidades - normalmente, medida em termos de PIB *per capita* - entre as regiões mais desenvolvidas e as menos desenvolvidas. A política tem como objectivo apoiar estas acções, com maior probabilidade de contribuir para a redução das disparidades económicas, sociais e territoriais na União (CEC 2001). Fá-lo através da concentração de recursos em áreas em situação de atraso. Estes fundos são atribuídos, na sua maioria, a regiões onde o PIB *per capita* é inferior a 75% da média da UE. A política não procura solucionar directamente estas disparidades; antes, concentra-se em intervenções que afectam os factores determinantes assumidos da infra-estrutura física de crescimento económico, incluindo os transportes, os recursos humanos e a capacidade para gerir investimentos e serviços, sobretudo através de um sector de PME viável, uma gestão pública eficaz e através das tecnologias da informação.
- A estratégia de Lisboa e o conseqüente processo, decorrente da sessão especial do Conselho Europeu em Março de 2000, acordou toda uma nova estratégia global, com a meta a 10 anos de tornar a União a economia de conhecimento mais competitiva e dinâmica do mundo, em 2010, com capacidade para sustentar o crescimento económico, com mais e melhor emprego, e maior coesão social. Embora o processo de Lisboa tenha continuado a evoluir (acrescentando o desenvolvimento ambiental sustentável, o empreendedorismo e elementos de competitividade nos Conselhos Europeus subsequentes), o núcleo central procura

integrar correntes bem enraizadas da política europeia. Em particular, reúne as políticas de emprego e de competição, com origem em 1993, no Livro Branco sobre Crescimento, Competitividade e Emprego, de Delors (COM [93] 700 final), e os sistemas de solidariedade social, bem-estar e emprego, que integram o que é frequentemente conhecido como o Modelo Social europeu.

- Este aspecto foi reforçado pela Estratégia Europeia para o Emprego, que, de certa forma, se sobrepõe e reforça o processo de Lisboa, por exemplo destacando dimensões sociais como a reconciliação do trabalho e da vida familiar, e melhorando o emprego de todos aqueles que se encontram em situação desfavorecida no mercado de trabalho, além de procurar melhorar a empregabilidade e contribuir para o pleno emprego. A Estratégia Europeia para o Emprego também está relacionada com os Fundos Estruturais, e existe um compromisso no sentido de concertar acções com os Planos de Acção para o Emprego dos Estados-membros.

São estes tipos de correntes políticas que influenciam a natureza do desenvolvimento socioeconómico num contexto europeu. As relações entre estas diversas políticas e os Fundos Estruturais são sobretudo evidentes no caso dos novos Estados-membros, destacando-se, por exemplo, as medidas ambientais relacionadas com a estratégia de desenvolvimento sustentável, endossada pelo Conselho Europeu de Gotemburgo, e as prioridades de recursos humanos relacionadas especificamente com a agenda de Lisboa na orientação da Comissão Europeia (Comunicação da Comissão: Orientações Indicativas Adicionais para os Países Candidatos à Adesão COM2003 110 final). Embora possa haver diferenças de ênfase entre aqueles que destacam a solidariedade e aqueles que privilegiam a desregulamentação do mercado de trabalho, é de consenso alargado que o desenvolvimento necessita de integrar as dimensões económica e social.

As avaliações de desenvolvimento socioeconómico devem, portanto, ter em conta este objecto de avaliação muito específico. Porém, a descrição das tendências políticas acima referida também destaca alguns dos desafios à avaliação colocados pelo desenvolvimento socioeconómico como objecto de avaliação. Por exemplo:

- Muitas das políticas e dos programas (por exemplo, desenvolvimento sustentável, empreendedorismo, etc.) constituem objectos compósitos complexos. Por isso, o desenvolvimento sustentável agrega recursos económicos e sociais, bem como recursos naturais e capacidades humanas. A sociedade da informação e do conhecimento combina elementos organizacionais, tecnológicos e institucionais. A avaliação destes elementos coloca problemas metodológicos relativamente à forma como se deve definir o objecto, bem como relativamente ao modo de ajuizar se as melhorias ou as metas foram alcançadas.
- Nos Fundos Estruturais europeus e noutras intervenções de desenvolvimento socioeconómico, as prioridades horizontais são reforçadas adicionalmente por políticas transversais, e muitas iniciativas socioeconómicas requerem a integração de recursos, medidas e mecanismos de implementação. Estas iniciativas baseadas em diversas medidas decorrem em contextos sobrepostos (que podem ser espaciais, institucionais ou sectoriais). Avaliar medidas integradas e a forma como as suas partes integrantes interagem faz parte da complexidade que caracteriza os programas de desenvolvimento socioeconómico.
- Devido, em parte, ao facto de os objectos compósitos e integrados estarem sujeitos a avaliação, deverá colocar-se, muitas vezes, o problema da escolha da unidade de análise mais adequada. Esta é, parcialmente, uma questão de decidir se é necessário avaliar uma medida em particular, um programa em geral, ou uma política mais alargada. Também se levantam questões de agregação e desagregação: decidir se devem dividir-se as iniciativas complexas nas suas partes e avaliar o sucesso de cada uma delas ou se é preferível procurar definir uma medida agregada, que poderá não ser muito significativa na prática. (Por

exemplo, poderá existir uma única medida de desenvolvimento sustentável otimizado?). Na prática, é provável que a escolha da unidade de análise se caracterize por um compromisso e avaliações de negociações entre a aprendizagem de aspectos mais específicos das intervenções e a possibilidade de proporcionar perspectivas pertinentes para o panorama geral.

- Alguns dos programas e políticas utilizados no desenvolvimento socioeconómico podem ser avaliados de acordo com diferentes lógicas e sistemas de valor. Em termos globais, por exemplo, deveria considerar-se as disparidades económicas ao nível local, regional, nacional ou europeu? Em relação a prioridades horizontais específicas, como por exemplo a igualdade de oportunidades, deverão estas ser julgadas no âmbito de um mercado de trabalho de referência ou tendo em conta critérios como o equilíbrio da vida familiar-profissional ou as contribuições destas políticas para a solidariedade social?

Muitas das grandes questões de avaliação que se pode colocar relativamente a estas políticas e a políticas semelhantes não podem ser respondidas ao nível do programa. Por exemplo, para estimar a extensão da convergência nos diversos Estados-membros e nas diversas regiões, é, normalmente, necessário recorrer a análises comparativas de dados a nível europeu, relacionadas com o crescimento económico, com a produtividade e com os rendimentos pessoais. Porém, as avaliações do programa podem contribuir para a interpretação dessas avaliações de política a larga escala. Também podem proporcionar alguns dados de tipo diferente, nomeadamente: o que está a acontecer realmente no terreno, a definição do alvo das medidas, a qualidade e a quantidade dos recursos e das realizações do programa, as reacções e satisfação das partes interessadas, e a forma como os programas são geridos e implementados.

Fundamentos teóricos do desenvolvimento socioeconómico

Já referimos que uma área teórica relevante na avaliação dos programas de desenvolvimento socioeconómico consiste no próprio desenvolvimento. Os principais conjuntos de pressupostos e de teorias utilizados para explicar e interpretar os resultados dos programas socioeconómicos decorrem da reflexão contemporânea acerca do desenvolvimento socioeconómico.

Embora tradicionalmente (provavelmente até ao final dos anos 70) o elemento em destaque fosse a gestão da procura através de subsídios regionais e outras subvenções (por exemplo, subsídios de desemprego), a corrente actual está mais orientada para a oferta ou para as capacidades. Isto pode assumir diversas formas, mobilizando recursos subutilizados, aumentando as capacidades e o valor dos recursos existentes e transferindo novos recursos para uma região ou sector.

Alguns exemplos destas teorias e pressupostos contemporâneos mais comuns incluem:

- **Economia do conhecimento.** O conceito de uma economia caracterizada pela produção, disseminação e aplicação da informação e do "know-how" como sendo produtos, e a utilização geral de novos modos de produção, constituídos pela aplicação de informação e conhecimento, na produção de produtos e serviços.
- **Capital humano.** Reconhece que os recursos humanos, em particular as taxas de literacia e de educação, saúde geral e esperança de vida, criam condições de produtividade que permitem aos grupos sociais transformar o seu capital humano em maior prosperidade económica.
- **Capital social.** Uma vez mais, relacionado com o bem-estar humano, mas a um nível social, em vez do nível individual, através das redes sociais e institucionais (incluindo, por exemplo, parcerias e associações), que apoiam uma acção social

eficaz. Isto inclui confiança social, normas e redes, bem como sistemas políticos e legais, que sustentam a coesão social.

- **Exclusão social.** Centra-se nas disparidades, entre as pessoas e as comunidades, no acesso a oportunidades de serviços, emprego e infra-estruturas. A exclusão social influencia o desenvolvimento económico equilibrado e sustentável, o desenvolvimento do emprego e dos recursos humanos, a actualização e a protecção ambiental, e a promoção da igualdade de oportunidades. Uma melhor coesão social e económica tornou-se um dos objectivos prioritários da UE, e constitui um conceito abrangente, relacionado com a política de emprego, protecção social, habitação, educação, saúde, informação e comunicações, mobilidade, segurança e justiça, cultura e lazer.
- **Transferência tecnológica.** O pressuposto da transferência tecnológica, possível através do acesso aos bens públicos, permite às regiões menos desenvolvidas recuperar a capacidade de absorver e imitar, é, à luz destas teorias, mais importante do que ser o inovador inicial.

Uma fonte importante de ideias em termos estruturais e regionais decorre do que, muitas vezes, se designa por nova geografia económica. Existem duas teorias que são, normalmente, associadas a esta escola:

- *Vantagem comparativa local.* Pressupõe que as regiões possuem potencial de crescimento ao explorar a sua vantagem comparativa distintiva. Isto pode assumir diversas formas em termos negociais (bens e serviços), e em termos de bens posicionais não transaccionáveis (paisagem e cultura, muitas vezes a base do turismo).
- *Pólos de crescimento.* Esse crescimento pode necessitar, inicialmente, de uma concentração, a nível regional, que conduza, inicialmente, a um aumento das disparidades, em vez da sua redução. Contudo, presume-se que estas disparidades sejam posteriormente atenuadas, e que o crescimento alastre.

Estas diferentes teorias subjacentes às prioridades políticas como a coesão, a Estratégia Europeia para o Emprego e o Processo de Lisboa possuem implicações na avaliação, como por exemplo:

- Dada a importância do investimento público adicional na teoria de crescimento neo-clássica, é importante determinar se o investimento público será, de facto, suplementar, ou se simplesmente substitui (ou destitui) outros tipos de investimento. Como se verá na previsão da adicionalidade, a inércia e a substituição constituem um elemento comum na avaliação do desenvolvimento socioeconómico.
- Dada a importância de orientar novos recursos para as regiões e áreas de investimento onde o potencial é superior, é importante que os avaliadores demonstrem em que medida os recursos são direccionados e, de facto, atingem os beneficiários pretendidos. A interacção entre os métodos de atribuição de recursos baseados na disparidade temática e regional também é relevante para os avaliadores de programas socioeconómicos, uma vez que os critérios temáticos podem contrariar a orientação para as regiões e para os beneficiários em situação de maior atraso.

Em muitos casos, a avaliação dos programas contribui para responder a grandes questões políticas através de uma compreensão qualitativa do que se passa realmente no terreno. Por exemplo, muitas das ideias e teorias subjacentes ao desenvolvimento socioeconómico realçam a importância da tecnologia como recurso para o crescimento. Os processos de avaliação de um programa socioeconómico neste quadro político necessitam, assim, de:

- Esclarecer quais os tipos de investimento tecnológico que estão a decorrer, por exemplo, a nível de infra-estruturas, capacidades, investimento das empresas em novo equipamento.
- Avaliar a qualidade destes investimentos em termos, por exemplo, da sua relevância, absorção e utilidade percebida pelos actores locais e regionais.
- Procurar dados relativamente à eficiência, rapidez e grau de consenso com que estas medidas tecnológicas estão a ser implementadas.
- Identificar os tipos de resultados preliminares e alterações consequenciais que parecem estar a surgir na sequência destes recursos e contribuições.
- Considerar em que medida as inovações observadas parecem ser sustentáveis, e em que medida estão a ser desenvolvidas novas capacidades, em vez de vantagens unilaterais.

Este tipo de informação qualitativa e do terreno poderá ser útil para a avaliação de diversas formas. Permite:

- Ajudar os gestores de programas e outras partes interessadas a compreender melhor o que está a acontecer, permitindo-lhes otimizar a orientação, a gestão e a implementação dos seus programas.
- Garantir aos decisores políticos que os recursos económicos estão a ser despendidos nos fins para que foram propostos, muito antes da conclusão do programa.
- Proporcionar informações contextuais que permitam aos avaliadores das políticas, responsáveis pela avaliação da eficácia dos instrumentos políticos, interpretar de forma mais geral as suas conclusões.

Valor acrescentado da Comunidade

Embora o GUIA não esteja exclusivamente relacionado com a avaliação das intervenções dos Fundos Estruturais europeus, muitos dos conselhos são pertinentes para estes programas. Geralmente, a avaliação destes programas considera especificamente o valor acrescentado das intervenções para a Comunidade. Não existe uma forma correcta ou errada de identificar e medir o valor acrescentado da Comunidade, e a sua consideração terá que ser concebida à medida do contexto e das intervenções específicas. Existem, no entanto, alguns pontos de partida:

- Em primeiro lugar, deverá considerar-se os princípios subjacentes à intervenção. Entre estes, incluem-se os princípios de parceria, subsidiariedade e adicionalidade. O processo de planeamento e de implementação envolveu os parceiros adequados? As principais decisões foram tomadas ao nível adequado mais reduzido? As intervenções foram complementares àquilo que teria ocorrido, noutras circunstâncias?
- Em segundo lugar, como é que se deve (ou deveria) avaliar o contributo da intervenção para a agenda política mais alargada da UE (política de coesão, Estratégia de Lisboa, Estratégia Europeia para o Emprego, a predominância de género, etc.)?
- Em terceiro lugar, deverá avaliar-se a medida em que se registaram impactos sobre as instituições e sistemas, incluindo o intercâmbio transnacional e a transferência de boas práticas ou de domínio político, que são uma consequência do financiamento da intervenção por parte da UE. (Neste contexto, conforme será discutido na secção do GUIA sobre a Capacitação para a Avaliação, as intervenções da UE foram, elas próprias, um estímulo para a promoção de uma cultura de avaliação).

- Em quarto lugar, a avaliação do valor acrescentado da Comunidade deve considerar o alcance da complementaridade entre as intervenções da UE e os instrumentos de política nacional.

Para proporcionar uma avaliação equilibrada, devem considerar-se estes e outros aspectos do valor acrescentado da Comunidade comparativamente aos custos transaccionais complementares.

Fazer o melhor num mundo imperfeito

Tanto os avaliadores como os decisores políticos poderão partir de pressupostos demasiado ambiciosos relativamente ao que a avaliação pode alcançar em qualquer programa específico ou contexto político. Num mundo ideal, os programas estão bem estruturados, os dados administrativos estão sempre disponíveis, as avaliações são pedidas atempadamente, os promotores dos programas têm objectivos claramente definidos, a filosofia das intervenções encontra-se devidamente investigada e articulada, os recursos adequados colocados à disposição estão em conformidade com as metas definidas, as capacidades necessárias estão disponíveis para agregar uma equipa de avaliação, os pressupostos políticos permanecem constantes no decorrer da vida do programa respectivo, e, através de um bom planeamento, as conclusões da avaliação chegam a tempo de informar as revisões de política e representam oportunidades de reprogramação numa fase anterior ao planeamento.

Previsivelmente, a maior parte das vezes o mundo real não é assim! As prioridades políticas mudam ou evoluem enquanto os programas estão a decorrer, as avaliações ex-ante não foram correctamente realizadas, os objectivos dos programas acabam por ser um compromisso entre as prioridades contraditórias dos diferentes parceiros e partes interessadas, sabem-se quais os indicadores que se pretende recolher mas os dados não estão disponíveis, e o ciclo de avaliação não se encontra sincronizado com o ciclo político. Nestas circunstâncias demasiado familiares, a avaliação poderá, mesmo assim, dar um grande contributo, mas requer uma dupla abordagem.

Em primeiro lugar, os avaliadores têm de estar dispostos a produzir o que efectivamente é possível, dados os recursos e contextos institucionais existentes, ao mesmo tempo que têm de reconhecer as limitações do que é possível afirmar com confiança. Corre-se o risco de, sob as diversas pressões contratuais, os avaliadores caírem na tentação de prometer mais do que podem realmente realizar, ou procurarem retirar conclusões mais firmes do que os dados disponíveis permitem.

Ao mesmo tempo, é necessário reconhecer os problemas que surgem quanto aos recursos, dados, prazos, a estrutura do programa, o planeamento e as capacidades disponíveis, no decorrer de uma avaliação. Nesta base, uma das conclusões mais importantes de uma avaliação pode ser a identificação das condições necessárias para melhorar a qualidade, reduzir os incumprimentos, aumentar a relevância e a utilização das avaliações no futuro. A responsabilidade de agir com base nestas conclusões pertence aos responsáveis que solicitam as avaliações e que gerem os programas. Estes aspectos são componentes essenciais da tarefa de construção da capacidade de avaliação, conforme se discute com algum pormenor na secção do GUIA relativa à Capacitação para a Avaliação e no Manual Técnico III, relacionado com este tema. Haverá, claro, determinadas situações em que as condições são de tal forma deficientes que seria pouco sensato realizar uma avaliação. Uma conclusão deste tipo poderia ser alcançada com base numa análise de viabilidade da avaliação (ver Manual Técnico II). Nestas circunstâncias, a razoabilidade de prosseguir com o programa em causa poderá ser questionada.

A maioria dos programas existe em circunstâncias menos extremas, independentemente da imperfeição das circunstâncias. Também não se deve subestimar o valor de informações sistemáticas, mesmo que reduzidos, quando nem sequer existiam anteriormente. No mínimo, o processo de planeamento de uma avaliação, de identificação da lógica de intervenção e dos recursos disponíveis, e de identificação dos momentos em que os resultados da avaliação podem contribuir para decisões de reprogramação, poderão ajudar a clarificar a reflexão, para além de gerar informações ou resultados.

1.5. REGRAS DE OURO

Esta secção do GUIA apresentou algumas das principais questões relacionadas com a avaliação do desenvolvimento socioeconómico. Integradas nos diversos temas discutidos, relativamente às vantagens da avaliação, à natureza da tarefa de avaliação, e aos requisitos específicos da política socioeconómica, encontram-se diversas regras de boas práticas de obtenção complexa, que, de acordo com a experiência, podem ajudar a planear, executar e a utilizar a avaliação. Em resumo, estas regras de ouro foram agregadas da seguinte forma:

1. É necessário lembrar que a intenção da avaliação é melhorar os programas, e não avaliar por avaliar. Sempre que projectar uma avaliação, pergunte-se de que forma é que os resultados irão melhorar a vida dos cidadãos, a prosperidade e o bem-estar das regiões e a competitividade dos actores económicos. Se não conseguir dar uma resposta positiva a estas questões, talvez deva observar novamente a necessidade de uma avaliação ou, pelo menos, observar a forma como foi concebida.
2. Alinhar os ciclos temporais da avaliação com os ciclos temporais dos programas e das políticas é uma meta que vale a pena alcançar! É deste modo que se pode assegurar a máxima contribuição das avaliações. É melhor realizar uma avaliação incompleta ou imperfeita dentro do prazo do que atingir uma melhoria de 10% na qualidade da avaliação e falhar o prazo oportuno, em que os decisores políticos e os gestores de programas podem utilizar os resultados da avaliação.
3. As diferentes partes interessadas, como por exemplo os decisores políticos, os profissionais, os gestores e os cidadãos, possuem diferentes expectativas em relação à avaliação. Se os interesses de uma das principais partes interessadas forem ignorados, isto poderá enfraquecer uma avaliação, porque esta terá uma fraca concepção e/ou os seus resultados não terão credibilidade. O envolvimento dos decisores políticos e de todos os responsáveis pelos programas permite que os resultados sejam levados a sério. Identifique as suas partes interessadas, descubra quais são os seus interesses numa avaliação e envolva-os!
4. As avaliações deverão estar totalmente integradas no planeamento e gestão do programa. Os gestores de programa têm de pensar na avaliação como um recurso: uma fonte de *feedback*, uma ferramenta para aumentar a eficiência, um alerta precoce para os problemas (e soluções), e uma forma de sistematizar o conhecimento. A avaliação não constitui simplesmente uma imposição externa. Obviamente, este truísmo possui implicações para os avaliadores, que têm de considerar as preocupações dos gestores de programa (e as suas parcerias), e de procurar levar a sério a sua necessidade de respostas a perguntas difíceis.
5. Obter um bom trabalho dos diversos grupos que constituem a comunidade profissional contemporânea de avaliações requer a construção de relações e de equipas. As relações têm de ser construídas ao nível nacional, regional e europeu, entre as diferentes tradições existentes entre os avaliadores, sejam

eles cientistas sociais, economistas, analistas políticos ou consultores de gestão. Por isso, realize conferências e apoie o intercâmbio profissional, no sentido de assegurar a difusão de conhecimento e de "know-how". Esta é uma forma de construir as capacidades. A um nível micro, a prioridade é a integração e a combinação de diferentes capacidades e competências, em equipas de avaliação.

6. A avaliação não consiste apenas em olhar para trás para classificar o sucesso ou o fracasso e para atribuir culpas. Ela tem uma contribuição a fazer em cada uma das fases do ciclo dos programas. Em particular, a avaliação pode, desde a sua fase mais precoce, reforçar os programas ajudando a seleccionar lógicas de intervenção e a revelar fraquezas na concepção do programa, permitindo adoptar medidas de correcção desde cedo.
7. Já não é aceitável reunir grandes quantidades de dados, na crença de que estes possam proporcionar respostas a todas as questões de avaliação. A dragagem de dados é quase sempre ineficiente. Isto não significa que os sistemas de dados não sejam essenciais; estes, deverão ser implementados numa fase inicial (ver a secção do GUIA relativamente a *A Seleção de Métodos e Técnicas*). Contudo, ao serem claras relativamente aos pressupostos, e ao basearem-se em teorias disponíveis, sendo claras relativamente ao tipo de avaliação que é necessário, as avaliações podem ser mais focadas e oferecer uma gama mais alargada dos recursos disponíveis.
8. O contexto político é um enquadramento importante no qual é necessário situar as avaliações. Obviamente, a política muda, ou está a ser constantemente reafirmada em termos diferentes, e com prioridades subtilmente mutáveis. Porém, é sempre necessário estar atento às discussões e às decisões políticas, de modo a assegurar que as avaliações se adequam às prioridades políticas. Os critérios mais alargados que têm de ser concebidos nas avaliações decorrem normalmente do quadro político mais alargado.
9. Embora se tenha defendido que todas as partes interessadas são importantes (ver 3 acima), a ênfase no desenvolvimento socioeconómico dá particular destaque a um grupo importante, e frequentemente negligenciado: os beneficiários destinatários das intervenções do programa. Integrar a sua voz, como comunidades locais, grupos marginalizados e novas entidades económicas nas avaliações significa mais do que pedir as suas opiniões. Também significa integrar os seus critérios e opiniões numa avaliação e admitir que a sua experiência e benefícios constituem a justificação das intervenções do programa. Isto está em consonância com a lógica das abordagens ascendentes, participativas e descentralizadas que são tão comuns actualmente no desenvolvimento socioeconómico. Também é por isso que os métodos reactivos e participativos se transformaram numa parte tão importante do conjunto de ferramentas dos avaliadores.
10. Sejamos pragmáticos! Vivemos num mundo imperfeito, onde os recursos são limitados, os administradores nem sempre são eficientes, a coordenação é imperfeita, os conhecimentos estão misturados e os dados nem sempre estão disponíveis. Vale a pena, no entanto, dar pequenos passos, trabalhar com aquilo que está disponível e aumentar, ainda que ligeiramente, a eficiência e a legitimidade dos programas públicos. Mesmo os resultados mais modestos podem fazer uma grande diferença, sobretudo se integrados numa perspectiva de longo prazo para construir capacidades e permitir avaliações mais ambiciosas no futuro.

(Índice da Parte II)

2. PARTE II: A CONCEPÇÃO E IMPLEMENTAÇÃO DA AVALIAÇÃO

2.1. A concepção e o planeamento da avaliação

2.2. A implementação e a gestão da avaliação

2.3. A utilização da avaliação e a gestão do conhecimento

2.4. Regras de ouro

2.1. A concepção e o planeamento da avaliação

- Avaliação e programação
- Planear o trabalho de avaliação
- Definir as questões e os critérios de avaliação
- Seleccionar métodos de avaliação, atribuir responsabilidades e distribuir os recursos
- Elaborar os Termos de Referência

Avaliação e programação

Na secção do GUIA relativa à avaliação e ao desenvolvimento socioeconómico, deu-se particular destaque ao facto do desenvolvimento socioeconómico, não sendo uma ciência exacta, ser complexo e incerto, quanto mais não seja porque existem diversas teorias que, individualmente, sustentam diferentes estratégias de desenvolvimento, que não são necessariamente compatíveis entre si.

O planeamento da documentação constitui, em primeiro lugar - e sobretudo -, uma parte essencial do ciclo de planeamento e de projecto e, como tal, é uma contribuição fundamental para a política de desenvolvimento socioeconómico. Também são, contudo, documentos políticos, que normalmente têm de obter o acordo dos diversos agentes a diferentes níveis territoriais, e com valores, metas e prioridades muito diferentes. Não é novidade, por isso, que estes documentos são em muitos casos vagos, que procuram cobrir todos os ângulos possíveis do problema, de uma forma por vezes genérica (mesmo que isto signifique dividir em partes muito pequenas os recursos disponíveis), e que alguns dos objectivos neles identificados são contraditórios entre si, para não dizer absolutamente incompatíveis.

Este perigo está sobretudo presente quando a complexidade dos programas aumenta, como acontece com a nova geração de políticas de desenvolvimento socioeconómico, com destaque para a abordagem territorial, o ênfase na sustentabilidade, a necessidade de parcerias alargadas e/ou de diversos princípios dominantes.

De acordo com uma perspectiva tradicional, estas incertezas tornam a tarefa de avaliação de resultados dos programas de desenvolvimento socioeconómico mais difícil, se não impossível. Sem metas claras, uma teoria de intervenção coerente e uma concepção do programa precisa - presume-se -, a identificação dos elementos a avaliar e dos critérios de avaliação torna-se arbitrária e subjectiva. Independentemente do mérito destes conhecimentos adquiridos, também existe outra forma de ver o problema. É precisamente devido à existência de diversos objectivos e de programas complexos que a avaliação se torna essencial.

Numa situação ideal, estas questões de avaliação deveriam ser consideradas na fase de formulação do programa, e isto deveria ajudar a evitar problemas como objectivos contraditórios. Conceptualizar os resultados esperados em termos operacionais e, por conseguinte, mensuráveis (por exemplo, construindo desde o início um sistema de monitorização e de indicadores) constitui um meio extremamente capaz de ajudar os decisores a formular melhores programas. Por isso, implicar a avaliação e os avaliadores tão cedo quando possível constitui um pré-requisito para um bom programa de desenvolvimento socioeconómico.

É importante referir que podem utilizar-se diversas técnicas de avaliação (tal como analisar a viabilidade da avaliação, construir uma teoria do programa e efectuar uma análise SWOT), desde uma fase inicial, para esclarecer as prioridades iniciais e a lógica subjacente às intervenções do programa.

Utilizar técnicas avaliativas, e adoptar uma abordagem avaliativa logo desde o início, irá ajudar as partes interessadas a desenvolver uma linguagem comum, enquanto a identificação de alguns resultados intermédios tangíveis e mensuráveis irá ajudar a fazer o enquadramento geral do processo de implementação. Isto também permite definir marcos que assegurem o registo do processo ao longo da vida do programa.

Mas mesmo depois de a concepção dos documentos estar concluída, o contributo da avaliação para a programação pode ser muito importante. Nesta fase, a avaliação pode ajudar a esclarecer um conjunto confuso de metas e/ou uma lista genérica de possíveis projectos. A utilização dos designados modelos lógicos ajuda a fazer um registo dos interesses e das prioridades políticas das diferentes partes interessadas. Se estes modelos forem desenvolvidos em interacção com os avaliadores e com os decisores políticos, podem, inclusivamente, conduzir à reestruturação do programa. Esta é, provavelmente, a contribuição mais importante da avaliação para a fase de programação.

É importante chamar a atenção para uma verdade essencial e frequentemente esquecida: o valor da avaliação para os decisores políticos reside tanto em colocar as questões correctas, como em fornecer respostas precisas. (Este assunto é debatido em seguida com maior detalhe). Uma metodologia sólida e dados fiáveis são muito importantes e comportam boas respostas a questões de avaliação. Porém, em última instância, o valor acrescentado da avaliação, para os decisores, consiste em colocar questões como: Quais os objectivos do exercício? É este o equilíbrio que realmente se pretende entre as diferentes metas? Poderá considerar-se este nível de eficiência como sendo satisfatório?

No entanto, criar as condições perante as quais as questões de avaliação podem ser respondidas o mais precisamente possível continua a ser uma meta importante. Neste contexto, quanto mais preciso for o programa, mais explícitas serão as potenciais compensações e sinergias entre as diferentes metas e os diferentes objectivos, mais exigente será a lógica do programa, mais fiável será a teoria do programa (i.e., as relações causais entre as acções projectadas e os resultados esperados), mais abrangente será o sistema de indicadores, mais coerente será a execução da sequência dos resultados intermédios e do processo de implementação, e mais fácil e mais precisa será a avaliação. Este tipo de situação ideal proporciona análises sólidas, fiáveis e úteis relativamente ao sucesso ou ao fracasso do programa. Irá aumentar a responsabilidade perante o sistema de governança e desenvolver uma melhor compreensão das formas como se poderá atingir a meta geral do desenvolvimento socioeconómico sustentável.

Na perspectiva de um potencial avaliador, existem vantagens sempre que:

- As metas sejam tão claras quanto possível, e os conceitos referidos são definidos e compreendidos por todos.

- Os objectivos estão relacionados com as intervenções (ou grupos de intervenções) e com os resultados mensuráveis.
- Se as intervenções possuírem diversos objectivos, está associado um peso explícito a cada um dos objectivos.
- Os objectivos integram alvos, cujo alcance pode ser medido. Os alvos devem possuir prazos explícitos e devem constituir o resultado da negociação entre os decisores políticos e os responsáveis pela implementação.
- Existem incentivos que estimulam inequivocamente o cumprimento dos objectivos.
-

Estas condições ajudam o processo de avaliação a definir critérios relativamente aos quais é preciso julgar se uma intervenção pode, está direccionada para, ou se alcançou um bom resultado.

Na prática, estas condições são pouco comuns e, como se verá de seguida, o próprio processo de avaliação deve, muitas vezes, contribuir para a definição dos objectivos e dos critérios que reflectem as ambições definidas para as intervenções.

Seja qual for a origem dos objectivos, estes proporcionam um conjunto importante de critérios relativamente aos quais é possível julgar o cumprimento das intervenções.

Planear o trabalho de avaliação

Esta secção discute mais precisamente as diversas actividades e questões necessárias para planear trabalho de avaliação, em particular:

- Definir o objecto e o âmbito da avaliação;
- Identificar e envolver as partes interessadas;
- Analisar a teoria do programa e os objectivos políticos subjacentes às intervenções.

Definir o objecto da avaliação

A decisão de avaliar constitui uma oportunidade para definir os limites do programa em termos das dimensões institucional, temporal, sectorial e geográfica. Isto designa-se âmbito da avaliação ou "objecto da avaliação". Definir o âmbito de uma avaliação consiste em colocar a questão: O que é que vai ser avaliado?

O âmbito da avaliação pode ser especificado em, pelo menos, quatro aspectos:

- institucional (nível europeu, nacional ou local);
- temporal (período temporal a ser considerado);
- sectorial (social, industrial, ambiental, rural, etc.); e
- geográfico (em que parte do território europeu, em que região, cidade, reserva natural, etc.).

Um programa é, por sistema, delimitado pelos aspectos financeiros e pelo território correspondente, bem como pelo período de programação. É, porém, útil considerar primeiro o seguinte:

- Pretende-se limitar a avaliação ao financiamento do programa ou incluir outro financiamento nacional, regional ou local, relacionado, num grau mais ou menos elevado, directamente com o programa?
- Pretende-se limitar a avaliação a intervenções na área elegível ou alargar as observações a determinadas áreas limítrofes, com problemas de desenvolvimento semelhantes?
- Pretende-se limitar a avaliação ao financiamento atribuído no âmbito do ciclo de programação em causa ou, em certa medida, incluir o financiamento de ciclos anteriores?

É, normalmente, útil adoptar uma definição relativamente rigorosa do âmbito da avaliação. A experiência mostra que, durante o processo de avaliação, as partes interessadas podem querer analisar quase tudo. Para obter conclusões, a avaliação deverá limitar-se a uma análise do programa e das suas interdependências mais essenciais, relativamente a outras políticas e intervenções públicas.

Este risco de alargamento do âmbito é particularmente elevado no caso das avaliações ex-ante. Estas podem tornar-se exercícios de previsão ou especulação muito afastados do objecto da avaliação. Numa avaliação ex-ante é melhor limitar rigorosamente o âmbito da avaliação às propostas do programa.

Os requerentes da avaliação não estão, muitas vezes, dispostos a restringir o âmbito da avaliação às questões que pretendem cobrir. Uma possível contribuição dos avaliadores consiste em identificar as questões mais centrais para o êxito do programa como resultados prováveis das teorias dos programas de desenvolvimento que identificam a lógica de intervenção e as cadeias de implementação. Por vezes, a melhor forma de estabelecer prioridades entre as questões de avaliação e o enfoque da avaliação consiste em discutir restrições de ordem prática, como sejam o tempo e os recursos. Este é, muito provavelmente, uma das conclusões essenciais da fase de concepção.

Para que uma avaliação seja útil, as decisões que podem tomar-se e que podem ser informadas pela avaliação devem ser explicitadas da forma mais precisa possível. Em muitos casos, os requerentes, não pretendendo influenciar demasiado a equipa de avaliação, estão relutantes em exprimir antecipadamente as alterações que consideram dever ser introduzidas no programa, ou as suas dúvidas relativamente à eficácia de uma acção específica. A intenção é louvável: não revelar nada antecipadamente para ver se a equipa de avaliação chega às mesmas conclusões! A experiência mostra, contudo, que é pouco provável que a avaliação possa documentar as decisões pretendidas, caso estas não sejam do conhecimento prévio de todos aqueles que vão recolher e analisar os dados no terreno. A realidade socioeconómica é altamente complexa e a equipa de avaliação defronta-se com um elevado número de observações e possibilidades de optimização. Verificar hipóteses que possam, dadas as circunstâncias, ser de pouco interesse para técnicos, gestores ou outras partes interessadas não é realista.

Identificar e envolver as partes interessadas

Como já foi referido, o desenvolvimento socioeconómico inclui diversos tipos diferentes de projectos, programas e políticas - o que significa que o número de actores ou partes interessadas é, normalmente, bastante elevado. A experiência de avaliação parece mostrar que este aspecto está longe de ser um obstáculo para uma boa avaliação. Pelo contrário, oferece oportunidades que devem ser exploradas de modo a colocar as questões mais adequadas e dar as respostas mais úteis.

As actividades no terreno influenciam o número de partes interessadas envolvidas na criação das políticas. Em particular, a ênfase no princípio do parceria baseia-se na perspectiva de que o envolvimento de instituições não governamentais e actores da sociedade civil irá aumentar a qualidade do desenvolvimento socioeconómico, quer do ponto de vista de definição de um vasto conjunto de objectivos, quer em termos de facilitar o processo de implementação.

Entre outros factores que reforçaram a tendência para o envolvimento de grupos de instituições e actores, grandes e diversificados, incluem-se a influência de parcerias verticais e horizontais, o surgimento de governança multi-nível e a aplicação de subsidiariedade, o estabelecimento de prioridades políticas transversais como o

desenvolvimento sustentável ou a igualdade de oportunidades e o reconhecimento do papel desempenhado pelo capital social no desenvolvimento socioeconómico.

O aparecimento do desenvolvimento local e territorial, onde diferentes sectores políticos e fontes de financiamento se integram numa tentativa de otimizar o desenvolvimento socioeconómico de uma área, torna a identificação das partes interessadas, e o seu envolvimento no processo de formulação de programas (uma abordagem ascendente), um passo essencial de todo o exercício.

Mesmo em programas e projectos mais simples, existe sempre uma série de actores cujos interesses são afectados, positiva ou negativamente, pela actividade planeada ou implementada.

Em todo o caso, no entanto, identificar os actores potencialmente afectados (em avaliações ex-ante), os actores realmente afectados (em exercícios intercalares ou ex-post), e envolvendo-os, de algum modo, no processo de avaliação é essencial para ter em conta pontos de vista, efeitos indirectos ou consequências não esperadas que podem ser muito significativas para descrever os efeitos, compreender as cadeias de causalidade e analisar os resultados.

A ênfase na identificação das partes interessadas foi, até ao momento, aplicada em termos dos seus benefícios práticos para melhor compreender o programa, para colocar questões mais pertinentes e para obter informação de elevada qualidade. Contudo, existe uma filosofia adicional para proceder à identificação e envolvimento das partes interessadas. Os avaliadores, juntamente com os gestores do programa, têm interesse em assegurar as questões de propriedade dos resultados de avaliação. Talvez apenas deste modo é que todos aqueles que estão envolvidos levarão a sério as avaliações, irão agir com base nas recomendações, ou definirão as suas próprias prioridades de acção.

A primeira questão que se deve colocar, depois de o âmbito da avaliação ser definido, é, portanto, bastante linear: Quem são as pessoas, os grupos ou as organizações interessados na avaliação da intervenção e/ou interessados no processo ou nos resultados da própria avaliação? Esta fase no discurso de avaliação é conhecida como identificação das partes interessadas.

Numa situação ideal, este exercício deveria decorrer antes de definir os detalhes da avaliação a realizar: tendo em conta os seus pontos de vista, é possível decidir quais são as questões mais relevantes a responder. Mas, mesmo no caso de isso não ser possível, por exemplo por não ser possível identificar todas as partes interessadas numa fase inicial, é desejável haver algum tipo de envolvimento.

A segunda questão que se deveria colocar é: como é que é possível assegurar que as partes interessadas fornecem os recursos e informações adequados para a concepção, gestão ou conteúdo do exercício de avaliação?

O envolvimento das partes interessadas pode ocorrer a níveis muito diferentes:

- No mínimo, os avaliadores devem certificar-se de que as partes interessadas proporcionam elementos (dados, informações, opiniões, etc.) como parte do processo de avaliação. Muitos dos métodos e técnicas descritos no Manual Técnico II podem utilizar-se com esta finalidade: entrevistas individuais, inquéritos, "focus grupos", etc.
- Do outro lado do espectro, as partes interessadas podem estar envolvidas na coordenação do estudo completo, incluindo a definição de prioridades, as questões de avaliação e os critérios associados. Muitas vezes, isto implica o

envolvimento da Comissão Directiva do projecto de avaliação, como se verá na discussão sobre a gestão do processo de avaliação.

Na prática, o envolvimento das partes interessadas na maior parte dos programas fica entre os dois. Se a participação das partes interessadas no Comité Directivo estiver restringida aos parceiros sociais e institucionais oficiais, uma forma de fornecer *feedback* a outros actores que podem proporcionar informações e opiniões é largamente praticada através da disseminação de relatórios, reuniões *ad hoc* e instrumentos semelhantes.

Teorias de programa e modelos lógicos

A fase de consulta das partes interessadas também proporciona uma oportunidade de reconstruir a lógica do programa, antes do seu lançamento. Como foi referido na Parte I do Guia, existem teorias diferentes, e muitas vezes concorrentes, subjacentes às intervenções neste campo.

Numa situação ideal, todos os programas ou políticas devem explicitar claramente o conjunto de pressupostos com base no qual é possível alcançar a meta desejada, no nosso caso o desenvolvimento socioeconómico, através dos recursos atribuídos e das intervenções financiadas. Estes pressupostos devem ser consistentes entre eles, e suportados por elementos concretos. Na prática, isto é raro, sobretudo no mundo complexo do desenvolvimento socioeconómico.

Um passo adicional no planeamento de avaliação consiste, por isso, em reconstruir a teoria do programa subjacente ao objecto de avaliação. Trata-se, sobretudo, de avaliar a capacidade do programa para alcançar os seus objectivos esperados (i.e., o desenvolvimento). Uma identificação inequívoca das razões pelas quais se deve esperar este facto constitui uma condição prévia importante para colocar as questões de avaliação correctas.

Isto realça a forma como a programação e a avaliação estão inter-relacionadas. Os gestores de programas e os decisores técnicos devem estar conscientes da existência de instrumentos disponíveis que ajudam a reconstruir a cadeia que liga os objectivos gerais do programa, os objectivos intermédios específicos, as actividades implementadas pelos responsáveis respectivos e, finalmente, os resultados e as consequências destas actividades. A secção A Seleção de Métodos e Técnicas do GUIA (e o Manual Técnico II) discute e dá exemplos de diversas ferramentas e técnicas (modelos lógicos, enquadramentos lógicos, teoria de programa, teorias da mudança) que podem ajudar na reconstrução da lógica de intervenção dos programas e nas cadeias de implementação.

Em conjunto com a consulta e a análise das partes interessadas, a aplicação destes métodos pode ajudar a identificar os possíveis aspectos críticos da implementação do programa, e, por conseguinte, concentrar-se na avaliação de forma adequada.

Definir as questões e os critérios de avaliação

Definir as questões de avaliação

A definição das questões de avaliação é uma parte essencial da fase inicial de qualquer exercício de avaliação.

As questões de avaliação podem colocar-se a diferentes níveis. Podem ser:

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- *Questões descritivas*, com o intuito de observar, descrever e medir a mudança (o que aconteceu?)
- *Questões causais*, que procuram compreender e avaliar as relações de causa e efeito (como e em que medida é que as ocorrências decorreram do programa?)
- *Questões normativas*, que aplicam critérios de avaliação (os resultados e os impactos são satisfatórios, tendo em conta os alvos, as metas, etc.?)
- *Questões prospectivas*, que procuram prever os acontecimentos resultantes das intervenções planeadas (as medidas de combate ao desemprego em determinado território irão originar efeitos negativos para o ambiente ou para os actuais activos empregados?)
- *Questões críticas* que têm por objectivo apoiar a mudança, em muitos casos, a partir de uma perspectiva de valor (como poderão as políticas de igualdade de oportunidades ter uma melhor aceitação por parte das PME? Ou quais são as estratégias eficazes para reduzir a exclusão social?)

Numa situação ideal, as questões de avaliação devem possuir as características que se seguem:

- Cada questão deve corresponder a uma necessidade real de informação, compreensão e/ou identificação de uma nova solução. Se uma questão tiver apenas interesse em termos de conhecimento novo, sem constituir um contributo imediato para a discussão pública ou para o processo de decisão, trata-se mais de uma questão de investigação científica e não deve ser incluída numa avaliação.
- Cada questão deve relacionar-se com um impacto, um grupo de impactos, um resultado ou uma necessidade. Ou seja, diz respeito, pelo menos parcialmente, a elementos externos ao programa, nomeadamente aos seus beneficiários ou ao seu ambiente económico e social. Se determinada questão estiver relacionada apenas com a gestão interna de recursos e de realizações, pode, provavelmente, ser tratada de forma mais eficiente no decorrer de uma monitorização ou auditoria.
- Cada questão refere-se apenas a um critério de opinião. Esta qualidade de uma questão de avaliação pode, por vezes, ser difícil de alcançar, mas a experiência tem vindo a demonstrar que se trata de um factor essencial para a utilidade da avaliação. Sem critérios de opinião claramente identificados desde o início, o relatório de avaliação final raramente proporciona conclusões claras.

Finalmente, deve notar-se que nem todas as questões colocadas pelos requerentes da avaliação e pelos gestores de programa são adequadas enquanto questões de avaliação. Algumas são demasiado complexas, aplicáveis a longo prazo e exigem dados que não estão disponíveis. Outras questões nem sequer requerem esforços de avaliação, mas podem ser tratadas através de sistemas de monitorização existentes, consultoria de gestores, ou recurso a processos de auditoria ou outros sistemas de controlo.

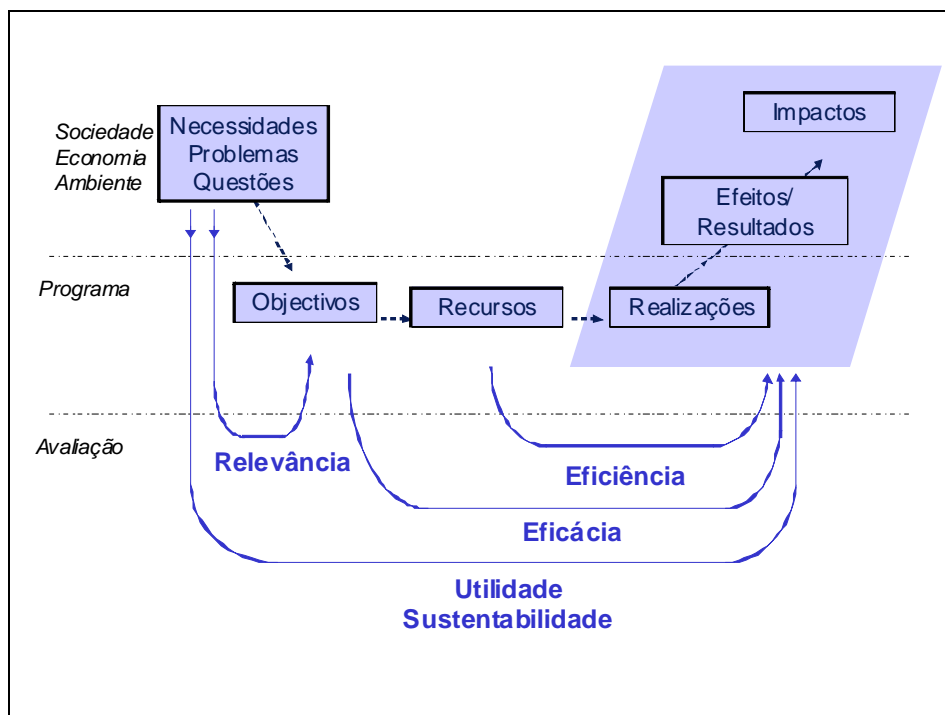
Critérios de avaliação

As questões de avaliação que incluem critérios de opinião pertencem a uma das quatro categorias seguintes:

- As que estão relacionadas com a relevância do programa;
- As que estão relacionadas com a sua eficácia;
- As que estão relacionadas com a sua eficiência; e
- As que estão relacionadas com a sua utilidade.

Estas quatro categorias principais estão representadas na caixa *Seleccionar questões de avaliação prioritárias*.

Caixa: Seleccionar questões de avaliação prioritárias



O termo *relevância*, no contexto de uma avaliação, refere-se à adequabilidade dos objectivos explícitos do programa em relação aos problemas socioeconómicos que devem solucionar. Na avaliação ex-ante, as questões de relevância são muito importantes, uma vez que o enfoque se encontra na escolha da melhor estratégia ou na justificação da estratégia proposta. Na avaliação intercalar, o objectivo consiste em verificar se o contexto socioeconómico evoluiu da forma esperada, e se a evolução coloca em causa um objectivo em particular.

O termo *eficácia* designa a verificação dos objectivos formulados no programa: se estão a ser alcançados, quais os sucessos e as dificuldades, e se as soluções seleccionadas são as adequadas, bem como qual a influência dos factores externos, com origem fora do programa.

O termo *eficiência* é avaliado comparando os resultados obtidos ou, de preferência, os impactos produzidos, e os recursos mobilizados. Por outras palavras, os efeitos obtidos estão em conformidade com os recursos aplicados? (Os termos economia e minimização de custos são, muitas vezes, equiparados à eficiência).

As questões básicas das avaliações intercalares e, mais especificamente, das avaliações ex-post, dizem respeito à eficácia e à eficiência das intervenções implementadas e de todo o programa.

Os termos *eficácia* e *eficiência* utilizam-se, normalmente, por gestores que procuram, no contexto da monitorização, ajuizar acerca dos efeitos (em vez de o fazer sobre os resultados ou impactos respectivos). De facto, as questões relativamente ao desempenho de um programa são cada vez mais comuns no âmbito da monitorização. Dada a relevância da monitorização e da avaliação, é provável que exista um grande conjunto de questões agrupadas na rubrica desempenho.

O critério da *utilidade* julga os impactos obtidos pelo programa em relação a necessidades sociais e económicas mais alargadas. A utilidade constitui um critério de avaliação muito específico, uma vez que não faz qualquer referência aos objectivos oficiais do programa. Pode ser prudente formular uma questão de utilidade sempre que os objectivos do programa estejam mal definidos, ou sempre que haja muitos impactos inesperados. Este critério deverá, no entanto, ser utilizado com cuidado, para assegurar que a equipa de avaliação não selecciona as necessidades ou questões importantes demasiado subjectivamente. Uma forma de precaução contra este risco consiste em envolver outras partes interessadas e, em particular, os beneficiários, na selecção das questões de utilidade.

O termo *sustentabilidade* refere-se à medida em que os resultados e as realizações da intervenção são duradouros. Muitas vezes, as avaliações consideram a sustentabilidade das mudanças institucionais, bem como os impactos socioeconómicos. (Os critérios de sustentabilidade também estão relacionados com o conceito de desenvolvimento sustentável, que pode ver-se como uma definição de *utilidade*, sobretudo se, conforme referido neste GUIA e no Manual Técnico I que o acompanha, o desenvolvimento sustentável for definido como sendo referente à manutenção de capitais humanos, produtivos, naturais e sociais, em vez de constituir apenas a manutenção do ambiente para gerações futuras).

A seguir, serão referidas algumas das questões de avaliação típicas, relacionadas com os principais critérios.

Questões de avaliação relacionadas com os principais critérios de avaliação

- *Relevância*: Em que medida é que os objectivos do programa se justificam pelas necessidades? Poderá comprovar-se o motivo da sua existência? Correspondem às prioridades locais, nacionais e europeias?
- *Eficácia*: Em que medida é que os objectivos foram alcançados? As intervenções e os instrumentos utilizados produziram os efeitos esperados? Poderá obter-se mais efeitos utilizando instrumentos diferentes?
- *Eficiência*: Os objectivos foram alcançados aos custos mais reduzidos? Seria possível obter-se melhores efeitos ao mesmo custo?
- *Utilidade*: Os efeitos esperados ou inesperados são globalmente satisfatórios, do ponto de vista dos beneficiários directos ou indirectos?
- *Sustentabilidade*: Os resultados e os impactos, incluindo as mudanças institucionais, são duradouros? Os impactos irão continuar se não houver mais financiamento público?

Estes critérios não são exclusivos. Outros critérios como a equidade, a coerência, a sinergia, e a reprodutibilidade também se utilizam frequentemente em avaliação. Além disso, os critérios de avaliação e as questões de avaliação daí decorrentes podem estar relacionados com as consequências negativas e positivas esperadas ou não esperadas das intervenções. Embora os programas possuam a sua própria lógica e os seus próprios objectivos, estão integrados em políticas que definem uma finalidade mais abrangente. Podem, em última instância, ser vistos como contribuindo para a inclusão social ou para uma maior competitividade económica, apesar de o seu objectivo imediato ser a formação profissional ou a criação de novas empresas. A avaliação também não pode ser limitada aos objectivos e às prioridades do programa.

Os avaliadores também devem considerar seriamente os possíveis resultados que não faziam parte da arquitectura do programa. Entre os resultados de um programa que vão além dos objectivos formais que os avaliadores deverão considerar, incluem-se:

- As experiências e as prioridades dos beneficiários, que possuem os seus próprios critérios de eficácia do programa, poderão não estar em conformidade com os dos criadores do programa e dos decisores técnicos;
- Os efeitos perversos que não sejam simplesmente não pretendidos, mas que podem, de facto, estar em contradição com as intenções do programa, reduzindo as oportunidades, em vez de as aumentar, e exportando o emprego em vez de o criar; e
- Os resultados sugeridos por outra investigação e avaliação, possivelmente com base em teorias de desenvolvimento socioeconómico ou em experiências comparativas de outros países.

Isto constitui, por isso, um argumento para que a avaliação não seja exclusivamente orientada para os objectivos, mas que, por vezes, se encontre ao lado da lógica dos programas e adopte uma natureza independente e, inclusivamente, crítica. Esta não é, contudo, uma justificação para ignorar os objectivos do programa, mas antes um argumento para ir mais longe na prossecução da aprendizagem e da melhoria do programa.

Um conjunto de conceitos normalmente aplicado às avaliações decorre da teoria económica, incluindo:

- *Adicionalidade*: a intervenção foi suplementar àquilo que decorreria noutras circunstâncias?
- *Inércia*: a intervenção gerou realizações, resultados e impactos que teriam, de qualquer forma, decorrido?
- *Deslocação*: a intervenção causou reduções do desenvolvimento socioeconómico noutros locais?

A avaliação pode contribuir para responder às questões sobre a inércia e a deslocação quando a escala de uma intervenção ou programa é grande em relação a outras possíveis explicações para os resultados obtidos. Este poderá não ser o caso em intervenções socioeconómicas de menor dimensão.

Avaliabilidade das questões de avaliação

Assim que as questões avaliativas forem identificadas, a sua viabilidade em termos de avaliação tem que ser considerada. É necessário fazer uma avaliação prévia, no sentido de apurar se as questões avaliativas podem ser respondidas, considerando os recursos e os dados disponíveis. A equipa de avaliação, com o tempo e recursos disponíveis, e utilizando ferramentas de avaliação adequadas, pode proporcionar respostas credíveis às questões colocadas? Este elemento requer a execução de um estudo de viabilidade da avaliação.

Relativamente a cada uma das questões de avaliação, é necessário verificar, ainda que resumidamente:

- se os conceitos são estáveis,
- se é possível formular hipóteses explanatórias,
- se é possível utilizar dados disponíveis para responder à questão, sem qualquer investigação adicional,
- se o acesso ao terreno coloca problemas graves.

Diversos factores podem fazer com que uma questão se torne difícil de responder como, por exemplo, se o programa for muito recente, se ainda não tiver produzido resultados significativos, se não tiver dados disponíveis, ou se os dados forem inadequados. Estas

razões poderão levar à decisão de não proceder à avaliação, de a adiar, ou de colocar questões mais realistas.

Como indica a Caixa Seleccionar questões de avaliação prioritárias, apresentada anteriormente, considerações importantes na fase da viabilidade da avaliação constituem as probabilidades de se obter e utilizar os resultados da avaliação.

Questões relevantes incluem, por conseguinte:

- As conclusões serão utilizadas? Por quem? Para que finalidade (decidir, discutir, informar)? Quando?
- É politicamente adequado efectuar uma avaliação deste tipo nesta altura específica ou neste contexto em particular? Existe uma situação de conflito que possa comprometer o sucesso do exercício?
- Existe algum estudo recente que já tenha respondido à maioria das questões?
- As prioridades de avaliação são suficientemente estáveis?

O Manual Técnico II apresenta um capítulo dedicado aos requisitos formais de uma análise da viabilidade da avaliação, que terá certamente informações adicionais úteis para os interessados neste tema.

Seleccionar os métodos de avaliação, atribuir responsabilidades e distribuir os recursos

Seleccionar os métodos

As questões de avaliação podem ser respondidas de diversas formas. A selecção dos métodos é, por conseguinte, crucial para tirar o máximo partido dos recursos de avaliação disponíveis. Esta é, normalmente, uma escolha operacional que só pode ser finalizada quando o campo de análise for reconstruído e houver informações suficientes relativamente à disponibilidade dos dados. Contudo, durante a fase de planeamento é necessário proceder a algumas escolhas. A selecção dos métodos é influenciada:

- pela fiabilidade da teoria do programa;
- pelo nível de consenso entre os parceiros e as partes interessadas;
- pelo tipo de programa a avaliar;
- pela fase no ciclo do programa no qual decorre a avaliação;
- pelo tema ou sector de intervenção do programa.

A selecção do GUIA relativamente à selecção dos métodos e técnicas proporciona informações adicionais e orientações relativamente à selecção dos métodos. O Manual Técnico II descreve métodos e técnicas específicas, e o Glossário proporciona uma definição dos termos utilizados com menor frequência.

O papel da orientação

O facto de a avaliação nos Fundos Estruturais ser, muitas vezes, obrigatória e de tanto a Comissão Europeia como as autoridades nacionais responsáveis pelos fundos distribuírem orientações acerca de quando e como executar os exercícios de avaliação, constitui uma moeda de duas faces. Por um lado, existe uma tendência para deixar cair em rotina a decisão de se proceder à avaliação. A avaliação torna-se uma obrigação no sentido de apenas dar cumprimento às exigências das instituições acima referidas, sem um valor adicional claro para os gestores do programa. Por outro lado, pode proporcionar uma orientação bastante necessária e muito apreciada, quer às autoridades com responsabilidade de planeamento, quer às equipas de avaliação, relativamente aos comportamentos e resultados esperados. Certamente, com base no pressuposto de que as avaliações deveriam ser executadas, e de que a disponibilidade de orientação

relativamente ao seu âmbito constituiu um estímulo importante para o desenvolvimento da capacidade de avaliação, conforme discutido pormenorizadamente na Parte III deste GUIA.

As orientações são particularmente úteis para definir os parâmetros de avaliação em programas relativamente descentralizados, onde é importante adoptar prioridades, critérios e procedimentos comuns. Isto permite assegurar um grau de comparabilidade. Estas orientações têm, normalmente, sido desenvolvidas por autoridades nacionais ou europeias. Também podem ser úteis se foram desenvolvidas por gestores de programas ou inclusivamente por avaliadores, sempre que for apontada como provável a descentralização geral do processo de avaliação de um programa individual. No desenvolvimento socioeconómico, sempre que for comum fazer avaliações participativas e auto-avaliações, algumas orientações básicas poderão ser extremamente eficazes, sobretudo se as prioridades forem desenvolvidas em colaboração com os actores locais.

Decisões sobre os recursos essenciais

As avaliações também podem ser vistas como uma tentativa de crítica retrospectiva das escolhas dos gestores dos programas. Em muitos casos, têm a sensação de que já sabem a maior parte do que os avaliadores lhes vão dizer.

Torna-se por essa razão importante logo à partida envolver as autoridades políticas, ou pelo menos a gestão de topo, juntamente com os principais parceiros externos do programa no planeamento da avaliação. Isto não significa envolvê-los nas decisões mais técnicas, mas sim assegurar que estes têm a possibilidade de exercer influência em relação às quatro questões fundamentais que se seguem:

- Quais as razões da avaliação?
- Quem é o responsável pelo exercício geral?
- Quanto despender no estudo?
- Quem irá realizar o trabalho?

a) Quais as razões da avaliação?

Esta é a questão mais fundamental. Como já foi referido, existem diferentes finalidades gerais possíveis num estudo de avaliação, existem diferentes objectivos específicos e existem diferentes questões de avaliação possíveis. Assegurar que os decisores de topo, bem como os parceiros mais relevantes do programa, partilham a escolha que se reflecte nos termos de referência proporciona uma maior credibilidade a todo o exercício.

b) Quem é o responsável?

Envolve:

- os membros do Comité Directivo;
- os responsáveis pela elaboração dos termos de referência; e
- todos os que funcionam como elo de ligação entre a administração e a equipa de avaliação.

Os responsáveis devem possuir um estatuto suficientemente sénior que lhes dê acesso directo aos decisores políticos, a todos os níveis, de modo a partilhar com eles os conhecimentos que se espera que o estudo produza. Também deverão estar em consonância relativamente aos problemas teóricos e metodológicos da investigação avaliativa. Isto é essencial para formar os seus próprios julgamentos relativamente à fiabilidade do produto, bem como para colocar as questões correctas à equipa de avaliação.

Assim, numa situação ideal os responsáveis pela avaliação devem possuir alguma experiência em trabalho prático de avaliação, e tê-la realizado no passado.

c) Quanto investir?

É difícil decidir quanto investir numa avaliação, numa base *a priori*.

Em termos genéricos, no caso de programas de relativa rotina, os orçamentos necessários para a avaliação constituem uma pequena parte dos recursos do programa (normalmente, menos de 1%). Por outro lado, no caso de intervenções que são relativamente inovadoras, e de piloto na matéria, e onde a avaliação possui um forte aspecto de aprendizagem e participação, os custos podem atingir uma proporção relativamente elevada do programa (até 10%). Existem casos onde até 5% do orçamento do programa se destinou a avaliações que fazem eficazmente parte da estratégia de implementação de gestão, como por exemplo, nos casos em que a avaliação inclui um forte elemento formativo destinado a apoiar os gestores e as partes interessadas relativamente ao seu trabalho.

A forma mais adequada para determinar o orçamento deve basear-se em critérios como a natureza e o âmbito do trabalho necessário. Uma avaliação sólida requer contributos por parte de bons avaliadores e níveis significativos de compromisso, tanto por parte de dos que solicitam o trabalho, como dos parceiros.

Na prática, os programas socioeconómicos têm tendência para incorrer desnecessariamente em despesas sempre que as avaliações dizem respeito a aspectos de rotina, mas não investir o suficiente quando os programas são inovadores. Este é, obviamente, um risco sempre que se pretende que a avaliação tenha sobretudo uma finalidade de responsabilização ou de monitorização.

Os recursos orçamentais não devem constituir um factor limitativo da qualidade de uma avaliação. Contudo, existem retornos progressivamente mais reduzidos. Na fase ex-ante, o tempo disponível para informar a formulação de programas e os dados disponíveis podem ser limitados. Na fase intercalar, a dimensão dos inquéritos aos beneficiários e a dimensão da consulta aos parceiros e às partes interessadas terão uma enorme influência sobre a necessidade de recursos. Na fase ex-post, a qualidade do trabalho contínuo de monitorização e de avaliação executado pode ser o principal factor limitativo, e não o orçamento propriamente dito.

d) Quem executa a avaliação?

Uma avaliação deverá ser realizada por uma equipa externa ou interna? Cada uma destas soluções possui vantagens e desvantagens. As equipas externas possuem, em muitos casos, maiores conhecimentos da especialidade, e podem ser vistas como independentes, o que poderá constituir um factor importante para a credibilidade da avaliação. Os avaliadores internos possuem um maior grau de familiaridade com os requisitos institucionais e de gestão, e poderão ter um acesso mais facilitado a informações e a recursos humanos essenciais. Estes podem, contudo, não ser vistos como independentes, e podem não ter conhecimentos especializados. Em parte, esta questão está relacionada com decisões acerca do desenvolvimento de capacidades na administração pública. Algumas pessoas comprometeram-se seriamente, a longo prazo, com a capacidade de avaliação interna, localizada junto de unidades especializadas. Sempre que estes estejam claramente separados da gestão operacional, podem ultrapassar as preocupações relativamente à sua independência.

Existe uma série de abordagens estruturais que se destinam a assegurar a independência dos avaliadores internos, face aos gestores de programa. Uma dessas abordagens consiste em centrar a função de avaliação numa unidade organizacional ou

divisão individual, por exemplo, de planeamento ou estratégia, e não no departamento operacional. Outra consiste em assegurar que os níveis mais elevados de gestão, separados tanto das operações, como da avaliação, se encontram explicitamente envolvidos no seguimento das recomendações e conclusões de avaliação. Isto pode funcionar como um contrapeso a qualquer tendência de ignorar os relatórios de avaliação, mantendo, por exemplo, todas as partes responsáveis pelo acompanhamento envolvidas.

Porém, a independência não constitui apenas uma questão estrutural. Desenvolver uma filosofia de independência junto dos avaliadores internos (e apoiar uma filosofia semelhante junto dos avaliadores externos) poderá constituir uma forma importante de garantir a independência em termos de comportamento. Além disso, desenvolver uma cultura de avaliação junto das unidades administrativas relevantes, com capacidade de auto-crítica e aberta aos novos dados e a ideias de melhoria, também pode permitir reforçar a independência da função de avaliação.

Poderá haver lógicas diferentes mais adequadas às diferentes fases do ciclo de avaliação e dos programas. Pode ser preferível basear-se mais nos recursos internos para obter contribuições para uma avaliação formativa ou para realizar exercícios ex-ante, mas estar mais dependente de recursos externos para proceder à avaliação ex-post dos impactos e dos resultados.

Elaborar os Termos de Referência

Os Termos de Referência (TdR) constituem o documento que serve de base a uma relação contratual entre o requerente de uma avaliação e a equipa responsável pela execução do trabalho. Conceber os Termos de Referência constitui, inequivocamente, um passo vital sempre que é necessário realizar uma avaliação por consultores externos. Este trabalho também é importante sempre que parte da avaliação for realizada internamente. Os TdR poderão referir-se quer às operações de avaliação como um todo (quando atribuídas a uma mesma equipa), quer a uma parte do trabalho de investigação previsto no projecto de avaliação (análise detalhada de uma questão avaliativa).

O documento dos TdR deverá ser breve (normalmente, entre 5 e 10 páginas), complementado, se necessário, por anexos administrativos. A caixa Estrutura padrão dos termos de referência apresenta um exemplo do conteúdo dos TdR, sendo depois descrito em pormenor.

Caixa: Estrutura padrão dos termos de referência

- Enquadramento regulamentar
- Âmbito da avaliação
- Principais utilizadores e partes interessadas no estudo
- Questões de avaliação e de pesquisa
- Conhecimentos disponíveis
- Principais métodos e técnicas a utilizar
- Calendarização
- Orçamento indicativo
- Qualificações necessárias da equipa
- Estrutura da proposta
- Regras de submissão e critérios de adjudicação

i. Enquadramento regulamentar

É necessário explicitar o enquadramento legal, contratual e institucional de um programa. Isto inclui, por exemplo, regulamentações das autoridades nacionais ou da União Europeia. Deverá esclarecer-se se determinada avaliação é obrigatória e legalmente necessária, ou se é executada porque os gestores do programa a perspectivam como sendo importante.

Os TdR deverão especificar quem iniciou o projecto de avaliação e, sempre que for relevante, quem esteve envolvido na formulação do resumo da avaliação. Também se deverá especificar os motivos e as intenções subjacentes. Por exemplo: a intenção é uma alteração da mudança de política? Em caso afirmativo, porquê? A intenção é modificar os procedimentos de implementação? A intenção é redistribuir fundos?

ii. Âmbito da avaliação

Já foi referida a importância de se definir o âmbito da avaliação. Os TdR devem esclarecer o projecto/programa/política/tema a avaliar, o período em apreço, o ponto do ciclo político/programa no qual se define a avaliação, e a área geográfica de referência do estudo.

iii. Principais utilizadores e partes interessadas no estudo

Foi realçada a importância de identificar a utilização e os utilizadores da avaliação nas fases mais precoces do planeamento. É, por isso, importante incluir cláusulas acerca da forma como os resultados da avaliação serão utilizados nos TdR. Se houver envolvimento do utilizador, por exemplo numa Comissão Directiva ou de Acompanhamento, isto deve ser mencionado.

iv. Questões de avaliação e de investigação

Foi referido que é possível responder a diferentes questões de avaliação e de pesquisa (descritivas, causais, críticas, prescritivas ou normativas), e que podem utilizar-se diferentes critérios para formular as opiniões de avaliação.

É importante definir as questões de avaliação, mas também é importante limitar o número de questões que a avaliação tem como objectivo responder. Centrar a avaliação numa lista de questões limitada, mas relevantes para o requerente, garante um melhor controlo de qualidade.

v. Conhecimentos disponíveis

Os TdR devem incluir uma revisão do actual estado da arte sobre o programa e os seus efeitos. Esta deverá incluir extractos ou referências sobre documentos de programação, listas de análises anteriores e avaliações com extractos relevantes, uma descrição do sistema de monitorização implementado, quantificação dos indicadores e os diversos relatórios e bases de dados disponíveis junto dos serviços de gestão da intervenção.

Este inventário é importante para que as equipas de avaliação possam ajustar os seus métodos propostos.

vi. Principais métodos e técnicas a utilizar

Cada avaliação terá os seus próprios métodos específicos, em princípio, os mais relevantes para o seu âmbito e para o seu conteúdo. Geralmente, não é uma boa prática especificar integralmente os métodos e as abordagens, mas antes deixar uma margem de manobra considerável para que os proponentes de uma avaliação possam indicar de que forma pretendem proceder. Para os requerentes de uma avaliação, a prioridade consiste em especificar o que consideram ser as suas necessidades em termos de resultados, como por exemplo, obter respostas, face às questões-chave. Poderão ou não

especificar métodos particulares, consistentes com as suas intenções, como por exemplo a necessidade de um inquérito aos beneficiários.

A escolha recai por norma na manutenção de um grau de flexibilidade suficientemente grande que permita aos que respondem aos TdR se possam diferenciar em termos de relevância e clareza nas suas propostas metodológicas. Este aspecto é particularmente importante na fase de selecção, uma vez que avaliar as qualidades metodológicas das propostas é uma etapa crucial para seleccionar o avaliador adequado.

Sempre que possível do ponto de vista administrativo, o melhor procedimento quanto a este aspecto consiste em estabelecer um orçamento (ver abaixo) e descrever apenas as principais linhas metodológicas dos TdR, para depois seleccionar a equipa que propõe a metodologia mais promissora. A seguir, os que seleccionam a equipa terão necessariamente de ter a capacidade de avaliar a qualidade metodológica de uma proposta.

vii. Calendarização

A calendarização da avaliação deverá ser definida tendo em conta diversas restrições, sobretudo as que estão relacionadas com a calendarização das tomadas de decisão, e a sua possível utilização. Também é necessário integrar os principais prazos, gerados pelos procedimentos inerentes aos convites à apresentação de propostas e às fases de recolha inicial de dados. Recomenda-se que as dimensões gerais do exercício sejam definidas nos TdR, e que se reserve o primeiro período, normalmente entre 10-20% da duração da avaliação geral, para o planeamento detalhado do trabalho. Esta fase deverá terminar com um Relatório Inicial, onde se descreve a concepção das actividades, bem como uma calendarização detalhada. Também se recomenda que se faça uma previsão dos diferentes resultados pretendidos do exercício, que devem incluir um referência específica à submissão de uma versão provisória do relatório final, permitindo assim que haja tempo suficiente para sugestões de alterações e correcções antes do término do estudo.

viii. Orçamento indicativo

Recomenda-se a apresentação de um orçamento indicativo, permitindo posteriormente que os candidatos ao concurso aberto para a realização da avaliação façam sugestões quanto ao que poderão efectivamente realizar face ao orçamento disponível. Este procedimento permitirá uma avaliação das propostas com base no critério da melhor relação preço-qualidade (*value-for-money*), para além de proporcionar aos requerentes da avaliação um maior controlo sobre os gastos. Uma alternativa a esta abordagem descendente consiste em permitir aos proponentes que submetam as suas próprias estimativas, com base nas tarefas que consideram ser necessárias. Em geral, os candidatos devem ter orientações claras para repartir os custos em categorias básicas, incluindo, por exemplo, a recolha de dados, a preparação de relatórios, o trabalho de campo, etc.

ix. Qualificações necessárias da equipa

Os TdR devem especificar um conjunto de requisitos em relação à equipa de avaliação, incluindo: as capacidades metodológicas necessárias, a experiência anterior em trabalhos de avaliação semelhantes, o conhecimento do contexto regional e institucional, a experiência profissional e a especialização na área, bem como a capacidade de gerir e concluir uma avaliação de forma atempada.

a) Independência da equipa de avaliação

Já foi referida a importância da independência em termos de credibilidade, factor que pode ser reforçado ao confiar a avaliação a uma equipa externa. Tornam-se igualmente relevantes os seguintes aspectos:

- Implementar procedimentos administrativos que apoiam a independência dos avaliadores seleccionados;
- Solicitar uma confirmação da inexistência de conflitos de interesse relativamente à potencial equipa.

Estas condições deverão encontrar-se especificadas nos TdR.

Ao mesmo tempo, também se deve descrever a forma como os avaliadores poderão ter acesso a colaboradores essenciais do programa em causa e à sua administração, bem como a informações necessárias para desempenhar o seu trabalho. (As questões relativamente à independência dos avaliadores discutem-se mais pormenorizadamente de seguida).

b) Perfil da equipa de avaliação

No âmbito dos Fundos Estruturais europeus, existem cada vez mais organizações presentes no mercado da avaliação, incluindo empresas de consultoria locais, nacionais ou internacionais. O sector comercial é responsável pela maioria do mercado, embora os centros de investigação universitários também representem uma proporção significativa.

Optar por uma empresa de consultoria ou pelo departamento de uma universidade pode ter implicações em termos de abordagem e, por conseguinte, em termos dos resultados da avaliação. Os académicos têm a vantagem de serem vistos como independentes e altamente credíveis, devido aos seus próprios requisitos institucionais e profissionais. Por outro lado, as empresas privadas têm em muitos casos mais disponibilidade em termos de tempo, e estão mais preocupadas em satisfazer as expectativas do requerente.

A escolha global não deverá depender tanto das origens institucionais da equipa de avaliação, mas sim das competências pretendidas, i.e., da sua experiência, das suas capacidades e dos seus conhecimentos prévios. Também deverá solicitar-se aos proponentes que indiquem de que forma as diferentes experiências, capacidades e conhecimentos numa equipa serão integradas e incentivadas de modo a funcionar em conjunto.

x. Estrutura da proposta

Para facilitar a adjudicação, e de modo a proporcionar orientação aos potenciais candidatos, os TdR deve especificar de que modo a proposta deverá ser estruturada, indicando possivelmente o número máximo de páginas para cada uma das secções do documento.

xi. Regras de submissão e critérios de adjudicação

A proposta deverá indicar: o prazo, a forma de envio (correio, fax, e-mail), período de validade da proposta, etc. Também deverá indicar os critérios de apreciação das propostas. Os TdR devem especificar, por exemplo, em pontos percentuais a importância relativa atribuída:

- à qualidade da abordagem metodológica
- às qualificações e experiência prévia da equipa
- aos custos

É obviamente importante que estes critérios de selecção sejam aplicados sistematicamente, após a recepção das propostas.

2.2. A implementação e a gestão da avaliação

- **Seleccionar os avaliadores adequados**
- **Gerir o processo de avaliação**
- **Gerir o processo de comunicação da avaliação**
- **Gerir a garantia e o controlo de qualidade**

Seleccionar os avaliadores adequados

O processo de selecção de avaliadores tem de ser transparente. A preparação cuidadosa de uma versão inicial dos Termos de Referência (TdR), juntamente com o recurso a uma comissão de selecção formal, é a melhor forma possível de cumprir com este objectivo. Este deverá incluir não só representantes das entidades responsáveis pela avaliação, mas também um perito independente e, sempre que possível, representantes dos utilizadores concretos e potenciais da avaliação.

Os membros da comissão de selecção deverão chegar às suas próprias conclusões em termos de candidaturas, face aos critérios estabelecidos nos TdR. As conclusões deverão, então, ser combinadas. Os critérios incluem, normalmente: a qualidade do método de abordagem proposto, a qualidade e a experiência da equipa de avaliação, e os custos.

Avaliar a qualidade do método proposto

A adequabilidade da abordagem e dos métodos propostos para responder às questões colocadas nos TdR deverão ser aspectos centrais na selecção de candidaturas. A comissão de selecção poderá assegurar este princípio, verificando cada uma das candidaturas de acordo com os pontos referidos na Caixa *Avaliar a qualidade do método de abordagem de uma proposta*.

Caixa: Avaliar a qualidade do método de abordagem de uma proposta

Para cada questão de avaliação

	Questão 1	Questão 2
<i>A proposta da equipa candidata:</i>			
<i>Inclui a recolha de informações suficientemente relevantes?</i>	++	+	
<i>Baseia-se em técnicas analíticas precisas?</i>	-	+	
<i>Tem capacidade para esclarecer os critérios de avaliação de forma imparcial?</i>	+	+/-	
<i>Poderá conduzir a conclusões credíveis?</i>	+	+	
<i>Compreendeu-se realmente a importância das questões?</i>		++	

Acima de tudo, deverá ter-se em conta que os processos de avaliação de qualidade do método proposto são qualitativos, e não quantitativos. Estes são constituídos por

avaliações que devem ser realizadas por pessoas experientes. Muitos dos itens sujeitos a avaliação também são qualitativos. Por exemplo, o tamanho da amostragem de um inquérito e/ou de uma série de estudos de caso poderá ser menos importante do que a qualidade do processo através do qual se extrai a amostra ou se identifica os estudos de caso.

Avaliar as qualificações e experiência prévia da equipa

As qualificações e a experiência prévia da equipa são sempre importantes, sobretudo se os métodos propostos forem experimentais ou não cumprirem na íntegra as especificações dos TdR. Poderia argumentar-se que a avaliação é comum e/ou os TdR são muito precisos em termos daquilo que deverá fazer-se; a qualidade dos recursos humanos e o preço são os aspectos que interessam realmente. Contudo, embora esta afirmação possa ser considerada verdadeira, o inverso não se aplica. Sempre que o processo de candidatura pede aos candidatos que proponham a metodologia que considerem adequada para a execução do exercício, o maior perigo consiste no facto de se prestar demasiada atenção à originalidade da abordagem, e não prestar a atenção suficiente à capacidade dos candidatos para cumprir realmente os resultados prometidos.

As capacidades da equipa deverão corresponder à metodologia proposta, de modo a evitar problemas que possam surgir no decurso da implementação da avaliação. Contudo, existe o perigo de discriminação da entrada de novos agentes, e, por conseguinte, poderá dificultar a criação e/ou manutenção de mercados concorrentes.

Uma forma útil de considerar a proposta de uma equipa é pedir exemplos anteriores do seu trabalho. Adicionalmente, este processo poderá ser apoiado pelo pedido de referências de clientes de avaliação anteriores, ou seja, a possibilidade de consultar pessoas concretas.

Finalmente, é sempre adequado prestar atenção, não só à presença de recursos humanos altamente qualificados na equipa, mas também ao momento em que estarão preparados para se dedicar a este exercício. Uma vez que as avaliações são extremamente morosas, o trabalho de campo não poderá ser desempenhado na íntegra pelas pessoas mais qualificadas. O tempo atribuído às pessoas mais experientes deverá ser suficiente para permitir a supervisão das pessoas que fazem o trabalho de campo. O facto de a equipa proposta ter trabalhado em conjunto, com êxito, é pertinente.

Avaliar o preço

A avaliação do preço proposto pelos serviços constitui um aspecto importante do processo de selecção, mas não deverá ser sobrevalorizado. Regra geral, em quaisquer avaliações que não sejam completamente de rotina, o critério financeiro não deverá exceder 20-25% da avaliação global.

Um segundo aspecto digno de nota é que deve ter-se em consideração, não só o preço total, como também o custo unitário por dia das diferentes categorias de colaboradores contratados. Por exemplo, se 80% do preço total for despendido em recursos humanos júnior a um *per diem*, digamos, reduzido, então as vantagens deste facto poderão ser comparadas com um cenário em que 50% do trabalho é executado por investigadores mais qualificados/experientes, que trabalham pelo dobro deste *per diem*.

Em alguns países, para evitar um decréscimo da qualidade, os custos são considerados, não em termos absolutos, mas relativamente à média proposta pelas equipas candidatas ao trabalho. Neste caso, se uma oferta for excepcionalmente baixa, o candidato poderá ter que justificar o motivo deste custo reduzido.

Gerir o processo de avaliação

Depois de iniciado o estudo de avaliação, existe a tentação de a organização requerente manter um contacto estreito com a equipa de avaliação. Esta perspectiva baseia-se no facto de uma abordagem de maior afastamento ajudar a assegurar a independência da equipa de avaliação.

A independência da equipa depende, de facto, de um conjunto de factores muito mais complexo do que a mera limitação de contactos com o cliente. A maior garantia de independência é constituída pela base científica e profissional da equipa seleccionada. A existência de um grande mercado, o aparecimento de normas profissionais e éticas, e a criação de uma comunidade envolvida na avaliação constituem dimensões estruturais importantes que, em última instância, contribuem para a independência.

Ao gerir avaliações, os requerentes e os gestores de programas deverão estar conscientes da existência de uma série de factores contínuos que poderão minar a independência dos avaliadores.

Factores que influenciam a independência dos avaliadores

Todo o trabalho de avaliação requer um determinado grau de independência entre o avaliador e o objecto de avaliação. Mesmo em abordagens de auto-avaliação, as pessoas envolvidas na implementação das intervenções necessitam de obter um determinado grau de distanciamento e independência, quer o processo seja ou não apoiado por avaliadores externos. Normalmente, o aumento do grau de independência do avaliador face ao objecto de avaliação irá aumentar a credibilidade dos resultados de avaliação. Em quaisquer circunstâncias, as possibilidades de conflito de interesse têm que ser minimizadas e, sempre que possível, eliminadas. Por vezes, esta medida baseia-se em declarações formais dos avaliadores e de potenciais avaliadores relativamente à ausência destes conflitos.

Porém, os avaliadores raramente são totalmente independentes dos objectos de avaliação, e a avaliação nunca é isenta de valor. Os avaliadores estão sujeitos a toda uma série de influências. De facto, o compromisso do avaliador para com os objectivos da intervenção em apreço poderá aumentar a qualidade dos resultados da avaliação, bem como a probabilidade de os resultados conduzirem a um maior desenvolvimento socioeconómico.

Os factores que influenciam a independência do avaliador são diversos e nem todos eles são fáceis de evitar. Por vezes, as influências externas da avaliação em curso podem trazer vantagens:

- Os avaliadores tendem a ser benevolentes com os objectivos de desenvolvimento socioeconómico das intervenções subjacentes. Poderão, provavelmente, residir no território em causa ou ter sido seleccionados, em parte, pela sua empatia com o grupo-alvo da intervenção. Muitas vezes, os avaliadores são seleccionados pelos seus conhecimentos aprofundados na matéria ou área política e contactos, bem como pela sua experiência de avaliação.
- Normalmente, os avaliadores pretendem ser ouvidos e ter influência. A actividade de avaliação é, normalmente, sumativa e formativa, em simultâneo, e o equilíbrio entre estes dois factores poderá mudar no decorrer da implementação do trabalho de avaliação. Se os requerentes de um trabalho de avaliação se defrontarem com uma nova escolha política, poderão entender conveniente alterar os TdR, ou

inclusivamente pedir as impressões iniciais do avaliador. Os resultados precoces da avaliação poderão levantar questões sérias, imprevistas, e identificar a necessidade de alterações radicais na intervenção proposta ou já em fase de implementação.

- A interpretação dos dados concretos depende do entendimento da forma como o mundo funciona. O avaliador terá as suas próprias perspectivas *a priori* relativamente às consequências prováveis dos diferentes tipos de intervenção, com base numa combinação de experiência directa, formação disciplinar educativa e valores pessoais. Nos relatórios finais e no decorrer da justificação de um método proposto, estas perspectivas, experiências e valores pré-existentes devem ser explicitados.
- O avaliador é, normalmente, remunerado. Na maior parte dos casos, os requerentes da avaliação são responsáveis, parcial ou integralmente, pelas intervenções. Nalguns casos, a avaliação constitui um requisito de terceiros, e poderá haver uma tendência de conluio entre os parceiros e os avaliadores. Uma avaliação de sucesso requer fortes medidas de confiança, que poderão ser reforçadas pelos tipos de normas e códigos de ética impostos aos avaliadores, descritos mais adiante nesta parte do GUIA, e uma intenção, por parte dos requerentes do trabalho, de observar as conclusões da avaliação e as perspectivas do avaliador.
- A avaliação do desenvolvimento socioeconómico nunca decorre num ambiente politicamente neutro. Os territórios ou os grupos-alvo prioritários poderão pretender manter a posição e o êxito das intervenções anteriores poderá ser um factor a considerar no futuro acesso aos recursos. Poderá haver rivalidades entre os responsáveis por diferentes intervenções. Os requerentes do trabalho de avaliação estão, muitas vezes, sob pressão para apresentar as conclusões do programa e mostrar resultados.

As circunstâncias e os papéis variáveis em que decorre a avaliação afecta o grau de independência que se poderá alcançar. Sempre que o avaliador oferecer sobretudo contributos de investigação e recolher dados, poderá alcançar-se um elevado grau de independência. Contudo, mesmo nestas circunstâncias, a escolha das questões colocadas e o método de questionamento poderão condicionar a independência das conclusões. Sempre que o trabalho de avaliação for executado sobretudo para efeitos de escrutínio, inspecção ou pré-auditoria, a independência do avaliador tende a ser superior. Sempre que os avaliadores trabalharem em estreita cooperação com as pessoas responsáveis pela preparação das intervenções, o papel do avaliador é caracterizado como sendo o de um amigo essencial. Este processo decorre, muitas vezes, ainda que não exclusivamente, numa fase ex-ante ou de um estudo de viabilidade. Estas avaliações são essencialmente estruturas de apoio, mas destinam-se a destacar dificuldades e fraquezas nas análises subjacentes a potenciais intervenções. Sempre que a intervenção for de natureza experimental ou executada com base num projecto-piloto, poderá ser difícil alcançar uma verdadeira independência. Neste caso, a intervenção é relativamente pequena e complexa, mas envolve diferentes partes que, provavelmente, trabalham em conjunto pela primeira vez, e o avaliador poderá ser um agente animador e catalisador de consenso, enquanto avaliador imparcial. Em muitos casos, o trabalho de avaliação implica uma combinação de processos de revisão e de estudos de caso, que podem ser utilizados para estruturar os argumentos. A selecção de casos e de exemplos concretos poderá restringir uma verdadeira imparcialidade. Sempre que um avaliador esteja efectivamente a desempenhar um papel de ligação e *feedback* auxiliar entre os actores a diferentes níveis governativos, é necessário que se tenha em particular atenção as normas profissionais e éticas em vigor, tanto entre os avaliadores, como entre os parceiros.

Interacção entre o requerente, os parceiros e o avaliador

Existe uma série de razões pelas quais a gestão de uma avaliação requer uma interacção contínua e significativa entre todos os parceiros envolvidos (incluindo a própria equipa de avaliação):

Recomenda-se uma primeira fase, durante a qual a equipa testa e aperfeiçoa a justificação declarada da avaliação, através da consulta a partes potencialmente interessadas, principalmente em todos os exercícios de avaliação intercalares, ex-post ou exercícios de avaliação temática.

Uma fase inicial ou de arranque destina-se, normalmente, a especificar os métodos e o plano de trabalho de forma mais pormenorizada do que era possível na fase de proposta. A equipa de avaliação só poderá, normalmente, propor uma abordagem operacional detalhada após um primeiro estudo do campo e uma análise dos dados disponíveis. Esta parte fundamental de concepção da avaliação deverá ser partilhada e acordada com o requerente e as restantes partes interessadas envolvidas. A Caixa *Avaliação ex-ante na República Checa* constitui um exemplo.

Caixa: Avaliação ex-ante na República Checa

No âmbito da avaliação ex-ante do Plano de Desenvolvimento Nacional, na República Checa, constatou-se que as recomendações da equipa de avaliação seriam mais facilmente aceites quando sugeridas no decorrer das fases de preparação inicial dos documentos do programa. O trabalho dos avaliadores possui funções de desenvolvimento muitas vezes relacionadas com a assistência técnica, e implica a percepção da avaliação como sendo mais útil após o estabelecimento em concreto de uma parceria.

Mesmo que o exercício de avaliação seja linear, os contextos de política externa mudam, muitas vezes, rapidamente. Por conseguinte, é útil assegurar uma ligação eficaz, não só com os membros da Comissão Directiva, mas também com os responsáveis políticos. A oportunidade de envolver – sempre que possível e, inclusivamente de forma indirecta – o nível de direcção estratégico constitui, conseqüentemente, uma outra razão pela qual o processo deverá ser interactivo.

Um mecanismo simples consiste em especificar a frequência das reuniões da Comissão Directiva, mesmo na fase de elaboração dos TdR. É normal a realização de, no mínimo, duas reuniões na fase inicial e para aprovar um primeiro rascunho do relatório final.

É importante deixar decorrer um determinado período de tempo entre a selecção da equipa de avaliação e o início do trabalho. Sobretudo nos casos em que a selecção implica um convite à apresentação de propostas, não é realista esperar que a equipa vencedora possa começar a trabalhar no dia seguinte à tomada de decisão. Face às incertezas com que se debate a escolha do contratante, a maioria dos candidatos irá necessitar de várias semanas para planear e constituir a equipa que irá, efectivamente, fazer o trabalho. Existem, pelo menos, duas formas de garantir esta flexibilidade:

- atrasar a efectiva assinatura do contrato – e, por conseguinte, a data de início do trabalho;
- deixar um período adequado para o Relatório Inicial.

Papel do Relatório Inicial

O Relatório Inicial é um documento que estabelece:

- a identificação das principais partes interessadas;
- as questões de avaliação mais importantes (elaboradas e possivelmente reafirmadas);
- os métodos a utilizar;
- um plano de trabalho pormenorizado, com a divisão de tarefas entre os diferentes membros da equipa;
- o calendário de trabalho (finalizado), incluindo as diferentes etapas; e
- os resultados e conclusões intercalares e finais.

Este documento deve ser discutido e acordado com a Comissão Directiva na primeira reunião após o início do trabalho, e, posteriormente, actualizado de forma contínua. Representa, ao longo de todo o exercício, o principal ponto de referência do processo de garantia de qualidade (ver abaixo), uma vez que especifica de forma pormenorizada o que poderá esperar-se do exercício, quais os momentos de execução das diferentes actividades, e qual o processo de produção dos relatórios de avaliação.

Relatórios intercalar e final

Em algumas avaliações, sobretudo nas mais prolongadas, existe um relatório intercalar, bem como um relatório inicial. Isto permite a partilha de primeiras impressões, e proporciona uma oportunidade de orientar as fases de avaliação subsequentes, sempre que as conclusões iniciais determinarem essa necessidade. Este aspecto é particularmente importante sempre que se prevê que as avaliações proporcionem informações ou conselhos para a gestão do programa. Nos fundos estruturais, esta fase intercalar inclui-se, muitas vezes, na avaliação intercalar. Destaca-se, assim, a vantagem da ligação das avaliações intercalar e ex-post. Os processos de avaliação contínua, que registam as alterações ao longo do tempo, possuem, normalmente, uma série de relatórios intercalares que proporcionam *feedback* aos gestores de programa e aos decisores políticos.

As versões em rascunho dos relatórios finais poderão exercer uma função de orientação semelhante, sobretudo se exigidas com a antecedência suficiente. Porém, estas orientam sobretudo o relatório, e não o programa, como seria o caso das avaliações intercalares e contínuas. Deverá, ainda, realçar-se que, em prol da independência, as comissões de direcção que recebem as versões preliminares dos relatórios finais deverão concentrar-se em questões de precisão e conformidade com as expectativas, em vez de tentar influenciar ou adivinhar as conclusões de avaliação.

A Comissão Directiva

Como vimos, a existência de um organismo como uma Comissão Directiva ou uma Comissão de Avaliação constitui uma parte importante do processo de gestão da avaliação dos programas de desenvolvimento socioeconómico.

A experiência resultante dos Fundos Estruturais demonstra a vantagem do envolvimento das partes interessadas mais importantes, e, sobretudo, dos parceiros-chave e institucionais, actores cuja cooperação é necessária para abordar os principais resultados do programa. As vantagens de uma Comissão Directiva abrangente são apresentadas na Caixa *As vantagens de uma Comissão Directiva inclusiva*.

Caixa: As vantagens de uma Comissão Directiva inclusiva

A criação de uma Comissão Directiva de avaliação constituída pelas diferentes

partes interessadas de um programa permite assegurar:

- uma melhor aceitação da avaliação pelos avaliados, criando relações de confiança;
- um acesso mais fácil à informação, e uma melhor compreensão dos factos e dos eventos ocorridos durante a execução do programa;
- oportunidades de aprendizagem e utilização de processos entre as partes interessadas, resultantes da interacção com a sua Comissão Directiva;
- interpretações e recomendações que considerem todos os pontos de vista importantes;
- a disseminação de conclusões e consideração de recomendações de forma mais rápida e informal; e
- uma maior probabilidade de as recomendações e conclusões conduzirem à acção e posterior acompanhamento.

Regra geral, a Comissão Directiva deverá incluir quatro tipos de pessoas:

- A *gestão estratégica* do programa ou intervenção, ou seja, as entidades financiadoras, a vertente política da administração e, sempre que adequado, os diferentes níveis de governo. Uma abordagem multi-nível, com vista a envolver a gestão estratégica na Comissão Directiva, é muito importante, uma vez que os programas são cada vez mais complexos, considerando aspectos com diferentes dimensões territoriais;
- A *gestão operacional* do programa, isto é, todos aqueles cujas actividades são avaliadas pelo estudo, embora, para garantir a imparcialidade da Comissão Directiva, a gestão operacional seja, normalmente, assegurada por gestores sénior, um pouco afastados da primeira linha de gestão quotidiana. Porém, constitui uma tarefa importante dos membros da Comissão garantir que nenhum membro, incluindo os gestores operacionais, procura influenciar os resultados da avaliação ou ignorar quaisquer elementos de prova;
- Os *parceiros sociais*, isto é, os representantes dos principais interesses afectados pelo programa. Entre estes, poderão incluir-se, não só as associações comerciais, os sindicatos e as associações de interesse económico, mas também os organismos institucionais e/ou sociais responsáveis por aspectos horizontais específicos, tais como o ambiente, a igualdade de oportunidades, o turismo e a protecção dos consumidores, etc.
- Os *peritos*, ou seja, pessoas com conhecimentos substanciais ou metodológicos, que possam ser úteis para definir as questões de avaliação ou interpretar os resultados. A presença de peritos independentes na Comissão Directiva poderá ser extremamente importante para proporcionar contributos considerados úteis para a equipa de avaliação, bem como para inaugurar o debate conducente a lições mais genéricas a retirar do exercício.

O papel principal da Comissão Directiva consiste em assegurar uma avaliação útil e de alta qualidade. Para isso, deve facilitar o trabalho dos avaliadores, por exemplo, dando acesso a informações e contactos e elaborando questões de avaliação e questões-chave, relativamente às quais considera que deve ser informada. A Comissão Directiva não deverá tentar influenciar os avaliadores, no sentido de omitir determinados dados ou chegar a conclusões da sua preferência, que não sejam fundamentadas pelos dados da avaliação. A Comissão Directiva também deve supervisionar o processo de comunicação dos resultados da avaliação.

Gerir o processo de comunicação da avaliação

A comunicação constitui uma parte importante do processo de avaliação. É preferível gerir a comunicação como um processo contínuo: uma oportunidade de diálogo e acumulação de conhecimentos, em vez de apostar todos os esforços de comunicação num grande exercício de disseminação único, após a elaboração de um relatório final. A comunicação deverá, por conseguinte, incluir:

- Uma melhoria do conhecimento do trabalho de avaliação a ser implementado;
- Transmissão de feedback relativamente aos resultados intercalares;
- Circulação e gestão de feedback relativamente aos relatórios e conclusões provisórios (por ex., instrumentos de recolha de dados);
- Comunicação dos resultados e conclusões da avaliação.

Melhorar a percepção e compreensão do trabalho de avaliação a ser implementado

Quando a equipa de avaliação estiver definida, é útil transmitir informações sobre a calendarização e metodologia às partes interessadas. O período de iniciação deve utilizar-se como uma oportunidade de explicar a abordagem planeada e de trocar opiniões relativamente à utilidade das questões de avaliação e ao êxito provável das propostas apresentadas. Além das informações formais transmitidas às partes interessadas, talvez através da Comissão Directiva, a divulgação de informações gerais junto do público e dos beneficiários, por exemplo em forma de comunicados à imprensa ou informações acerca de websites, também poderá representar uma forma útil de preparar o terreno para a avaliação.

Assegurar feedback sobre os resultados intercalares

A comunicação de resultados intercalares coloca grandes desafios. Por um lado, é muito provável que as partes interessadas queiram conhecer os resultados iniciais, sobretudo se apontarem os resultados finais como essenciais. Ao mesmo tempo, a equipa de avaliação poderá hesitar em inferir grandes conclusões, e apresentar algum nervosismo relativamente ao peso das provas de base nas suas observações. Poderão, também, - embora não devam - perspectivar a elaboração de resultados intercalares como sendo pouco mais do que uma necessidade burocrática (não é pouco frequente serem necessários relatórios intercalares para activarem pagamentos intercalares). O ideal é prestar atenção ao relatório inicial relativamente ao âmbito e conteúdo prováveis dos resultados intercalares, bem como à metodologia e abrangência da sua circulação. Poderá não ser adequado não existir um relatório intercalar. Este processo poderá evitar críticas suscitadas por resultados da avaliação parciais e não fundamentados. Na melhor das hipóteses, os resultados intercalares poderão proporcionar um feedback útil relativamente à metodologia e implementação (por exemplo, propor alterações ao processo) e ajudar a aumentar o compromisso das partes interessadas e de todos os intervenientes, tanto nos programas, como na sua avaliação. Os resultados que possam representar questões e apontadores poderão satisfazer um objectivo de aprendizagem importante, em termos de avaliação.

Circulação e gestão do feedback sobre os relatórios e resultados preliminares

A elaboração do relatório final provisório constitui, frequentemente, uma fase difícil, quer para os avaliadores, quer para as partes interessadas. Aquilo que constituía, anteriormente, uma previsão abstracta dos resultados ou das realizações torna-se agora concreto e, por vezes, um elemento intimidador e desanimador. As partes interessadas, sobretudo as que têm responsabilidades na gestão do programa, poderão ser tentadas a desacreditar os resultados que não sejam do seu agrado. Os avaliadores, por seu lado, poderão criar argumentos com base num número limitado de dados, ou ser insensíveis à importância política do que apresentam, sobretudo na fase provisória. Produzir um

relatório final considerado aceitável pela equipa de avaliação e pela organização responsável pela encomenda, e respeitado pelas partes interessadas envolvidas no processo, poderá representar um enorme desafio e exigir uma grande disponibilidade de tempo. As sugestões que se seguem poderão facilitar o processo:

- A estrutura do relatório deverá ser acordada o mais cedo possível;
- A Comissão Directiva deverá constituir o principal fórum de discussão do documento provisório;
- A entidade contratante deverá evitar a tentação de influenciar demasiado a formulação das conclusões e recomendações; antes, deverá desafiar a equipa de avaliação, no sentido de justificar as suas conclusões e recomendações, com base nos dados apresentados;
- Deverá dedicar-se tempo suficiente aos comentários escritos;
- A entidade de gestão deverá assumir a responsabilidade pela circulação do relatório e pela compilação do feedback.

Comunicar os resultados da avaliação

O trabalho de avaliação não possui quaisquer consequências se os resultados não forem transmitidos. A principal forma de comunicação é constituída por um relatório escrito. Embora a adequabilidade dos meios de comunicação específicos varie, existem diversas lições de boas práticas:

- O relatório escrito deverá ser redigido de forma clara e concisa. Normalmente, rondará as cem páginas, incluindo um sumário. Os dados de avaliação pormenorizados, tais como estudos de caso e análise quantitativa, devem ser apresentados, em anexos, ou separadamente.
- O relatório deve incluir um sumário de 5-10 páginas, escrito num estilo adequado aos decisores políticos.
- As relações entre as conclusões e a análise dos dados devem ser claras.
- A redacção provisória do relatório deve indicar a base de sustentação das observações efectuadas: os dados de avaliação ou uma combinação de dados e de opiniões dos avaliadores.
- O relatório deve incluir uma descrição e avaliação do método utilizado que seja suficientemente pormenorizada e auto-crítica, de modo a permitir ao leitor avaliar o peso dos dados conducentes às conclusões.
- Devem utilizar-se quadros e diagramas sempre que estes permitam melhorar a apresentação dos resultados.
- Devem referir-se exemplos de boas práticas de intervenções para ilustrar os argumentos apresentados, mas os relatórios da avaliação não devem assumir o papel de orientações de boas práticas. A pressão muitas vezes exercida sobre os avaliadores, no sentido de produzir boas histórias noticiosas é, frequentemente, contra-productiva: estes resultados são vistos com desconfiança, tanto pelo público, como pelos decisores políticos.
- As recomendações devem ser claras nas acções de acompanhamento necessárias.

Canais de comunicação dos resultados de avaliação e alcance dos respectivos utilizadores

Os responsáveis pela encomenda e pela realização da avaliação devem assegurar a transmissão e utilização dos resultados. É necessário identificar cuidadosamente os potenciais utilizadores, dos decisores políticos ao público em geral, passando pelos beneficiários, e seleccionar os canais de comunicação mais adequados.

Os relatórios de avaliação são, normalmente, publicados, e cada vez mais na internet. Os relatórios escritos também deverão incluir resultados mais populares, para que possam ser adoptados pelos meios de comunicação social. Muitos programas produzem as suas próprias *newsletters*, e estas proporcionam mais uma oportunidade de disseminação. As apresentações orais à Comissão Directiva e a outras partes interessadas (por exemplo, em *workshops* especializados) também são úteis.

Gerir a garantia e o controlo de qualidade

Ver também *Controlo de Qualidade*.

(http://www.pm.gov.uk/files/pdf/Quality_framework.pdf).

Avaliar a qualidade de uma avaliação constitui parte integrante e fundamental do processo de avaliação. De facto, uma avaliação que não satisfaça alguns padrões de qualidade mínimos apresenta uma grande probabilidade de enganar os decisores políticos e os gestores de programa.

Contudo, avaliar a qualidade de avaliação é um processo complexo e difícil. As avaliações realizadas no contexto dos programas e políticas de desenvolvimento socioeconómico são muito diferentes entre si, de modo a permitir a existência de algumas regras simples, que garantam a qualidade em toda a sua dimensão.

De um modo geral, poderá dizer-se que a qualidade da avaliação como um todo depende da existência de três aspectos distintos, mas inter-relacionados:

- a qualidade da fase de concepção e planeamento, incluindo a encomenda da avaliação;
- a qualidade da implementação da própria avaliação;
- a qualidade do sistema de monitorização e dos dados disponíveis.

Estes aspectos encontram-se inter-relacionados, na medida em que um fraco desempenho do avaliador pode, muito provavelmente, ser decorrente da fraca qualidade dos dados e/ou das falhas da fase de planeamento e concepção. Infelizmente, as pessoas envolvidas nestes três conjuntos de actividades são diferentes, e, muitas vezes, os seus objectivos, bem como os seus critérios de qualidade, também são diferentes. Por exemplo, o sistema de monitorização concebido para a gestão do programa no dia-a-dia não produz necessariamente os dados necessários para uma avaliação dos impactos.

Além disso, estes aspectos podem ser vistos de duas perspectivas diferentes.

Em primeiro lugar, poderá considerar-se a qualidade como *uma característica do processo* através do qual se executam as actividades de avaliação. A avaliação da qualidade poderá incluir a forma como a entidade requerente estrutura a decisão para avançar para uma avaliação, como define o seu âmbito e os recursos disponíveis. Esta análise poderá permitir compreender se os procedimentos adoptados foram adequados à alocação das diferentes responsabilidades, se a contribuição das diversas partes interessadas foi tida em consideração, etc. O mesmo se aplica à execução da avaliação. Poderemos focar-nos na forma de gestão da equipa, bem como na sua interacção com o requerente e com os avaliadores, nas verificações aplicadas de modo a assegurar o tratamento adequado dos dados recolhidos, etc. Também é possível avaliar a organização do processo de monitorização.

Em segundo lugar, poderá considerar-se a qualidade como *uma característica dos produtos* do processo de avaliação. Assim, poderá analisar-se os TdR de acordo com os critérios já descritos. A qualidade dos relatórios de avaliação final e intercalar pode ser

avaliada, no sentido de verificar se eles satisfazem alguns critérios básicos de boas práticas profissionais, e se os dados são quantitativamente suficientes, e suficientemente fidedignos para garantir conclusões sólidas.

Teoricamente, os dois aspectos (o processo e o produto) estão relacionados: um bom processo deverá dar origem a um bom produto, e vice-versa, no sentido em que um bom produto deverá ser o resultado de um processo de produção suficientemente bom.

A Colecção MEANS (1999) referia:

“Não existe um sistema de certificação profissional, em parte alguma do mundo, que institucionalize as regras e os critérios de qualidade. Propõem-se grelhas de critérios algo díspares, com base nos modelos de avaliação elaborados por diversos autores, mas não existe consenso nesta matéria. Além disso, a natureza dos critérios referidos nem sempre é apresentada de forma clara, sempre que se tenta colocá-los em prática.”

Desde então, contudo, muitas coisas mudaram. Especificamente, conforme ilustrado na Caixa Normas, Directrizes e Códigos de Ética, cada vez mais é prática comum definir normas de boas práticas na avaliação. Estas foram elaboradas por organismos internacionais (tais como a OCDE), Governos Nacionais (por exemplo, o Departamento Italiano de Economia e Finanças) ou associações profissionais, tais como associações nacionais de avaliação. Muitos deles decorrem de esforços anteriores realizados nos Estados Unidos, pela American Evaluation Association (AEA): Guiding Principles for Evaluators (1992) e o Joint Committee on Standards for Educational Evaluation: Program Evaluation Standards (1994).

Caixa: Normas, Directrizes e Códigos de Ética

- USA - Program Evaluation Standards (1994) Joint Committee on Standards for Educational Evaluation, Program Evaluation Standards
- Deutsche Gesellschaft fr Evaluation (DeGEval): Standards fr Evaluation (2001)
- Société canadienne d'évaluation (SC) Lignes directrices en matière d'éthique / Guideline for Ethical Conduct
- Switzerland: SEVAL Evaluation Standards
- The African Evaluation Guidelines 2000
- American Evaluation Association (AEA), Guiding Principles for Evaluators
- Australasian Evaluation Society (AES), Guidelines for the Ethical Conduct of Evaluations
- UK Guidelines for good practice
- PUMA Best Practice Guidelines
- Italy Treasury Guidelines

A Caixa Normas, Directrizes e Códigos de Ética apresenta uma síntese de um conjunto de normas e códigos de avaliação actuais, correspondentes a uma série de categorias. A maioria – sobretudo tratando-se dos princípios decorrentes das Normas Conjuntas da AEA, como por exemplo as sociedades de avaliação alemãs (DeGEval) e as *African Evaluation Guidelines* – está orientada, principalmente, para a execução técnica da avaliação pelos avaliadores; referem-se, por exemplo, à forma de recolha dos dados e à apresentação das conclusões. (Uma distinção entre directrizes e normas mais rigorosas e

ambiciosas seria também instrutiva.) Uma outra categoria (representada, por exemplo, pelas normas ou as orientações das associações de avaliação canadianas, australasiáticas e, em certa medida, britânicas) incide nos códigos éticos de boas práticas em detrimento da execução técnica da avaliação. Mas, também se centra mais na ética dos avaliadores do que em outros intervenientes. Mais recentemente, surgiu uma nova categoria de orientações, mais direccionada para as administrações e para os requerentes da avaliação, do que para os avaliadores. Podemos encontrar exemplos disso na OCDE (directrizes PUMA e DAC), e, mais recentemente, na Comissão Europeia.

Apesar deste conjunto crescente de directrizes, normas e códigos referentes à qualidade da avaliação, não existe actualmente uma orientação comum, universalmente aceite.

Embora não haja ainda um consenso relativamente a todos os componentes de um sistema de garantia de qualidade na avaliação, já começamos a verificar um maior enfoque no controlo de qualidade, isto é, para as formas de avaliar a qualidade dos relatórios e das conclusões. Esta mudança foi confirmada por um estudo recente sobre a utilização da avaliação pela Comissão Europeia (Caixa Citação de um estudo da UE sobre a utilização da avaliação).

Caixa: Citação de um estudo da UE sobre a utilização da avaliação

(O estudo...) tende a apoiar o valor de normas inclusivas que abrangem tanto os interesses dos requerentes da avaliação, como os dos avaliadores e cidadãos. As normas de avaliação europeias mais abrangentes (instrumentais e éticas), conforme utilizadas pela *European Evaluation Society* e por diversas associações de avaliação nacionais na Europa, poderão complementar a evolução para normas desenvolvidas pela Comissão Europeia e por alguns Governos nacionais (*The Use of Evaluation in Commission Services*, Outubro de 2002).

A Caixa Critérios de controlo de qualidade e de garantia de qualidade identifica critérios, quer de controlo de qualidade, quer de garantia de qualidade. Ambos são necessários como forma de avaliar os relatórios e as conclusões da avaliação. Normalmente, a pessoa responsável pela gestão da avaliação no organismo requerente deveria assumir a responsabilidade pela aplicação dos critérios de controlo de qualidade. Numa situação ideal, a performance dos critérios de garantia de qualidade deverá ser informada pelas perspectivas dos membros da Comissão Directiva, por outras partes interessadas, pela equipa de avaliação, pelas pessoas responsáveis pela gestão da avaliação, e pelo organismo requerente. A Comissão Directiva deverá estabelecer os critérios o mais cedo possível, na atribuição da avaliação, e encontra-se, normalmente, mais bem posicionada para fazer a avaliação global, na fase de conclusão do trabalho. Porém, no caso da garantia de qualidade, que assenta em critérios processuais, será necessário consultar outras partes interessadas, não necessariamente representadas numa Comissão Directiva. Para efeitos de controlo de qualidade, poderá ser útil consultar peritos ou árbitros externos. Destaque-se que a aplicação de critérios de controlo de qualidade/critérios de conteúdo e garantia de qualidade/critérios de processo se realiza tendo em consideração diferentes objectivos. O controlo de qualidade do conteúdo do relatório oferece algumas garantias de que o trabalho se realizou de forma adequada, e que as suas conclusões são fiáveis. A garantia de qualidade do processo de avaliação irá contribuir para conhecer melhor a gestão da avaliação, e oferece importantes contributos para melhorar a gestão futura da avaliação. A Caixa Critérios de controlo de qualidade e de garantia de qualidade descreve os critérios de controlo de qualidade e de garantia de qualidade.

Caixa: Critérios de controlo de qualidade e de garantia de qualidade

Controlo de qualidade: Critérios de conteúdo para os resultados finais	Garantia de qualidade: Critérios processuais
Satisfação das necessidades descritas nos TdR	Objectivos coerentes e avaliáveis
Relevância da cobertura e do âmbito	Termos de referência bem delineados
Concepção e métodos justificáveis	Processo de selecção de propostas sólido
Utilização de dados fidedignos	Diálogo e feedback eficaz ao longo do processo de avaliação
Análise consolidada	Disponibilidade de informação adequada
Resultados credíveis, relacionados com a análise e com os dados	Boa gestão e coordenação por parte da equipa de avaliação
Conclusões imparciais, não tendenciosas, e com conclusões sólidas	Disseminação eficaz dos relatórios/conclusões junto da Comissão Directiva e gestores políticos/de programa
Relatório claro, com sumários e dados de suporte em anexo	Disseminação eficaz junto das partes interessadas

Os responsáveis pelo processo de controlo de qualidade e de garantia de qualidade variam consoante o contexto institucional. Em programas sectoriais nacionais, poderá constituir uma responsabilidade do governo central, e em programas de desenvolvimento locais a responsabilidade poderá ser dos actores locais. Do mesmo modo, os métodos de aplicação são diversos: por vezes, poderá pedir-se a pessoas-chave para preencher uma grelha e, de seguida, agregá-la; porém, noutras alturas, poderá realizar-se um *workshop* ou uma conferência para garantir conclusões mais equilibradas.

Controlo de qualidade – critérios de conteúdo para os resultados finais

Satisfação de necessidades

A avaliação respondeu às questões incluídas nos TdR de forma satisfatória e o relatório proporciona informações adicionais, possivelmente essenciais para os requerentes? Em concreto:

- Os objectivos do programa evoluíram e foram interpretados e analisados?
- O relatório abrange o programa completo? Em caso negativo, a selecção é justificável em termos de prioridades definidas pelos requerentes nos TdR, e posteriormente acompanhadas?
- A avaliação transmite feedback útil aos gestores do programa?
- Inclui lições sobre êxitos e falhas que poderão ser interessantes para outros programas, regiões ou países?

No caso das avaliações ex-post, é importante verificar se a avaliação consegue atingir um compromisso razoável entre dois requisitos contraditórios: obter informações rapidamente para introduzir no novo ciclo do programa e não tirar conclusões precipitadas antes de se observarem todos os impactos.

Âmbito relevante

O âmbito de uma avaliação deve abranger questões relevantes do ponto de vista dos programas.

No sentido de verificar a relevância do âmbito de uma avaliação, é necessário, em primeiro lugar, observar se as características essenciais do programa foram devidamente descritas, e se os problemas e êxitos na implementação do programa foram devidamente clarificados.

Em segundo lugar, uma vez que os resultados e os impactos do programa têm que ser analisados de forma a avaliar em que medida se alcançou os objectivos do programa, é necessário verificar se os mesmos foram incluídos na avaliação. Também é necessário verificar se a avaliação não sobrestimou outros resultados ou impactos potenciais ou futuros, bem como quaisquer efeitos e resultados inesperados, ainda que significativos.

Finalmente, o âmbito de uma avaliação depende do alvo do programa que se pode definir em termos de áreas geográficas elegíveis ou grupos-alvo não localizados (por exemplo, pessoas em situação de desemprego prolongado). É, portanto, necessário verificar se:

- os limites do âmbito, em termos de áreas ou grupos sociais, se encontram definidos de acordo com a lógica da intervenção;
- o âmbito inclui áreas periféricas ou grupos não elegíveis que são, no entanto, provavelmente afectados pelas intervenções avaliadas;
- em último lugar, se a avaliação considera o programa avaliado isoladamente ou inclui as suas interacções com outros programas Europeus ou nacionais.

Concepção justificável

Este critério relaciona-se com qualidades técnicas da avaliação. As escolhas metodológicas deverão decorrer das questões de avaliação. A avaliação deve, além disso, fazer a melhor utilização possível de análises e investigação existente. É necessário colocar três tipos de questões:

- Os conhecimentos relevantes foram devidamente recolhidos e utilizados?
- A construção do método e a escolha de ferramentas justificam, de facto, de forma adequada a resposta às questões de avaliação?
- As situações de referência escolhidas (contrafactuais ou semelhantes) foram as adequadas para fazer comparações válidas?

Qualquer relatório de avaliação deverá incluir uma descrição do método utilizado e definir claramente as fontes dos dados. Do mesmo modo, deve descrever-se claramente os limites do método e as ferramentas utilizadas. É necessário verificar se:

- o método se encontra descrito de forma suficientemente pormenorizada para avaliar a qualidade;
- a validade dos dados recolhidos e das ferramentas utilizadas se encontra claramente identificada;
- os dados disponíveis correspondem às ferramentas utilizadas.

Uma vez que a análise causal dos efeitos constitui a questão mais importante das avaliações ex-post, o método utilizado para analisar estas relações causais é prioritário neste tipo de avaliação. É necessário verificar se a avaliação analisa de forma adequada as relações de causa e efeito nas questões mais essenciais.

Dados fidedignos

Os avaliadores utilizam dados existentes (dados secundários) do sistema de monitorização e de outras fontes de informação, ou, então, dados primários recolhidos para a avaliação. Neste caso, os métodos utilizados para recolher e processar os dados (escolha e aplicação das ferramentas utilizadas para este efeito) constituem factores muito importantes para a fiabilidade e validade dos resultados.

Para avaliar a fiabilidade dos dados utilizados, é necessário verificar se:

- as fontes de informação disponíveis e a fiabilidade destes dados foram identificadas;
- Os estudos prévios e fontes de informação retiradas do sistema de monitorização foram utilizados de forma óptima;
- as técnicas utilizadas para recolher os dados seleccionados foram exaustivas e adequadas para responder às questões de avaliação;

Independentemente de a recolha de dados ter utilizado técnicas quantitativas ou qualitativas, ou uma combinação das duas, é necessário questionar se:

- a combinação de dados qualitativos e quantitativos é adequada a uma análise válida do fenómeno;
- as "populações" utilizadas na recolha de dados foram correctamente definidas;
- as amostras do estudo ou os casos estudados foram seleccionados em relação aos critérios definidos;
- as principais técnicas de recolha de dados foram implementadas com instrumentos adequados, e de forma a garantir um grau apropriado de fiabilidade e validade dos resultados.

Análise sólida

A análise quantitativa é constituída pela análise dos dados, em forma de quadros ou em qualquer outra forma de análise estatística. A análise qualitativa é constituída pela comparação e interpretação sistemática de fontes de informação, em forma de referências cruzadas. Em ambos os casos, é necessário determinar se os métodos de análise utilizados são relevantes em termos de tipo de dados recolhidos, e se a análise foi realizada de acordo com as instruções dos manuais técnicos correspondentes.

No caso do desenvolvimento socioeconómico, as relações de causa e efeito são complexas e, por conseguinte, constituem um desafio muito específico para a avaliação. É necessário verificar:

- se as relações de causa e efeito subjacentes ao programa são suficientemente explícitas e relevantes, de modo a que se possa focar o objecto da análise,
- e em que medida a análise utiliza técnicas adequadas.

Por isso, recomenda-se a comparação do antes e depois, ou, sempre que não existirem grupos semelhantes, uma comparação entre os beneficiários e um grupo de controlo. No primeiro caso, esta comparação do antes e depois deverá ser executada adequadamente; caso contrário, deverá referir-se se isso teria sido possível. No último caso, a análise comparativa deverá poder basear-se na recolha de dados entre grupos semelhantes ou de controlo.

Resultados credíveis

A credibilidade dos resultados define-se, aqui, como o facto de eles seguirem logicamente, e serem justificados, pela análise de dados e interpretações, com base em hipóteses explicativas cuidadosamente apresentadas. A validade dos resultados deverá ser satisfatória. Isto significa que o equilíbrio entre a validade interna (ausência de influências técnicas na recolha e processamento dos dados) e a validade externa (representatividade dos resultados) deverá ser justificável. Também é necessário verificar se os resultados da análise foram produzidos de forma equilibrada e fidedigna.

A necessidade de executar uma análise aprofundada de uma parte do programa coloca o problema da extrapolação, a partir de estudos de caso, para o programa como um todo. Neste contexto, é necessário verificar:

- se as hipóteses de interpretação e extrapolação são justificáveis, e se os limites de validade foram definidos;
- se a selecção de casos e amostras torna possível generalizar as conclusões.

Conclusões imparciais

Entre as conclusões incluem-se sugestões e até, por vezes, recomendações que constituem mais do que resultados. Enquanto os resultados são "técnicos" e podem ser analisados sem demasiado risco de imparcialidade, as conclusões e, mais ainda, as recomendações são emitidas com base em julgamentos de valor. A qualidade destes, por conseguinte, é decisiva.

Para responder à questão se as conclusões são justas, isentas de considerações pessoais ou parciais e suficientemente pormenorizadas para serem implementadas em concreto, é necessário verificar:

- se os elementos em que se baseiam as conclusões são claros;
- se as conclusões são operacionais e suficientemente explícitas para serem implementadas;
- se as questões controversas são apresentadas de forma justa e equilibrada.

As principais questões como a relevância, a eficácia e a eficiência do programa deverão processar-se no âmbito de uma avaliação, e deverão, conseqüentemente, ser respondidas de forma adequada. O relatório de avaliação também deve mostrar a adequabilidade do orçamento do programa, tanto globalmente, como em termos de alocação interna de acordo com os diferentes eixos e medidas.

Também é necessário estudar questões essenciais como o valor acrescentado do programa e o progresso em termos de objectivos transversais como coesão, subsidiariedade, boa governação, desenvolvimento sustentável e igualdade de oportunidades. No caso dos exercícios ex-ante, é necessário formular conclusões, com vista a alimentar o processo de negociação do programa avaliado. O relatório deverá possibilitar a melhoria da capacidade de avaliação do programa.

Relatório claro

Os resultados da avaliação podem ser disseminados e transmitidos às partes interessadas, por escrito ou oralmente. O relatório final constitui apenas uma forma de divulgação entre muitas outras, sendo desejável uma comunicação contínua dos resultados. A legibilidade do relatório depende da qualidade da apresentação dos resultados e dos limites do trabalho executado. É necessário verificar se:

- o relatório foi escrito de forma clara, e estruturado logicamente;
- utiliza conceitos específicos apenas quando absolutamente necessário, e, em caso afirmativo, se os mesmos estão claramente definidos;
- a apresentação, os quadros e os gráficos melhoram a legibilidade e inteligibilidade do relatório;
- os limites da avaliação, em termos de âmbito, de métodos e de conclusões, são claramente apresentados.
- Em muitos casos, só se lê o resumo do relatório. É, por isso, essencial que este resumo seja claro e conciso. Deverá apresentar as principais conclusões e recomendações de forma equilibrada e imparcial. Deverá ser fácil ler sem necessitar de consultar o resto do relatório.

Critérios de garantia de qualidade

O conjunto de critérios seguinte relaciona-se com o processo e com o contexto global da avaliação: garantia de qualidade e não controlo de qualidade. Este irá permitir a todos os que avaliam a qualidade compreender o que poderá representar aspectos positivos e negativos das conclusões da avaliação, e retirar lições passíveis de aplicação futura, de modo a melhorar a qualidade das avaliações futuras.

Objectivos coerentes e avaliáveis

A coerência dos objectivos do programa: anteriormente, discutiu-se em que medida estes são específicos, relacionados com as intervenções, não contraditórios, etc.. Também se afirmou que a utilização de abordagens como modelos lógicos, da teoria dos programas e da teoria da mudança constituem formas úteis para clarificar os objectivos de um programa e a lógica das intervenções nas suas fases iniciais, antes do lançamento de uma avaliação. Nesta fase, estamos interessados nos resultados deste processo inicial. Em que medida é que os avaliadores estavam a lidar com um programa coerente em termos de objectivos e intervenções? Algumas dificuldades de avaliação resultaram de objectivos mal articulados ou de outros problemas relacionados com a capacidade de avaliação?

Termos de referência bem redigidos

Termos de referência sólidos permitem avaliações eficazes. Em certa medida, é possível, no momento da sua redacção, avaliar a adequação dos TdR. Também se torna mais fácil, com algumas pistas, identificar aquilo que poderia ser útil ter incluído. Este processo é importante em termos de aprendizagem futura, i.e., como melhorar os TdR no futuro.

A elaboração de TdR deficientes ou incompletos poderá levar os avaliadores a distribuir os seus recursos de forma inadequada. Também poderá conduzir a outros efeitos negativos. Uma consequência comum acontece quando as omissões nos TdR se tornam visíveis no decorrer de uma avaliação, e o requerente procura redireccionar a avaliação a meio do processo ou solicitar resultados adicionais, que não foram previamente planeados ou orçamentados.

Processo sólido de selecção de propostas

O processo de selecção de propostas foi devidamente orientado? Esta constitui tanto uma questão processual como uma questão substancial. Processualmente, deverá fazer-se uma avaliação da aplicação sistemática de critérios relevantes no momento da selecção. Substantivamente, interessa saber se a decisão tomada foi a correcta. Por exemplo, tomou-se uma decisão para favorecer uma empresa conhecida, mas o processo de afectação de pessoal-chave foi inadequado. O método foi pouco especificado? Ou será que se favoreceu um método experimental de alto risco, que poderia justificar problemas ocorridos posteriormente?

Diálogo e feedback eficaz ao longo do processo de avaliação

Manter um registo da avaliação, dar *feedback* e organizar um fórum para que as partes interessadas possam aprender através do diálogo, mutuamente e com os avaliadores, constitui um pré-requisito bem conhecido da qualidade da avaliação. Esta constitui, parcialmente, uma questão do fórum criado para esse efeito. A materialização mais óbvia talvez seja uma Comissão Directiva, mas também, possivelmente, reuniões de debate e workshops específicos, tais como workshops de discussão destinados a decisores

políticos e a políticos locais. É necessário avaliar a inclusão dos membros dessas reuniões: envolveu-se todos os públicos e partes interessadas correctos?

No entanto, o objectivo destas oportunidades de debate e de troca consiste no diálogo e no *feedback* que eles proporcionam. Fez-se uma utilização adequada das reuniões da Comissão Directiva? As agendas foram adequadas? As partes interessadas consideraram que estas oportunidades foram produtivas e que melhoraram a sua compreensão? Em última instância, ajudaram a formular e a melhorar a qualidade e a utilidade da avaliação?

Disponibilidade de informação adequada

Os avaliadores necessitam de informação. A Parte 4 deste GUIA destaca a importância da disponibilidade dos dados e dos sistemas de monitorização. Sem recursos de informação adequados, é difícil os avaliadores fazerem um bom trabalho. É necessário, portanto, fazer uma avaliação da adequação da informação. Obviamente, trata-se de monitorizar a informação e os sistemas. Muitas vezes, os sistemas de monitorização destacam as necessidades de patrocinadores e agentes de financiamento externos. Também necessitam de poder ajudar os gestores de programas, e uma avaliação mostra, normalmente, esta necessidade. Os avaliadores também necessitam de se basear em dados administrativos secundários, recolhidos pelas autoridades locais, regionais e nacionais, muitas vezes para outros fins.

Muita da informação de uma avaliação mantém-se na mente dos principais informadores, sobretudo, no caso de informação contextual e qualitativa considerada importante, não só para compreender o programa, mas também para interpretar dados mais formais.

De um modo geral, para avaliar a qualidade do processo e o contexto da avaliação, é necessário haver primeiro uma avaliação da existência de informação e, em segundo lugar, se a mesma foi colocada à disposição. Por exemplo, em alguns programas poderá haver dados disponíveis, tais como dados administrativos de emprego a nível local ou minutas das comissões de gestão de projectos ou subprogramas específicos, mas difíceis de aceder. Também poderá acontecer que o informador-chave se recuse a fornecer informação aos avaliadores, talvez devido a relações fracas entre as partes interessadas e as administrações. Neste contexto, a avaliação relativamente à disponibilidade de informação e de dados junto dos avaliadores poderá, ela própria, fornecer dados acerca do estado concreto da cooperação entre a parceria e inter-agências.

Uma boa gestão e coordenação por parte da equipa de avaliação

Independentemente do plano de trabalho e do relatório inicial estarem bem planeados e serem convincentes, todas as avaliações deverão ser convenientemente executadas. Os avaliadores deverão seguir os planos e conseguir adaptar-se a eventos inesperados, que tornam os planos (ou alguns dos seus aspectos) redundantes. É necessário que as equipas se mantenham unidas, que os diferentes componentes de trabalho sejam coordenados, e que os seus resultados sejam integrados. É necessário, também, gerir as relações com os requerentes da avaliação, com os gestores de programa e com uma vasta diversidade de informantes, locais de trabalho no terreno, instituições, associações e grupos envolvidos.

Estes aspectos da gestão constituem, sobretudo, uma responsabilidade da equipa de avaliação e dos seus gestores. Contudo, também existem elementos que são partilhados pelos gestores de programa e pelos que requisitaram a avaliação. Por exemplo, a forma

como o sistema de requisição responde a pedidos para adaptar um plano de trabalho elaborado anteriormente não se encontra sob o controlo apenas da equipa de avaliação.

Disseminação eficaz dos relatórios e das conclusões junto da Comissão Directiva e gestores políticos/de programa

A disseminação dos relatórios constitui outra responsabilidade partilhada. Em parte, depende da capacidade da equipa de avaliação em apresentar resultados de alta qualidade e bem planeados (em termos de controlo de qualidade, conforme especificado acima). Também requer um conhecimento do valor e das oportunidades de divulgação na equipa de avaliação. Existe, por exemplo, uma grande diferença entre os avaliadores que limitam o seu *feedback* ao mínimo contratual e os que consideram ser da sua responsabilidade proporcionar um *feedback ad hoc* sempre que ocorrem novos problemas ou sempre que é necessário solucionar questões essenciais.

Este tipo de divulgação também requer sensibilidade para com as necessidades de informação e interesses dos principais intervenientes. Por vezes, é necessário redimensionar as conclusões de acordo com os diferentes interesses. Por exemplo, os gestores de programa têm uma perspectiva diferente das PME locais, embora também partilhem determinados interesses comuns.

Divulgação eficaz junto das partes interessadas

É necessário divulgar os relatórios e as conclusões para que estes possam facilitar a aprendizagem das organizações e agências. Os outros grupos comerciais, informais e da sociedade civil também estarão interessados nas conclusões da avaliação, seja como contribuintes, eleitores locais ou potenciais beneficiários de um programa e das suas intervenções. Um processo de avaliação não deverá ser considerado como concluído até se implementar um programa de divulgação. Os requisitos gerais deste tipo de divulgação deverão estar assinalados nos TdR. Contudo, a responsabilidade não pertence exclusivamente aos avaliadores. Os gestores de programa e os requerentes da avaliação também deverão assumir a sua responsabilidade pela difusão junto das partes interessadas, incluindo o público em geral.

Avaliação sintética

A avaliação sintética recapitula todos os critérios de qualidade acima referidos. É difícil recomendar um peso específico a aplicar aos diferentes critérios, uma vez que a sua importância varia de situação para situação.

A Caixa *Critérios de controlo de qualidade e garantia de qualidade* (apresentada anteriormente) apresenta uma grelha de controlo de qualidade do relatório de avaliação. A Caixa *Grelha de análise da qualidade do processo de avaliação* proporciona uma grelha de garantia de qualidade.

Nos dois casos, utiliza-se uma escala de classificação de 5 pontos. Esta escala vai do positivo (com a classificação de muito positivo no extremo) ao negativo (onde muito negativo se encontra no outro extremo). Assim, existem duas possibilidades positivas e duas possibilidades negativas, e um ponto intermédio, onde o equilíbrio da análise não é seguro.

Caixa: Grelha de análise da qualidade do processo de avaliação

Por favor, classifique o relatório de avaliação quanto ao modo, positivo ou negativo, como	Muito positivo	Muito negativo
--	----------------	----------------

foi ao encontro de cada um dos critérios especificados de seguida:		
1. Satisfação de necessidades: O relatório de avaliação responde adequadamente aos pedidos de informação formulados pelos requerentes e corresponde aos termos de referência:		
2. Âmbito relevante: A filosofia do programa, as suas realizações, resultados, impactos, interacções com outras políticas e efeitos inesperados foram cuidadosamente estudados:		
3. Processo aberto: As partes interessadas – tanto os parceiros do programa, como as outras partes interessadas – estiveram envolvidas na concepção da avaliação e na discussão dos resultados, de modo a ter em conta as suas diferentes perspectivas:		
4. Concepção justificável: A concepção da avaliação foi apropriada e adequada de modo a obter resultados (no âmbito dos seus limites de validade) necessários para responder às principais questões da avaliação:		
5. Dados fidedignos: Os dados primários e secundários recolhidos ou seleccionados são adequados e fiáveis em termos da utilização esperada:		
6. Análise consolidada: Os dados quantitativos e qualitativos foram analisados de acordo com convenções estabelecidas, e de forma adequada a responder correctamente às questões de avaliação:		
7. Resultados creíveis: Os resultados são lógicos e justificados pela análise dos dados e por interpretações e hipóteses adequadas:		
8. Conclusões imparciais: As conclusões são justificadas e não tendenciosas:		
9. Relatório claro: O relatório descreve o contexto e o objectivo, bem como a organização e os resultados do programa de tal modo que a informação fornecida é facilmente compreendida:		
10. Recomendações úteis: O relatório faz recomendações úteis para as partes interessadas, e suficientemente detalhadas para serem implementadas:		
Considerando as restrições contextuais relativamente à avaliação, o relatório de avaliação é considerado:		

2.3. A utilização da avaliação e a gestão de conhecimento

Ver também *Gestão de Conhecimento* (<http://www.undp.org/gef/05/index.html>).

Executar um trabalho de avaliação e garantir a sua qualidade só compensa se conduzir a uma utilização das conclusões da avaliação e contribuir para um melhor conhecimento entre os que têm maior capacidade para tirar partido dele.

Existem, pelo menos, três formas diferentes de utilização do trabalho de avaliação.

- As utilizações individuais podem realizar-se directamente ou de forma instrumental, assumindo-se os resultados, as conclusões e as recomendações. Na prática, este processo é pouco habitual, e, sempre que implementado, tem tendência para decorrer apenas parcialmente.
- Mais frequentemente, utiliza-se cumulativamente diversas avaliações ou avaliações individuais em conjunto com outros dados e opiniões, de modo a informar debates e a influenciar a tomada de decisão. O trabalho de avaliação estimula, assim, o processo de discussão, desafio e contra-desafio dos dados e da sua interpretação.
- Mesmo quando os resultados de avaliação não são utilizados, o processo de iniciação e reflexão da avaliação poderá ser útil, na medida em que oferece oportunidades de trocar informações, clarificar o raciocínio e desenvolver enquadramentos.

A medida em que a avaliação é utilizada e o seu impacto são influenciados por uma série de factores:

- Os processos organizativos da disseminação. O tempo e os recursos disponíveis para a disseminação e o grau de acompanhamento do processo por todos os responsáveis pelo trabalho influenciam a compreensão, a comunicação e a utilização das conclusões.
- A qualidade do trabalho de avaliação. Sempre que os padrões de avaliação forem elevados, não é possível negligenciar facilmente os resultados.
- O envolvimento das partes interessadas nas diferentes etapas do ciclo de avaliação, juntamente com os avaliadores e os administradores, é essencial para incentivar a utilização da avaliação.
- O envolvimento dos directores e gestores sénior. Ajuda a garantir que a atribuição de políticas e de recursos, bem como a sua aplicação prática, são influenciadas pelas conclusões da avaliação.
- A aplicação de um sistema de acompanhamento sistemático das conclusões do trabalho de avaliação. Este processo chama a atenção para a utilização, ou não, das conclusões, e reduz a tendência para uma re-aprendizagem da mesma lição. A aplicação do processo não é comum.
- Os processos institucionais de realização da avaliação. Não existem modelos perfeitos. É muito provável que as conclusões da avaliação sejam úteis para os decisores, para todos os intervenientes no planeamento e concepção das intervenções, e para todos aqueles que se encontram envolvidos operacionalmente. Perante uma tendência de separação organizativa da avaliação, as funções operacionais e políticas podem levar a uma maior independência e qualidade do trabalho de avaliação. As preocupações políticas e operacionais podem, por exemplo, dar demasiado destaque àquilo que pode alcançar-se através da avaliação. Por outro lado, a separação poderá ser menos útil se conduzir a um enfoque exagerado na gestão da avaliação e limitar a utilização da avaliação. (Os processos institucionais discutem-se mais pormenorizadamente na secção do GUIA correspondente à Capacitação Institucional e Administrativa para a Avaliação.)

É possível concluir que a criação de uma cultura de avaliação é essencial para a aprendizagem organizativa. Entre os principais componentes de uma cultura de avaliação que orienta a geração de trabalhos de avaliação de qualidade, incluem-se: um pressuposto de que as intervenções devem ser concebidas e implementadas de forma a facilitar a posterior avaliação; uma apreciação de um conjunto de objectivos da avaliação; um reconhecimento dos limites da avaliação, o âmbito da interpretação e a necessidade

de combinar dados quantitativos e qualitativos; e um reconhecimento das necessidades de diferentes utilizadores da avaliação.

2.4. Regras de ouro

1. A avaliação da competência deverá ser convocada numa fase inicial, pelos planeadores do programa. Em concreto, isto poderá ajudar a clarificar os objectivos e a lógica de intervenção dos programas. Esta actividade, apesar de utilizar as competências da avaliação, está bastante afastada das actividades de avaliação dominantes. É necessário que ocorra na fase de concepção e planeamento do programa. Porém, poderá facilitar e tornar mais bem-sucedida a avaliação posterior. Para esse efeito, poderá implementar-se diversas técnicas, tais como aferição da capacidade de avaliação e preparação de uma análise da teoria do programa. Geralmente, para assegurar a independência da avaliação principal, o ideal seria utilizar, neste trabalho de planeamento do programa, equipas de avaliação ou recursos diferentes dos da avaliação principal.
2. Deverá executar-se uma aferição da capacidade de avaliação semelhante, pelos avaliadores, sempre que iniciarem o seu trabalho. De certa forma, isto poderá sobrepor-se ou repetir o que já ocorreu na fase de planeamento do programa. Contudo, neste caso o objectivo é diferente, destinando-se a assegurar a produção de um plano de avaliação viável e a clarificar a forma de utilização das conclusões de avaliação. Este processo é consistente com uma expectativa geral dos avaliadores, no sentido de que deveriam preocupar-se com a forma como os seus resultados, conclusões e recomendações se utilizam desde a fase mais inicial possível do seu trabalho.
3. As partes interessadas, os gestores de programas e os decisores políticos, os potenciais beneficiários e os parceiros devem envolver-se na avaliação desde as fases iniciais, sempre que possível. Isto irá garantir que o plano e a concepção da avaliação incluem as suas prioridades e ordens de trabalhos. Também irá garantir um sentimento de propriedade em relação às conclusões da avaliação e uma maior probabilidade, portanto, de as considerarem úteis e utilizarem essas conclusões. Por outro lado, poderá ser necessário ser-se selectivo, numa fase em que finalmente se determina a ordem de trabalhos da avaliação, de modo a manter o enfoque e a garantir que é possível gerir a avaliação. As prioridades mais abrangentes deverão ser formuladas pelas intenções e pela lógica do programa, permanecendo abertas a consequências inesperadas, sobretudo no caso de beneficiários alvo.
4. É necessário que as avaliações sejam geridas activamente, mas de forma sensata. Isto irá assegurar que os comissários estão conscientes das escolhas que será necessário adoptar ao longo do percurso. Também irá garantir que os avaliadores obtêm o apoio suficiente, acesso à informação e dados relativamente às alterações políticas e contextuais. As pessoas responsáveis pela requisição da avaliação e os gestores do programa são as pessoas mais adequadas para gerir a avaliação, uma vez que têm que estar conscientes do seu histórico e da sua filosofia.
5. Não é habitual retirar critérios a partir de uma avaliação, ou seja, a partir dos objectivos de um programa fazer uma avaliação positiva ou negativa da evolução. Também é importante incluir um conjunto de critérios mais abrangente, decorrente das necessidades sociais. Por exemplo, este programa é útil e ajuda as pessoas a

quem se destina? Apoia a equidade ou não? O programa é consistente com outras iniciativas políticas? E é implementado de uma forma eficaz e legítima? Manter esta perspectiva mais abrangente assegura que, pelo menos relativamente a parte do seu trabalho, os avaliadores podem permanecer fora da lógica do programa e assumir uma perspectiva crítica relativamente àquilo que o programa tenta alcançar e relativamente à forma como o faz.

6. A importância das questões de avaliação no processo de concepção da avaliação não pode ser sobrevalorizada. Antes pelo contrário, existe a tentação de recolher grandes quantidades de dados e de produzir, por vezes, indicadores tecnicamente sofisticados, que contribuem muito pouco para a aplicação prática ou política. Existe, obviamente, um problema na formulação das questões de avaliação, de tal forma que é provável dar-lhes uma resposta. Embora esta seja uma questão técnica e esta parte do GUIA tenha apresentado sugestões relativamente à forma adequada de formulação das questões, também existe uma preocupação constante com a utilização. Deverá procurar colocar questões que alguém considerará úteis. No entanto, a utilização não deve definir-se de forma demasiado restrita. Não nos referimos só à utilização instrumental da avaliação que os gestores possuem. Também nos referimos a possíveis formas de utilização da avaliação, por grupos de cidadãos e da sociedade, como forma de apoiar os processos democráticos e a responsabilização.
7. Especificámos com algum pormenor o conteúdo e a forma de um documento ideal de Termos de Referência para uma avaliação. Isto faz parte da questão geral da concepção, e as escolhas que poderá adoptar-se na fase de concepção podem influenciar a qualidade e a orientação de toda uma avaliação. É importante, por conseguinte, não seguir simplesmente um quadro padrão com parágrafos provisórios pré-redigidos. Deverá reconhecer-se que a definição do âmbito, a clarificação dos utilizadores da avaliação e a decisão das capacidades necessárias numa equipa de avaliação, encontram-se entre as decisões mais importantes que se deve adoptar no decurso de uma avaliação.
8. Era normal ver a utilização da avaliação como estando limitada a agir sobre as recomendações e sobre os relatórios finais. Compreende-se, agora, que a utilização da avaliação pode ser apoiada e decorre ao longo da avaliação. A denominada utilização de processo ou de diálogo deverá envolver as partes interessadas na concepção da avaliação, desde o início. Existem, inclusivamente, avaliações cujas conclusões e recomendações são rejeitadas, mas as partes interessadas, sobretudo as principais partes interessadas envolvidas na Comissão Directiva, consideram, no entanto, a avaliação útil. Poderá ajudá-los a clarificar o seu próprio raciocínio e a sua própria compreensão do programa, e despoletar ideias inovadoras para melhorar o programa. Este processo de comunicação contínuo proporciona um contexto específico para a disseminação dos relatórios e das conclusões da avaliação. A promoção do diálogo no decurso de uma avaliação apresenta uma grande probabilidade de garantir que, quando as partes interessadas receberem os relatórios, estarão mais preparadas e receptivas.
9. Muitas vezes, é mais fácil para os gestores de programa e para os requerentes da avaliação limitar a apreciação da qualidade da avaliação às conclusões da própria avaliação. No entanto, este processo de controlo de qualidade proporciona poucas oportunidades de aprendizagem e de melhoria da própria forma de gestão da avaliação. Uma perspectiva de garantia de qualidade do tipo que se defendeu nesta parte do GUIA proporciona um contexto para explicar as forças e as fraquezas dos resultados ou conclusões de avaliação. Também proporciona aos

requerentes da avaliação uma oportunidade para aprender a melhorar todo o processo no futuro.

10. Deverá considerar-se, numa fase inicial, a forma como as conclusões da avaliação serão colocadas em prática. Algumas formas de utilização decorrerão directamente das conclusões e recomendações do trabalho. Pode também ser útil combinar as avaliações individuais com outros dados, para informar as discussões. O processo de avaliação poderá trazer vantagens em termos de estruturação do inquérito e reflexão institucional. A consideração dos factores que influenciam a utilização do trabalho de avaliação irá maximizar o seu contributo.

(Índice da Parte III)

3. PARTE III: A CAPACITAÇÃO ADMINISTRATIVA E INSTITUCIONAL PARA A AVALIAÇÃO SOCIOECONÓMICA

3.1. Introdução

3.2. A capacidade de avaliação como parte da capacidade institucional

3.3. Em que consiste a capacidade de avaliação?

3.4. O desenvolvimento da capacidade de avaliação

3.5. A "arquitectura" dos sistemas de avaliação

3.6. O reforço da procura de avaliação

3.7. A perspectiva da oferta

3.8. A institucionalização da avaliação

3.9. Regras de ouro

3.1. Introdução

Esta secção do GUIA procura discutir uma das condições prévias da avaliação do desenvolvimento socioeconómico, que consiste em possuir e desenvolver a capacidade de executar avaliações. Começa-se por explicar a importância da capacidade de avaliação, e o que significa. Segue-se uma descrição resumida das acções a considerar para reforçar as capacidades, incluindo diversas estratégias consideradas eficazes em diferentes contextos nacionais. Discute-se a capacidade de avaliação numa perspectiva, quer de procura, quer de oferta: a necessidade que as organizações têm em requisitar, acompanhar e utilizar as avaliações e a necessidade de recursos humanos, capacidades e instituições disponíveis para que se possa satisfazer estas necessidades. O ponto de partida reside nas necessidades da administração pública nacional, mas também se reconhece o papel das administrações locais, descentralizadas e dos gestores de programas locais, bem como da sociedade civil, incluindo consultores, universidades, comunidades e o sector privado. O conteúdo desta secção do GUIA baseia-se num dos Manuais Técnicos com ele relacionado, que inclui estudos de caso da forma como diversos países têm vindo a reforçar, ao longo dos anos, as suas capacidades de avaliação.

3.2. A capacidade de avaliação como parte da capacidade institucional

O GUIA já apresentou um panorama ambicioso do que se espera de uma avaliação, da forma como deverá ser organizada e do que se espera serem os seus resultados. A secção anterior, por exemplo, oferecia aos sectores da administração que encomendam ou utilizam as avaliações algumas orientações sobre como conceber e implementar um processo de avaliação. Partiu-se do pressuposto de que estariam disponíveis as correspondentes capacidades institucionais e administrativas. Tendo em conta os principais objectivos deste Guia, define-se a capacidade institucional de uma forma abrangente, incluindo disposições legais (por exemplo, as que regulam o emprego e o aprovisionamento), pressupostos normativos (por exemplo, em termos de igualdade de oportunidades ou concorrência aberta), governança (por exemplo, responsabilização democrática e partilha de responsabilidade entre os diversos níveis administrativos e a sociedade civil), bem como disposições administrativas e organizativas (por exemplo, a forma como os ministérios estão estruturados e a respectiva distribuição de recursos –

ver Desenvolvimento da capacidade institucional e administrativa no Manual Técnico I). Esta secção do GUIA centra-se na forma como se procede à criação dessa capacidade, para que as ambições da avaliação sejam mais praticáveis, de modo a contribuir efectivamente para o desenvolvimento socioeconómico.

As capacidades não podem ser criadas de um momento para o outro, e o processo implica custos. Porém, as potenciais vantagens da avaliação são suficientemente abrangentes para justificar o investimento inicial, sendo necessária a inovação dos custos recorrentes, quer em termos de produtos, quer em termos de processos de avaliação. O desenvolvimento de capacidades requer tempo, e não é possível implementar os sistemas necessários ao mesmo tempo. Estes sistemas têm que ser alimentados a longo prazo, de modo a proporcionarem benefícios sustentáveis.

O desenvolvimento da capacidade de avaliação tem que ser uma preocupação partilhada pela comunidade responsável pelas decisões políticas mais alargada. Os decisores políticos e os responsáveis pela concepção de programas deverão, em primeiro lugar, estar convencidos de que existe uma necessidade de avaliação. Ao mesmo tempo, embora o seu apoio seja essencial, não se deverá permitir que se tornem detentores do processo. Conforme se realçou várias vezes ao longo deste GUIA, a independência constitui um ingrediente essencial da avaliação. Em primeiro lugar, inclui todos os que gerem e encomendam avaliações, todos os que estão interessados nos resultados da avaliação ao nível do programa e da implementação da política, bem como todos os responsáveis pela execução da avaliação. No entanto, a avaliação não é uma função independente; necessita de um enquadramento institucional e tem de estar integrada nos procedimentos institucionais. Especificamente, a capacidade das instituições públicas de executarem avaliações constitui parte integrante dos requisitos mais abrangentes que cada Estado tem que satisfazer, para dar resposta às necessidades económicas e sociais contemporâneas. De facto, verifica-se que os sectores em que se tem desenvolvido mais intensamente a capacidade de avaliação são precisamente os que a consideram uma parte integral de uma estratégia mais abrangente de inovação e modernização do sector público.

Esta necessidade de construir uma capacidade institucional e administrativa está, por vezes, implícita, mas constitui, cada vez mais, um objectivo transversal explícito das políticas de desenvolvimento socioeconómico. Este objectivo decorre de duas condições:

1. Ultrapassar a falta de eficiência da administração pública tradicional, passando para uma abordagem de gestão pública baseada em boas práticas de gestão, não só do sector público, mas também aplicando os princípios decorrentes das organizações privadas e sem fins lucrativos mais bem-sucedidas. Muitas das limitações assumidas quanto à eficiência dos organismos da administração pública encontram-se disfarçadas no designado movimento da Nova Administração Pública (NPM), que preconiza a reforma, a modernização, a descentralização e a privatização, entre outras, como soluções para o sector público. O NPM tem como objectivo melhorar os resultados, a implementação e o *Value for Money* (proveitos proporcionais ao investimento). (Este processo também se descreve mais em pormenor em Desenvolvimento da capacidade institucional e administrativa.)
2. Ultrapassar a distância percebida e a separação dos organismos públicos da sociedade em geral, e, assim, abrir a decisão política à contribuição dos parceiros externos, dos representantes da sociedade civil e dos cidadãos (a orientação global para a governança democrática).

A difusão da avaliação pode contribuir para as duas condições que se seguem:

- Em primeiro lugar, a identificação sistemática das melhores alternativas, bem como uma cuidadosa consideração da possibilidade de os programas em curso, passados e futuros, alcançarem os seus objectivos de uma forma eficiente. Esta situação foi identificada na secção anterior como sendo uma contribuição importante para a avaliação, que pode tornar-se um potente utilitário para a modernização do sector público, com vista a uma maior economia e uma maior capacidade de resposta junto dos cidadãos.
- Em segundo lugar, a abertura da caixa negra administrativa ao escrutínio dos parceiros externos, bem como a consideração dos interesses dos parceiros e dos cidadãos (por exemplo, incluindo grupos da sociedade civil nas comissões directivas, consultas sobre os resultados e as conclusões da avaliação, etc.), constitui uma incorporação dos princípios de governança democrática. Uma vez que o pedido de avaliação sistemática dos programas públicos permite aumentar a confiança dos cidadãos no governo, contribui para o aumento e manutenção do capital social. Uma vez que as teorias contemporâneas de desenvolvimento socioeconómico assentam, sobretudo, na mobilização de potencial e recursos com base territorial (endógenos), estes aumentos do capital social ajudam a promover o desenvolvimento socioeconómico sustentável.

3.3. Em que consiste a capacidade de avaliação?

A capacidade de avaliação é multifacetada e necessita de ser integrada a diversos níveis diferentes, que se reforçam mutuamente. Por exemplo, existe:

- Um nível individual, constituído pelas capacidades e competências necessárias;
- Um nível organizacional de procedimentos e estruturas de gestão;
- Um nível inter-organizacional, que liga organismos públicos e privados, através de redes, processos e parcerias; e
- Um nível social, que integra considerações avaliativas na sociedade civil, incluindo organizações profissionais – bem como no sector público.

Existem, transversalmente a todos estes níveis, regulamentações, legislação, recursos, normas e factores institucionais que pressupõem necessariamente determinadas condições e mecanismos organizacionais.

Existem diversos indicadores da capacidade de avaliação, sobretudo condições que se espera estejam disponíveis, para que um sistema de avaliação seja sustentável. Embora não seja viável tentar implementar todas estas condições ao mesmo tempo, vale a pena descrever um estado ideal, que permita o planeamento de uma estratégia de implementação. Numa situação ideal, seria de esperar que se encontrasse as seguintes condições, embora os exemplos que se seguem não sejam exaustivos ou abrangentes:

A nível individual:

- Existem pessoas em todos os níveis de administração com experiência e capacidades de avaliação, e este factor é reforçado e renovado por actividades de recrutamento e formação adequados, bem como de desenvolvimento profissional;
- Estão disponíveis orientadores de recursos formativos, cursos de formação e diplomas, a nível nacional ou regional, muitas vezes oferecidos por universidades, organismos de formação privados e organismos profissionais, à disposição de profissionais e requerentes.

A nível organizacional:

- A avaliação é executada de forma rotineira em cada uma das fases de concepção e de programação da política: do planeamento à implementação e ao acompanhamento;

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- Os resultados da avaliação estão integrados no processo de tomada de decisão relativamente às opções de política, à melhor forma de proceder à sua implementação e oferta, e à verificação do que foi verdadeiramente eficaz;
- Os gestores vêem a avaliação como um recurso que irá ajudá-los a melhorar o seu desempenho e a gerir os seus resultados;
- É pedido um fluxo regular de avaliações que abrange um vasto espectro de políticas e programas;
- Executa-se procedimentos de acompanhamento, no sentido de assegurar que as recomendações da avaliação são devidamente consideradas, e, quando viável, que se adoptam medidas conducentes à sua intervenção;
- Existem procedimentos que permitem agregar os resultados da avaliação e as lições daí tiradas, para que os gestores de programa e os decisores políticos possuam uma base de sustentação acessível, e uma memória organizacional.

A nível inter-organizacional:

- Existe uma coordenação baseada numa rede de funções ou unidades de avaliação dedicadas, de modo a assegurar a consistência necessária na forma como as avaliações são requisitadas, geridas e executadas em toda a administração, e, em última instância, em todo o sector público;
- Existem requisitos para que os processos de avaliação decorram de acordo com um enquadramento legislativo, políticas articuladas e no âmbito de actividades reguladoras (por exemplo, auditorias ou revisões parlamentares);
- Existe um mercado bem definido, com regras claras, para que as potenciais entidades de avaliação possam organizar-se para responder a concursos, concluir trabalhos de avaliação atempadamente, desenvolver conhecimentos especializados sectoriais e técnicos, e compreender as prioridades das políticas e os clientes dos programas;
- Existe uma cultura de avaliação que valoriza elevados padrões profissionais, a independência, a aprendizagem com a experiência e as políticas constituídas com base em factos concretos.

A nível societal:

- Mantém-se um diálogo aberto e sistemático entre os decisores políticos e os especialistas da avaliação, no sentido de identificar e escrutinar as prioridades da avaliação;
- Existe uma comunidade de avaliação prática, cujos membros podem trabalhar em universidades ou consultoras, ou então trabalhar como consultores e agentes de avaliação independentes, partilhando, no entanto, as mesmas normas e códigos de ética;
- Existem associações de avaliação que, agregando em torno de si todos os que requerem, prestam, utilizam a avaliação, reforçam uma cultura de avaliação, disseminam boas práticas e contribuem para salvaguardar a independência das funções de avaliação e dos seus agentes;
- Existe uma maior sensibilização para as actividades de avaliação e os seus resultados, bem como uma disseminação dos relatórios e das conclusões, de tal forma que a avaliação virá a ser utilizada de forma regular pelos diversos parceiros (em contextos parlamentares, na sociedade civil, etc.), apoiando a transparência e a responsabilização democrática.

3.4. O desenvolvimento da capacidade de avaliação

Em alguns países, as capacidades de avaliação evoluem organicamente, por um longo período de tempo. É este o caso de países como o Reino Unido, a Holanda e a Suécia. Na maior parte dos países, no entanto, é necessária a adopção de planos e estratégias

para que a capacidade de avaliação se torne útil, num intervalo de tempo bem definido. É este o caso da maior parte dos novos Estados-membros da União Europeia. Sempre que se começa a desenvolver a capacidade da avaliação de forma estratégica, é necessário tomar diversas decisões de base. Entre as mais importantes, incluem-se:

- Arquitectura: localizar e estruturar as funções de avaliação e a sua coordenação;
- Reforço da procura de avaliação: assegurar a existência de uma procura de avaliação eficaz e devidamente gerida;
- Reforço da prestação de avaliação: garantir a implementação das capacidades e competências, com um apoio organizacional adequado;
- Institucionalização da avaliação: desenvolver e integrar a avaliação em sistemas de elaboração de política, e no sistema de políticas mais abrangente.

Ver o exemplo na Caixa [A capacidade de avaliação na Lituânia](#).

Caixa: A capacidade de avaliação na Lituânia

Antes da sua adesão, a Lituânia possuía uma experiência de avaliação muito reduzida, bem como uma capacidade limitada de avaliação dos Fundos Estruturais. A estratégia adoptada consistiu em centrar a responsabilidade global da avaliação no Ministério das Finanças, envolvendo simultaneamente todos os parceiros num Grupo de Gestão da Avaliação. Foi contratado um consórcio, constituído por consultores locais e estrangeiros, para executar um projecto de capacitação em termos de avaliação.

Entre as medidas adoptadas, encontram-se:

- Uma conferência regional sobre avaliação, organizada para aumentar a consciencialização;
- Formação de peritos e funcionários públicos em avaliação;
- Preparação de medidas orientadoras e de documentos metodológicos.

Fonte: Conference Presentation, Director EU Programme Management Department Ministry of Finance, Lithuania

3.4. A "arquitectura" dos sistemas de avaliação

É importante que todos os responsáveis pelo desenvolvimento da capacidade de avaliação tenham consciência de que estão a desenvolver um sistema onde os diversos elementos interagem entre si, e se reforçam mutuamente. Entre estes elementos encontram-se, por exemplo, divisões ou unidades responsáveis nos departamentos, procedimentos e protocolos, políticas de recursos humanos e pessoal, sistemas de tecnologias da informação, etc. Nestas circunstâncias, é necessário que exista um responsável que irá delinear uma estratégia. Geralmente, a responsabilidade, pelo menos nas fases iniciais do desenvolvimento da capacidade de avaliação, pertence a uma unidade dedicada da administração central. Esta responsabilidade poderá pertencer a um departamento de planeamento, às Finanças, a um departamento coordenador, ao gabinete do Primeiro-Ministro, ou a um departamento responsável pelos Fundos Estruturais da UE.

Na fase inicial do desenvolvimento das capacidades de avaliação, encontram-se normalmente dois tipos de arquitectura:

- Uma função ou unidade de avaliação central, que também está associada, desde o início, a uma rede de outras unidades centradas noutros ministérios, sectores ou programas. A unidade central proporciona uma organização abrangente, onde são acordadas as directrizes e os procedimentos básicos, bem como protocolos de

requisição e aprovisionamento público. Uma vez estabelecido um sistema, a unidade central continua a orientar os trabalhos, mas as restantes unidades passam a ter uma autonomia cada vez maior, representando as suas próprias prioridades no seio da rede.

- Uma unidade central, localizada num determinado departamento desde o início, concentrando aí o seu trabalho. Porém, uma vez estabelecida a unidade, pretende-se disseminar o projecto-piloto e mostrar os seus próprios resultados concretos. Esta disseminação poderá ser formalizada quando adoptada a decisão, a nível central, de que, no futuro, todos os ministérios terão unidades de avaliação. Em alternativa, poderá escolher-se um processo mais informal: difundir as boas práticas e incentivar a sua replicação. Ver o exemplo apresentado na Caixa Coordenação da avaliação na Roménia.

Caixa: Coordenação da avaliação na Roménia

A Unidade de Avaliação da Comissão de Gestão da Roménia possui uma série de tarefas essenciais, incluindo:

- Presidir e assegurar o Secretariado de um Grupo de Trabalho de Avaliação transversal da Comissão de Gestão;
- Coordenar e gerir a avaliação, num quadro nacional (NRSF);
- Supervisionar a preparação de uma estratégia de avaliação nacional (capacidade).

Fonte: Entrevistas com as autoridades romenas

Existem vantagens e riscos associados a estas duas abordagens. É necessário que uma função ou unidade de avaliação central consiga proporcionar um equilíbrio perfeito entre a garantia de evolução de um sistema coerente e uma gestão que não seja demasiado centralizada. É necessário manter um equilíbrio entre a adaptação dos sistemas aos condicionamentos locais e a garantia de consistência. A centralização num ministério coordenador poderá pôr em causa a responsabilidade e o apoio da avaliação noutros locais do governo, no sector público ou entre outros parceiros. Uma publicação recente do Banco Mundial indica que isto terá acontecido em alguns países: a centralização conduziu à reduzida utilização da avaliação noutras partes do governo em causa. Por outro lado, uma descentralização exacerbada (como, por exemplo, possuir procedimentos de coordenação diferentes para os diferentes programas, políticas ou territórios) corre o risco de ser incoerente.

Limitar as fases iniciais do desenvolvimento da capacidade de avaliação a um ministério-piloto poderá levantar um problema comum de se criar enclaves de inovação. A localização do desenvolvimento da avaliação poderia adaptar-se de uma forma demasiado próxima a um único contexto político, e ser inflexível e difícil de transferir. Este problema ocorreu em alguns Estados-membros da UE, que fizeram um grande investimento no desenvolvimento das capacidades de avaliação nos Fundos Estruturais, considerando depois difícil alargar os sistemas e práticas de avaliação a outros programas financiados a nível nacional. Registaram-se casos de países candidatos à adesão que adoptaram os seus sistemas de avaliação aos requisitos do PHARE (programas de pré-adesão e de transição), constatando a dificuldade de transferir modelos, capacidades e procedimentos institucionais para os Fundos Estruturais. Como se pode ver no caso irlandês (ver o Manual Técnico III e a Caixa Experiência das avaliações intercalares na Irlanda), torna-se de algum modo comum verificar que as unidades de avaliação se restringem a programas financiados pela UE no início, e só posteriormente alargam a sua aplicação a programas financiados através de instituições nacionais.

Caixa: Experiência das avaliações intercalares na Irlanda

Entre 1992 e 1995, foram criadas três unidades de avaliação na Irlanda com um estatuto independente: a Unidade de Avaliação de Programas do FSE, a Unidade de Avaliação Industrial e a Unidade de Avaliação Agrícola. A maioria dos gestores e alguns dos avaliadores das unidades provinham de departamentos governamentais ou de agências associadas e, por conseguinte, conheciam bem o sistema. Esta proximidade facilitou o acesso à informação e, em geral, as unidades usufruíram de um acesso livre à informação nos diferentes departamentos. Em termos de avaliação interna face à avaliação externa, poder-se-á dizer que estas unidades se tornaram híbridas, procurando integrar elementos dos dois tipos de avaliação. As unidades eram sobretudo responsáveis por executar actividades de avaliação interna, centrada em aspectos de gestão dos programas.

Fonte: Estudo de caso sobre a Irlanda no Manual Técnico III, A Capacitação Administrativa e Institucional para a Avaliação

Uma vez estabelecido um sistema de avaliação, passam a estar disponíveis opções adicionais em termos de localização e de estrutura. Por exemplo, alguns ministérios ou agências públicas podem concentrar os seus esforços de avaliação na sua própria unidade central, enquanto outros podem basear-se em redes internas, com uma ou duas pessoas num ponto central, e as restantes distribuídas pelos diferentes departamentos, unidades ou divisões de implementação política. Um exemplo desta estrutura é uma rede de consultores de avaliação, um em cada um dos Programas Operacionais dos Fundos Estruturais, mas sob a coordenação, de uma forma transversal a todos os Programas Operacionais, de uma unidade central.

Se a intenção for efectivamente transformar a avaliação numa função valorizada no âmbito do sector público, terá necessariamente de reflectir a arquitectura de governação do respectivo país. Por exemplo, se for importante a participação de níveis regionais ou federais do governo, a avaliação deverá por conseguinte estar plenamente integrada nessas estruturas de administração sub-nacional. Embora possa não ser este o principal enfoque da avaliação, é necessário conceber novos sistemas de avaliação de raiz, com base numa visão global da forma como serão posteriormente adaptados na generalidade.

3.5. O reforço da procura de avaliação

- Os governos como fonte de procura da avaliação
- A sociedade civil como fonte de procura da avaliação

Os governos como fonte de procura da avaliação

A procura de avaliação, o primeiro impulso de lançamento do processo de desenvolvimento de capacidades, poderá advir maioritariamente de fontes internas, ou maioritariamente de fontes externas. Em diversos países europeus, a União Europeia e, particularmente, as regulamentações dos Fundos Estruturais proporcionaram o primeiro impulso, pelo que a procura foi inicialmente externa. Este foi o caso de países do sul da Europa, tais como a Grécia, Itália e Portugal, e também é o caso da maior parte dos novos Estados-membros. Esta procura externa constitui, em parte, um *outsourcing* das responsabilidades das instituições europeias, que se vêem obrigadas a demonstrar que são igualmente responsáveis. A Comissão Europeia possui obrigações para com os Estados-membros, para com o Conselho Europeu e para com o Parlamento Europeu, de

garantir que os recursos colocados à disposição para efeitos de desenvolvimento socioeconómico são distribuídos de uma forma sensata, e que os objectivos políticos de uma melhor coesão económica e social são cumpridos. Uma vez que os destinatários dos recursos financeiros são os mesmos governos que proporcionam o financiamento, é perfeitamente natural que a avaliação fique, pelo menos parcialmente, a seu cargo, e não desenvolvida a nível europeu. Um nível de procura externa da avaliação semelhante já se encontra consolidada quando se trata do desenvolvimento internacional. Os financiadores, sejam eles países individuais ou agências multilaterais, tais como a própria UE, os bancos internacionais (o Banco Mundial, o Banco Europeu de Reconstrução e Desenvolvimento, etc.), ou mesmo as agências das Nações Unidas exigem que os beneficiários dos fundos demonstrem que estes foram bem aplicados e utilizados.

Em paralelo com o factor da responsabilização, os actores externos que iniciam uma procura de avaliação também estão interessados na eficácia, nas boas práticas e no *value for money* (proveitos proporcionais ao investimento) das acções, ou seja, estão interessados na aprendizagem, bem como na responsabilização. Porém, esta motivação é ainda mais forte quando o ímpeto da avaliação é interno.

A maioria das entidades que entraram no desenvolvimento da capacidade de avaliação por razões de exigência nacional, fizeram-no como parte de um compromisso com um sector público eficaz, eficiente, transparente e responsável. A avaliação no Canadá, na Holanda, na Suécia e no Reino Unido não começou com pedidos externos. Antes, a procura da avaliação foi induzida por ciclos sucessivos de reformas internas do sector público, com o intuito de otimizar a sua gestão e eficácia, bem como novas noções de governança. Foi este o caso, também, das fases iniciais de reforma do sector público na Austrália, orientadas sobretudo por uma filosofia assente em resultados, reflexo do destaque atribuído à avaliação e à sua utilização (Caixa *Estímulo da capacidade de avaliação na Austrália*).

Caixa: Estímulo da capacidade de avaliação na Austrália

Foi implementada uma série de reformas de gestão do sector público nos primeiros anos do novo governo. Um aspecto destas reformas foi a vontade de ver os gestores a desenvolverem o seu trabalho com base na descentralização de poderes e responsabilidades, reflectindo uma filosofia em que os gestores do sector público seriam altamente incentivados a melhorar o seu desempenho se lhes fosse dada uma maior autonomia e potencial para gerir os seus departamentos com um menor nível de controlo por parte da agência central e com um menor grau de interferência. Um outro aspecto das reformas consistiu em atribuir aos gestores uma maior capacidade de gestão, e esta orientação estava mais directamente relacionada com fortes pressões orçamentais.

The Development of Australia's Evaluation System, Keith Mackay, World Bank ECD Working Paper 1998

Isto não significa que a própria avaliação interna não seja, muitas vezes, externalizada. Mesmo quando a força motriz por detrás da implementação da avaliação decorre de contextos nacionais, é provável a existência de escrutínio externo. Isto poderá assumir diversas formas:

- Parlamentos a nível nacional, regional ou local que procuram responsabilizar o governo pela implementação eficiente e eficaz das decisões tomadas;
- Tribunais de Contas e outros organismos semelhantes, que pretendam alargar o seu papel da verificação da legalidade e do controlo financeiro, de forma a integrar noções de eficácia, eficiência e *value for money*;

- Ministros das finanças que pretendam relatórios de departamentos aos quais foram atribuídas dotações orçamentais;
- Administração central que financia as actividades ou o investimento das autoridades sub-nacionais, sob condição de que estas implementem a avaliação na sua prática e/ou se abram a alguma forma de escrutínio externo.

Se o ímpeto da avaliação for integralmente externo (ou mesmo transposto externamente por agências ou departamentos governamentais a organismos de escrutínio nacional), é provável que a base de sustentação da construção da capacidade de avaliação seja fraca. Para que a procura seja levada a sério, deverá haver um reconhecimento, intrinsecamente interessado, de que a avaliação é mais do que uma obrigação para com interposta pessoa. Este princípio é central em debates sobre a necessidade de força legal ou regulamentar por detrás da avaliação. Há pessoas que argumentam que, sem uma procura legal clara, apoiada por mecanismos de conformidade adequados, a avaliação não será levada a sério. Por outro lado, também existe um forte argumento de que uma regulamentação sem mudanças culturais, administrativas e de gestão não será eficaz. Haverá, assim, uma conformidade simbólica, em vez de uma adopção intencional, de novas práticas. (ver exemplo na Caixa *A cultura de avaliação na Holanda*).

Caixa: A cultura de avaliação na Holanda

Uma lição que é possível retirar da experiência na Holanda é que, para que seja possível desenvolver uma cultura de avaliação num país, é importante que a motivação para executar e utilizar a avaliação não seja orientada meramente por factores externos. A motivação interna do governo Holandês para melhorar as políticas públicas, incentivada por noções de credibilidade e de economia da gestão pública, poderá, de facto, ser vista como um factor importante para o sucesso relativo da introdução e utilização do modelo de gestão baseado em resultados.

Fonte: O estudo de caso da Holanda, Manual Técnico III, A Capacitação Administrativa e Institucional para a Avaliação

Existem diversas razões pelas quais os sistemas baseados exclusivamente em regras e regulamentações poderão não funcionar:

- A sua implementação depende de determinado tipo de sanção. No entanto, a capacidade ou vontade das autoridades políticas competentes, sejam elas a Comissão Europeia, a administração central dos Estados-membros ou as autoridades sub-nacionais, em exercer esta sanção são, muitas vezes, limitadas. Estas limitações poderão decorrer, tanto de considerações políticas, como de capacidades legais ou administrativas.
- Paralelamente à probabilidade limitada de interposição de sanções, os requisitos da avaliação são levados menos a sério, e os documentos e processos de avaliação têm tendência para serem menosprezados pelos mesmos organismos que os prescreveram.
- Os gestores preocupados com as suas responsabilidades quotidianas irão minimizar o seu compromisso temporal para com a avaliação, se esta não for vista por eles como sendo directamente relevante e útil. É pouco provável que as avaliações concebidas para satisfazer as necessidades de intervenientes externos sejam consideradas importantes e úteis.

As implicações deste último aspecto são profundas. O reforço da procura da avaliação requer, em primeiro lugar, a mobilização de uma coligação abrangente de parceiros internos, incluindo gestores e decisores políticos. Não se pode criar simplesmente por um ou dois especialistas em avaliação empenhados, ou decidindo estabelecer uma pequena unidade central. Sempre que for criada uma coligação deste género, é possível

desenvolver procedimentos e políticas adicionais, que reforcem a procura inicial. Por exemplo, muitas das actividades específicas identificadas na secção anterior do GUIA, relacionadas com o pedido e gestão da avaliação, tornam-se, assim, importantes. A criação de procedimentos de consulta de gestores e decisores políticos durante o desenvolvimento de termos de referência ou de um plano de avaliação anual, a criação de comissões directivas, e a revisão da qualidade de relatórios de avaliação permitem gerar situações de procura adicional de avaliação. Porém, o incentivo mais forte decorre da integração da avaliação nas práticas comuns, quotidianas, das agências públicas ou dos ministérios governamentais. Esta questão discute-se mais pormenorizadamente de seguida - ver secção sobre a *Institucionalização da avaliação*.

3.6. A sociedade civil como fonte de procura de avaliação

Um sistema de avaliação sustentável não pode estar confinado aos governos ou mesmo ao sector público. A sociedade civil também pode originar procura de avaliação. O reforço da avaliação é visto, muitas vezes, como um processo descendente nacional ou, mesmo, transnacional. A responsabilização e a aprendizagem decorrem ao nível dos decisores políticos, esperando-se que outros actores, de níveis hierarquicamente inferiores, satisfaçam os requisitos da comunidade política. Em termos de concepção contemporânea de políticas, esta perspectiva descendente é anacrónica. A participação, o envolvimento activo, a responsabilidade e a transparência pressupõem que os actores da sociedade civil se encontram, eles próprios, empenhados na prestação de serviços e na decisão política pública, e tem responsabilidades de eficiência e capacidade de resposta, partilhadas com os governos. O desenvolvimento de capacidades tem que incluir actores da sociedade civil para ser fiel à lógica e aos valores da reforma do sector público. Este processo poderá alcançar-se através da consulta pública, procurando contribuições para as prioridades de um plano de avaliação anual, e garantindo, conforme referido anteriormente, uma representação alargada dos interesses das comissões directivas e dos grupos consultivos.

Um exemplo interessante desta situação na União Europeia reside no caso da região basca, em Espanha, onde se fez uma tentativa deliberada em termos de desenvolvimento da avaliação, no caso das avaliações do plano de desenvolvimento rural sustentável do País Basco (no âmbito dos Fundos Estruturais da UE) para adoptar uma concepção de avaliação de natureza participativa (ver exemplo na Caixa *Capacidade de avaliação participativa*).

Caixa: Capacidade de avaliação participativa

A avaliação deve, em primeiro lugar, constituir uma prática útil a todos os responsáveis pela execução do programa; porém, também deverá proporcionar a mesma função a todos os potenciais utilizadores e entidades colectivas afectadas pelos seus resultados, beneficiários, organismos sociais, associações civis, etc. Nesta perspectiva, a avaliação torna-se, assim, um exercício destinado a dar resposta às necessidades de informação dos agentes sociais na sua totalidade, e vai, inclusivamente, ao ponto de promover a sua participação activa no processo de avaliação.

Fonte Diez, Malagon e Izquierdo

Citando uma classificação das capacidades de avaliação executada pela C3E em Lyon, os autores indicam que este tipo de avaliação é comum numa fase avançada da avaliação (ver a Caixa *A sociedade civil e a procura de avaliação*).

Caixa: A sociedade civil e a procura de avaliação

[A avaliação] começou por ser um acto social e político, e as suas conclusões estão sujeitas a discussão pública (utilização política da avaliação). O objectivo é fornecer ao público informações acerca da eficácia da utilização de recursos públicos pela administração. Neste contexto, a avaliação torna-se num instrumento do processo democrático, na medida em que informa os cidadãos, proporciona responsabilização e obtém aprovação social para as intervenções do Estado. Entre os países onde o desenvolvimento da avaliação atingiu esta fase, encontram-se a Dinamarca, a Holanda, a Suécia e o Reino Unido.

Fonte Diez, Malagon e Izquierdo, com base numa tipologia desenvolvida pela C3E Lyon, no âmbito do Programa MEANS

Impõe-se, naturalmente, o caso da avaliação ser executada nos países referidos, em primeiro lugar, em resposta à procura a nível nacional, mais do que com o objectivo de satisfazer necessidades de responsabilização externas (por exemplo, a UE), e estes processos de avaliação precederam ou, pelo menos, desenvolveram-se independentemente dos regulamentos dos Fundos Estruturais e dos sistemas de avaliação a eles associados. Porém, ainda mais surpreendente é a natureza descentralizada e multi-nível da avaliação, que reflecte as realidades da governança multi-nível.

No Reino Unido, por exemplo, é já uma prática consolidada a de complementar as avaliações de políticas nacionais com avaliações locais, que têm como objectivo não só considerar questões descentralizadas de prestação e implementação, como também responder a questões de avaliação, com interesse para os beneficiários das políticas e dos programas, bem como de comunidades locais, ONGs e actores de política nacionais (Biott and Cook, 2000). Os actores da sociedade civil já mobilizados tornam-se também uma fonte de procura de avaliação, se houver a perspectiva de que os potenciais benefícios se reverterão a seu favor.

3.7. A perspectiva da oferta

- Tradições e condições de partida
- Relações com os sectores da educação, formação e universidades
- Consultores independentes: incentivar novos prestadores
- Uma comunidade de avaliação profissional

Tradições e condições de partida

Independentemente de haver um elevado nível de procura de avaliação, a resposta - e a celeridade da resposta - a essa procura irá depender dos recursos, das capacidades e das instituições previamente existentes. Em parte, como já foi referido anteriormente, esta será uma questão de oferta intermédia, integrada nas rotinas governativas e administrativas. No entanto, irá depender em grande medida das capacidades, dos recursos e das instituições que são externas ao sistema de gestão pública e administrativa.

Não existe uma forma única de aumentar a oferta de avaliação. Existem três abordagens às quais se recorre com maior frequência:

- A construção de relações com instituições educativas, em particular com universidades;

- O desenvolvimento e reforço de uma comunidade de consultores independentes;
- O apoio ao desenvolvimento de uma comunidade de avaliação profissional.

Relações com os sectores da educação, da formação e as universidades

As instituições de ensino, e em particular as universidades, são importantes para a perspectiva da oferta de avaliação, por duas razões. Em primeiro lugar, as universidades são, em muitos países, o local natural para instituir programas de educação e formação, necessários para desenvolver conhecimentos e capacidades de avaliação. Em segundo lugar, muitos dos recursos humanos a trabalhar em instituições e estabelecimentos de ensino já possuem muitos dos conhecimentos subjacentes à investigação social e económica, à recolha e análise de dados, e à preparação de relatórios, necessários para os avaliadores. No entanto, as perspectivas de ensino e de investigação terão de ser complementadas por outros métodos, teorias e práticas para se tornarem úteis à avaliação.

O estabelecimento de relações de colaboração com universidades e outras instituições de ensino poderá contribuir para estimular a oferta de programas de formação formais, educação contínua e a criação de um ou mais centros de excelência, relacionados com a investigação (por exemplo, em gestão do sector público ou em desenvolvimento regional), e com capacidade para disseminar os conhecimentos. Os cursos podem ser orientados para diferentes grupos-alvo. Por exemplo, estudantes graduados em ciências sociais e económicas, especialistas em estudos regionais e políticos, e profissionais do desenvolvimento socioeconómico. Dependendo do grupo-alvo, poderá ser mais adequado criar novos cursos especializados ou internalizar os módulos de avaliação (por exemplo, integrar as teorias e métodos de avaliação socioeconómica nos *curricula* formais de economistas, engenheiros, sociólogos, planeadores, etc.). Deverá incluir-se nestes programas questões específicas, importantes para a área do desenvolvimento socioeconómico (por exemplo, prioridades transversais, tais como a sociedade da informação, a igualdade de oportunidades, e temas afins).

Numa fase em que se procede ao planeamento de programas de educação e formação, é importante reconhecer aspectos profissionais da prática da avaliação no domínio socioeconómico. O trabalho desenvolvido num contexto que envolve clientes e condições contratuais, a avaliação do que será mais vantajoso para os gestores de programa, o planeamento e a gestão de equipas, e a gestão de diversos interesses, muitas vezes em situação de conflito, são todos aspectos que tendem a surgir com maior regularidade num contexto de avaliação do que no ambiente académico. São necessárias capacidades profissionais, bem como conhecimentos académicos. É possível obter este nível de integração em ambientes que implicam abordagens de ensino e aprendizagem menos formais, como, por exemplo, *workshops*, seminários de formação, sessões práticas com profissionais, estágios com consultores de avaliação, visitas de estudo a entidades administrativas com responsabilidades no desenvolvimento socioeconómico, etc.

Alguns países possuem uma tradição consolidada em que universidades e institutos independentes estão activamente envolvidos em actividades de investigação relacionadas com a formulação de políticas e práticas implementadas, e onde existirá conseqüentemente uma bacia de capacidades à qual se pode recorrer para aumentar a oferta de avaliadores num prazo relativamente curto. No entanto, mesmo nesses países, isso poderá implicar transformações culturais entre todos os que estão habituados a um ritmo e a um estilo de investigação mais académico. Nestes casos, o aumento da oferta de avaliação levará algum tempo a mobilizar-se, começando, talvez, com *workshops* e seminários de sensibilização, passando posteriormente a tarefas específicas, de pequena escala, que permitirão aos potenciais avaliadores testar novas formas de trabalhar num ambiente político.

Consultores independentes: incentivar novos prestadores

Em diversas áreas da avaliação, existe um mercado internacional para a avaliação. Este é o caso dos processos de avaliação de desenvolvimento socioeconómico. Existem grandes consultoras internacionais que se candidatam a concursos de avaliação numa diversidade de países, e não apenas no local onde se encontram sediadas. Está, por isso, a tornar-se cada vez mais frequente que os novos operadores, a entrar em novos mercados de avaliação, sejam provenientes destas consultoras internacionais. Estes possuem o potencial para importar uma ampla experiência específica, em termos internacionais e de política, bem como experiência de trabalho com decisores políticos e gestores de programa. Porém, considera-se igualmente importante que exista um sector de consultadoria de avaliação local, com consultores locais, familiarizados com as instituições, as políticas e as línguas nacionais. Naturalmente, em muitos casos, os consultores internacionais também procedem ao recrutamento junto de homólogos nacionais, incentivando, assim, o intercâmbio internacional, mesmo dentro de uma única equipa de avaliação. Este trabalho conjunto poderá constituir um método eficaz de transferência de capacidades e de disseminação de conhecimentos. Este factor é amplamente reconhecido em diferentes sectores da avaliação. Por exemplo, em programas de desenvolvimento internacional exige-se frequentemente que as equipas de avaliação incluam peritos dos hemisférios norte e sul. Esta prática também foi incentivada na UE, no âmbito de uma noção abrangente de aproximação entre os 15 Estados-membros e os novos Estados-membros, reflectindo-se mais especificamente na composição das equipas de avaliação constituídas por parcerias internacionais semelhantes, sobretudo na contratação de avaliações multinacionais, regionais e sectoriais nos anos que precederam a adesão (ver exemplo da Caixa *Desenvolvimento das capacidades de avaliação na República Checa*).

Caixa: Desenvolvimento das capacidades de avaliação na República Checa

Na República Checa, existia uma base de peritos, sobretudo na comunidade académica, à qual se recorria para efeitos de avaliação. Porém, eram inexistentes as empresas checas especializadas na oferta de avaliação, as ofertas das empresas estrangeiras eram demasiado dispendiosas, e o conhecimento do ambiente checo era inadequado. No domínio prático, muitas entidades domésticas tinham experiência na preparação de planos, programas, metodologias, e redacção de políticas para apoiar o desenvolvimento socioeconómico. Entre estes, encontram-se o Plano de Desenvolvimento Nacional e os Programas Operacionais. Esta experiência prática de uma área de avaliação relacionada revelou-se extremamente útil enquanto base para o desenvolvimento de capacidades de avaliação.

Considerando os custos elevados da avaliação, a necessidade de um conhecimento do ambiente checo, e os requisitos relacionados com o desenvolvimento da capacidade de avaliação na República Checa, ao encomendar avaliações ex-ante a avaliadores externos, foi dada preferência aos contratantes domésticos. Em termos de avaliação, estes contratantes eram novos em campo, comparativamente aos seus homólogos estrangeiros, cujos serviços não podiam ser utilizados, sobretudo porque eram demasiado dispendiosos, mas também porque nem sempre havia um conhecimento suficientemente sólido do contexto checo. Em muitos casos, as empresas checas contratadas para executar a avaliação recrutavam um perito estrangeiro para transmitir conhecimentos sobre os processos de avaliação geralmente aplicados de uma forma informal, e para garantir o cumprimento com as normas europeias. Estes consórcios cumpriam, assim, com o que se esperava das suas actividades práticas. Uma conclusão positiva foi o facto de as operações das empresas checas, que não integravam quaisquer peritos externos, obterem

praticamente os mesmos resultados.

Em: Avaliação do Quadro Comunitário de Apoio Checo: Mario Hladek, David Grolig (com Eva Písova, Chefe da Unidade de Avaliação) Praga, Setembro de 2004

Alguns dos novos consultores serão provenientes de (ou estarão parcialmente baseados em) universidades, mas outros não. Também é uma característica das novas abordagens da gestão pública que este processo se encontre associado à criação de uma série de novos mercados para serviços de consultoria. Nessa medida, a promoção do sector da consultoria independente possui vantagens mais amplas do que a avaliação considerada isoladamente.

Existem diversas abordagens para incentivar a emergência de consultores nacionais. Entre estas, incluem-se:

- Requerentes de avaliação que insistem em orçamentos de consórcios ou parcerias que incluem sempre alguns consultores locais;
- A adaptação de contratos de avaliação de forma que avaliações de risco relativamente reduzido podem ser executadas por operadores nacionais, novos no mercado da avaliação;
- Garantia de que os requisitos financeiros e de aprovisionamento associados aos concursos de avaliação não sejam demasiado restritivos (são necessárias respostas em períodos de tempo reduzidos, garantias bancárias e diversos anos de contas certificadas por auditorias);
- Ênfase em critérios técnicos e de *know-how* na elaboração dos TdR, e não em procedimentos administrativos complexos, com os quais os consultores com menos experiência podem não estar familiarizados;
- Reuniões de trabalho com potenciais consultores para esclarecer dúvidas e incentivar a participação em concursos num ambiente competitivo;
- Apoio à colaboração entre consultores de avaliação relativamente isolados, de modo a incentivar o trabalho em equipa, a formação de consórcios e outras redes e associações profissionais (ver abaixo).
- Reconhecimento, por parte dos requerentes de avaliação, de que poderão necessitar de uma gestão mais prática dos novos contratantes para acelerar a sua aquisição de conhecimento tácito, que os consultores experientes já possuem.

Uma comunidade de avaliação profissional

É do interesse dos que requerem e executam avaliações que trabalhem num enquadramento que incentive elevados padrões de qualidade. Isto é, normalmente, o que se compreende quando se descreve a avaliação como uma profissão. As normas profissionais são essenciais para que os avaliadores sejam vistos como credíveis, e também para que seja garantida aos avaliadores a independência necessária por parte dos que requerem e utilizam as avaliações. Este aspecto é particularmente importante, face aos diferentes e, muitas vezes, fortes interesses envolvidos na avaliação de programas e de políticas. Muitos avaliadores já estiverem sujeitos a tentativas de interferência na forma como as avaliações são executadas. Esta interferência pode assumir diversas formas:

- TdR desequilibrados;
- Influência indevida num processo de avaliação pelos responsáveis pelo programa em avaliação (e.g., uma Comissão de Gestão dos Fundos Estruturais);
- Tentativas de seleccionar previamente uma equipa de avaliação, de modo a incluir elementos que poderão apresentar o que se consideram ser as respostas mais adequadas;

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- Pressão para reformular os resultados e as conclusões, de modo a serem menos críticos e mais aceitáveis para os principais parceiros;
- A disponibilização de dados limitados ou seleccionados.

Na realidade, ninguém está interessado em que isto aconteça. Para uma administração moderna, as vantagens de avaliações imparciais e fiáveis compensam largamente as aparentes vantagens de relatórios conservadores e prudentes. Contudo, este aspecto só se torna evidente com o desenvolvimento de uma cultura de avaliação.

Assume-se por norma que as orientações profissionais no campo da avaliação incluem um compromisso entre os avaliadores e os requerentes, para que:

- Se aplique os melhores métodos ou os mais adequados;
- Se promova o auto-desenvolvimento e actualização de competências;
- Se tenha em particular atenção os interesses dos parceiros e dos responsáveis políticos, incluindo os gestores de programa;
- Se assuma uma postura independente, que seja imparcial e verdadeira em relação às provas recolhidas;
- Se tenha um comportamento ético para com os diversos intervenientes: requerentes, fornecedores de informações e de dados, potenciais beneficiários de programas e de políticas.

Existem várias fontes de independência e capacitação profissionais. Estas incluem:

- Uma cultura de avaliação que reforce os padrões de conduta profissionais;
- Respeito pela independência, por parte dos requerentes de avaliação;
- Códigos de ética em relação aos quais exista um consenso, e que sejam, por conseguinte, amplamente reconhecidos e divulgados;
- Instituições independentes onde os avaliadores trabalham, que poderão conceder às suas considerações um maior peso, permitindo-lhes resistir a pressões externas;
- Educação e formação de elevada qualidade, bem como desenvolvimento profissional;
- Associações e sociedades profissionais, que integrem todos os que estão envolvidos na comunidade da avaliação, onde é possível partilhar experiências e discutir problemas práticos.
- O desenvolvimento e aplicação de normas de qualidade por parte dos profissionais e requerentes de avaliação – como, por exemplo, as que foram indicadas na secção anterior deste GUIA (ver a Caixa *Normas de avaliação*).

Caixa: Normas de avaliação

[O estudo]... tende a apoiar o valor de normas abrangentes que incluem os interesses dos requerentes da avaliação, dos avaliadores e dos cidadãos. As normas de avaliação europeias mais abrangentes (instrumentais e éticas), conforme utilizadas pela *European Evaluation Society* e por várias outras associações de avaliação nacionais na Europa, poderão complementar a passagem para normas desenvolvidas pela Comissão Europeia e por alguns Governos nacionais.

A Utilização da Avaliação nos Serviços da Comissão. Direcção-Geral do Orçamento, Outubro de 2002

Na prática, muitas destas características desejáveis de uma comunidade de avaliação profissional poderão ser incentivadas ao abrigo de uma sociedade nacional de avaliação ou associações. (Estão disponíveis diversos materiais sobre sociedades de avaliação nacionais e regionais em diversos websites de avaliação internacionais. Ver, em particular, o *IOCE - International Organisation for Cooperation in Evaluation* e o *IDEAS* -

International Development Evaluations Association). Estas reúnem por norma pessoas de diferentes instituições (académicas, administração pública, centros de investigação e consultoras), com diferentes especializações disciplinares (sociólogos, economistas, planeadores, especialistas em ciências políticas, pedagogos, etc.) e diferentes campos de especialização (serviços sociais, educação, investigação e desenvolvimento e, claro, desenvolvimento socioeconómico). Estas sociedades disseminaram-se em toda a Europa ao longo dos últimos anos (ver a Caixa [Lista de sociedades e associações de avaliação](#) onde se apresentam algumas das actuais sociedades europeias de avaliação). Estas sociedades proporcionam uma oportunidade para a germinação cruzada de ideias, a discussão de problemas partilhados fora do contexto de contratos de avaliação particulares, a disseminação de boas práticas e a emergência de normas e códigos de ética comuns. Como referido na Parte 1 deste GUIA, a maioria das normas de avaliação e códigos de ética emergentes nos últimos anos decorreram do trabalho das sociedades e associações de avaliação (ver a Caixa [Lista de sociedades e associações de avaliação](#)).

Caixa: Lista de sociedades e associações de avaliação

- Sociedade Europeia de Avaliação: <http://www.europeanevaluation.org/>
- Sociedade Dinamarquesa de Avaliação: <http://www.danskevalueringsselskab.dk/>
- Sociedade Finlandesa de Avaliação: <http://www.finnishevaluationsociety.net/>
- Sociedade Francesa de Avaliação: <http://www.sfe.asso.fr/>
- Sociedade Alemã de Avaliação: <http://www.degeval.de/>
- Sociedade Italiana de Avaliação: <http://www.valutazioneitaliana.it/>
- Sociedade Espanhola de Avaliação: http://www.idr.es/idr_in/index_intro.html
- Sociedade de Avaliação do Reino Unido: <http://www.evaluation.org.uk/>

3.8. A institucionalização da avaliação

Grande parte deste GUIA dedica-se à iniciação da capacidade de avaliação, apontando os motivos da sua importância para o desenvolvimento socioeconómico e para uma adequada gestão pública. A institucionalização está mais relacionada com a integração e o aprofundamento da avaliação nas instituições do sector público. Isto implica normalmente:

- alargar a avaliação no sector público, e
- obter uma maior integração dos processos de avaliação e ampliar a sua aplicação nos processos de decisão política e na gestão de programas (ver a Caixa [Definição de institucionalização](#)).

Caixa: Definição de institucionalização

A institucionalização da avaliação é vista, em muitos casos, quer como estando assente em regras, quer como um processo. Segundo Boyle, Lemaire e Rist, é com base nas regras, nos procedimentos e nas disposições organizacionais que as avaliações de políticas públicas são produzidas, e é com base no processo de legitimação que as práticas da avaliação se tornam formalmente parte dos processos de tomada de decisão das administrações nacionais.

Building Effective Evaluation Capacity, Boyle and Lemaire (Eds), Transaction Publishers,

1999

É provável que a introdução da avaliação se concentre, em primeiro lugar, numa parte do governo, num único ministério ou num conjunto de programas. O alargamento da avaliação permite estender as práticas de avaliação a outros ministérios ou programas e a outras instituições. Em sistemas de avaliação desenvolvidos, a avaliação é utilizada por parlamentos e organismos de auditoria nacionais, bem como por ministérios, decisores políticos e gestores de programa.

Uma maior integração da avaliação nos processos de decisão política e na gestão de programas poderá envolver:

- O planeamento conjunto de políticas e de avaliações com base em ciclos plurianuais de consulta e de estabelecimento de prioridades por parte de agências públicas, levando à inclusão da avaliação em planos plurianuais ou departamentais;
- O acompanhamento sistemático dos resultados e recomendações da avaliação, tornando-se difícil para os gestores intermédios ignorar as avaliações em relação às quais terão de, pelo menos, justificar por que motivo estão ou não a seguir as recomendações;
- Incentivos positivos para que os gestores façam uso das conclusões das avaliações e modifiquem os programas e a sua orientação, se necessário;
- A criação de fóruns, onde os gestores, os decisores políticos e os gestores de avaliação se reúnem para discutir os resultados da avaliação e as lições daí retiradas;
- O alargamento da avaliação ao longo de todo o ciclo político, desde a identificação de opções ao planeamento, programação, implementação e reprogramação (consulte a Caixa *Planeamento do alargamento e aprofundamento da avaliação na Roménia*).

Caixa: Planeamento do alargamento e aprofundamento da avaliação na Roménia

Na Roménia, durante a preparação para a adesão, os esforços da avaliação concentraram-se inicialmente nas fases de financiamento, planeamento e concepção das novas políticas e programas europeus. Os esforços centraram-se igualmente nos programas PHARE e na preparação para os Fundos Estruturais. No entanto, elaboraram-se planos para desenvolver uma cultura de avaliação que seria relevante para todos os programas dos Fundos Estruturais financiados pela UE. O ministério incumbido da coordenação central encontrava-se paralelamente a organizar um programa de modernização e de reforma transversal a todos os ministérios. Esta medida estava ainda relacionada com planos para, numa fase posterior, alargar a avaliação a programas e políticas de financiamento nacional.

Fonte: Entrevistas a representantes romenos

A criação de uma cultura de avaliação é uma expressão que tem como objectivo deixar claro que um sistema de avaliação consolidado consiste em mais do que apenas uma organização estrutural. Por exemplo:

- Existe um compromisso para com a melhoria e a aprendizagem;
- Existe uma tendência para evitar uma cultura de culpabilização, que desincentiva a aprendizagem;
- Os decisores políticos estão comprometidos com políticas assentes em provas concretas, no sentido lato, e de forma mais ambiciosa;
- Existe um empenho na excelência, na manutenção de elevados padrões de qualidade e na melhoria contínua;

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- A avaliação é utilizada como um elemento num processo geral, que visa a transparência e a múltipla responsabilização dos cidadãos e das comunidades, bem como dos Ministros e dos parlamentos;
- O governo e o sector público estão empenhados numa adaptação contínua, no sentido de se transformarem numa organização de aprendizagem (ver a Caixa *Avaliar o desenvolvimento da capacidade de avaliação*).

Caixa: Avaliar o desenvolvimento da capacidade de avaliação

Os autores de um estudo comparativo internacional da capacidade de avaliação utilizaram os nove critérios que se seguem para avaliar a dimensão da capacidade de avaliação:

1. A avaliação decorre em diversos domínios da política;
2. Existe uma oferta de avaliadores locais, em diferentes áreas disciplinares;
3. Existe um discurso nacional sobre a avaliação;
4. Existem organizações profissionais de avaliadores;
5. Existe um determinado grau de institucionalização da avaliação a nível do governo;
6. Existe um determinado grau de institucionalização no Parlamento;
7. Existem diversas instituições de avaliadores a efectuar avaliações, em cada um dos domínios de política;
8. Existe uma função de avaliação na instituição de auditoria nacional;
9. Existe um equilíbrio entre as avaliações de resultados e as avaliações de realizações e de processos.

International Atlas of Evaluation, Editors Furubo, Rist and Sandahl, Transaction Publishers, 2002

O Canadá é largamente reconhecido como um país bem sucedido na institucionalização da avaliação nos seus sistemas de gestão pública e de decisão política, a nível federal (i.e., nacional). A responsabilidade geral de coordenação da avaliação reside no *Treasury Board* (Conselho do Tesouro Público) do Canadá, que deverá indicar as orientações centrais para a avaliação do Governo do Canadá, e, para o efeito, deverá (ver a ligação: http://www.tbs-sct.gc.ca/pubs_pol/dcgpubs/TBM_161/ep-pe1_e.asp#_Toc505657338):

1. Criar um Centro de Excelência, no sentido de proporcionar liderança, orientações e apoio à prática da avaliação;
2. Utilizar os resultados da avaliação, sempre que necessário, na tomada de decisões do centro;
3. Definir normas; e
4. Monitorizar a capacidade de avaliação do governo.

Alguns dos traços diferenciadores da política de avaliação canadiana residem no facto de:

- Basear-se, em grande parte, na devolução das responsabilidades para desenvolver e utilizar avaliações junto de departamentos governamentais (ministérios) e de gestores nesses departamentos.
- Centrar-se numa gestão com base nos resultados como um componente essencial da gestão pública, o que difere, por exemplo, de um enfoque da responsabilização voltado para o passado, embora também destaque a responsabilização.

Ver a Caixa *Excertos da política de avaliação: Treasury Board (Conselho do Tesouro Público) do Canadá*.

Caixa: Excertos da política de avaliação: *Treasury Board* (Conselho do Tesouro Público) do Canadá

Esta política assenta em três princípios fundamentais:

1. O princípio de que alcançar e reportar com precisão os resultados é uma responsabilidade primária dos gestores da administração pública;
2. O princípio de que uma avaliação rigorosa e objectiva é uma ferramenta importante para ajudar os gestores na gestão dos resultados; e
3. O princípio de que os departamentos com o apoio do *Treasury Board Secretariat* (Secretariado do Conselho do Tesouro Público) são responsáveis por assegurar a devida implementação da avaliação, com rigor e disciplina, dentro da sua jurisdição.

A gestão orientada para os resultados

A gestão orientada para os resultados é a principal responsabilidade dos gestores da administração pública. Os gestores devem ser responsáveis pelo seu desempenho perante a equipa de gestão hierarquicamente superior, perante os ministros, perante o Parlamento e perante os canadianos.

A avaliação como ferramenta de gestão

A avaliação decorre num ambiente complexo, que inclui parcerias com outras organizações federais, com outros níveis de governo, com o sector privado ou com entidades sem fins lucrativos. Além disso, à medida que a implementação dos programas passa para outras jurisdições, a importância da avaliação das políticas subjacentes aumenta. A avaliação deverá contribuir para implementar melhorias nas políticas, bem como para a concepção e implementação de programas.

Neste contexto, os departamentos deverão integrar a disciplina da avaliação na gestão do ciclo de vida das políticas, dos programas e iniciativas para:

- desenvolver quadros de responsabilização e de gestão com base nos resultados, com vista a obter políticas, programas e iniciativas novos ou melhorados;
- implementar uma monitorização contínua do desempenho e práticas de medição do desempenho;
- avaliar questões relacionadas com as fases iniciais de implementação e de administração das políticas, programas ou iniciativas, incluindo os que são desenvolvidos através de disposições em parceria (avaliação formativa ou intercalar); e
- avaliar questões relacionadas com a relevância, os resultados, e a eficácia em termos de custos.

Desenvolveram-se diversas ferramentas de diagnóstico para avaliar o grau de institucionalização da avaliação e para proporcionar orientações relativamente à forma de se reforçar este aspecto. Na Lituânia, por exemplo, as tipologias da OCDE foram adaptadas de modo a avaliar o grau de institucionalização da avaliação. Entre os factores considerados por Nakrosis e por membros do *Public Policy and Management Institute* (Instituto de Gestão e Políticas Públicas) encontram-se:

- Aplicação atempada da avaliação em todas as fases da implementação da política;

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- Grau de centralização;
- Localização interna ou externa dos avaliadores;
- Integração com o processo de tomada de decisão.

Esta aplicação de padrões permite a identificação de forças e fraquezas, e um plano de optimização.

O Banco Mundial publicou um quadro de referência para a avaliação de opções sob diferentes condições. Baseia-se no reconhecimento de que, em qualquer contexto das administrações nacionais, as condições de partida para o desenvolvimento das capacidades de avaliação irão revelar variações significativas. Nalguns casos, a procura (sobretudo a procura interna) é mais forte, e, noutros, mais fraca. Nalguns casos, a disponibilidade de capacidades e de redes de instituições capazes de mobilizar essas capacidades terá uma maior capacidade de responder à procura de avaliação do que noutros. Independentemente do equilíbrio entre a procura e a oferta, existem estratégias adequadas que é possível adoptar para reflectir as circunstâncias prevaletentes. O quadro retirado do relatório do Banco Mundial (1999) apresenta uma síntese de algumas destas diferentes opções (ver a Caixa Menu de algumas opções ECD - Evaluation Capacity Development - Desenvolvimento da Capacidade de Avaliação).

Caixa: Menu de algumas opções ECD (Evaluation Capacity Development - Desenvolvimento da Capacidade de Avaliação).

		PROCURA	
		Forte	Fraca
O F E R T A	F O R T E	<ul style="list-style-type: none"> • apoiar a avaliação de políticas, programas e projectos • estabelecer ligações entre a avaliação, o planeamento estratégico, distribuição de recursos e orçamentos • recorrer a comissões de peritos para avaliar as políticas • reforçar a avaliação na legislatura (parlamento) • divulgar os resultados das avaliações junto do público • organizar e sistematizar a função de avaliação • apoiar sistemas de monitorização financeira e de informação 	<ul style="list-style-type: none"> • divulgar métodos e práticas de avaliação • apoiar a avaliação contínua de programas e projectos • participar em avaliações executadas por agências de financiamento externas • apoiar o desenvolvimento profissional em avaliação • construir uma rede de avaliação no seio do governo e/ou a nível regional • organizar seminários de ECD (Desenvolvimento da Capacidade de Avaliação) para funcionários superiores, com vista à construção de consensos • mandar a avaliação através de decretos e regulamentos
	F R A C A	<ul style="list-style-type: none"> • divulgar lições aprendidas com experiências passadas e boas práticas de avaliação • estabelecer comissões para avaliar projectos e programas importantes • dar formação a e recorrer a organizações privadas em exercícios de avaliação • apoiar a formação universitária em avaliação, incluindo cursos de curta duração ou seminários • prestar assistência técnica a agências governamentais • construir uma rede de avaliação dentro do governo • promover a avaliação por ONGs • introduzir normas de avaliação 	<ul style="list-style-type: none"> • reforçar a auditoria e a contabilidade • realizar avaliações conjuntas com as agências financiadoras • divulgar experiências nacionais e internacionais de avaliação • apoiar a formação em avaliação em instituições de ensino • promover a cooperação transnacional na avaliação • consciencializar os decisores

A institucionalização da avaliação constitui um processo contínuo, que necessita de ser integrado, ao longo do tempo, com outras ferramentas de planeamento e avaliação (por exemplo, avaliação de impactos, técnicas de planeamento de projecto) e outras vias de identificação, divulgação e implementação de boas práticas (por exemplo, consulta pública, descentralização nos parceiros locais).

Ao mesmo tempo, não se pode ignorar potenciais obstáculos à institucionalização da avaliação:

- Em primeiro lugar, conforme se tem vindo a destacar nesta secção do GUIA, a avaliação é mais sólida quando vista como parte integrante do desenvolvimento institucional e da reforma do sector público. Se não se cumprir com estas alterações organizacionais e culturais em simultâneo, será mais difícil institucionalizar a avaliação.
- Em segundo lugar, um processo de institucionalização coerente exige recursos financeiros e humanos. A institucionalização pode ser minada se houver um investimento insuficiente nas capacidades e nos sistemas de gestão das unidades de avaliação especializadas.
- Em terceiro lugar, o empenhamento político de alto nível é igualmente essencial no processo de institucionalização da avaliação. Este é particularmente o caso quando o processo inevitavelmente é moroso e necessita de ser desenvolvido por fases. As alterações de direcção sempre que ministros ou oficiais mudam podem constituir um grande obstáculo ao desenvolvimento bem sucedido das capacidades de avaliação.

Como noutras secções do GUIA, apresentamos de seguida um resumo das principais lições retiradas, em termos de regras de ouro, que se destinam a todos os que procuram reforçar as capacidades de avaliação.

3.9. Regras de ouro

1. A avaliação terá maior utilidade e aplicação quando estiver plenamente integrada enquanto função em sistemas de apoio administrativo e institucional, que procuram igualmente incluir parceiros da sociedade civil. É isto que se entende como capacidade de avaliação.
2. A capacidade de avaliação poderá desenvolver-se organicamente ao longo do tempo mas, na maioria dos contextos nacionais e internacionais, necessita de ser devidamente planeada, conscientemente introduzida e ser provida dos recursos adequados.
3. O princípio da independência e da imparcialidade constitui um princípio essencial a integrar nas funções de avaliação, considerando os interesses adquiridos que podem influenciar a execução das avaliações e as suas conclusões.
4. A capacidade de avaliação é constituída por elementos a nível individual, organizacional, inter-organizacional e societal, que deverão reforçar-se mutuamente. A capacidade de avaliação deverá ser concebida como um sistema.
5. As actividades de avaliação devem ser coordenadas de uma forma transversal a todos os níveis das administrações, de modo a garantir a coerência e a proporcionar uma base de sustentação para o desenvolvimento de capacidades e para a profissionalização. Ao mesmo tempo, os procedimentos e as abordagens deverão ser suficientemente flexíveis para se adaptarem às especificidades das políticas, dos programas e dos territórios.
6. A capacidade de avaliação poderá começar por centrar-se em programas específicos, como por exemplo os programas financiados pela UE ou os programas que decorrem numa parte específica do governo (por exemplo,

departamentos do governo central), mas deve ser concebida de modo que as actividades de avaliação possam ser alargadas e incluir, por exemplo, programas com financiamento nacional ou regional, bem como actividades do governo central.

7. A avaliação deverá ser dotada de uma configuração adequada, para satisfazer necessidades de gestão e de tomada de decisão reais, que irão melhorar a eficácia e a eficiência da acção pública, e que representam vantagens concretas para os cidadãos e parceiros. Embora a avaliação seja inevitavelmente útil para satisfazer necessidades de responsabilização externas, nunca deverá ser vista como uma obrigação principalmente externa.
8. A inclusão da sociedade civil e a participação na avaliação, seja por consulta, seja por representação directa, irá incentivar diferentes interesses (ex. representantes da comunidade, grupos profissionais e sector privado) de procura e utilização de avaliação, reforçando, assim, a responsabilização, aumentando a transparência e reforçando o capital social.
9. A capacidade de avaliação deverá prestar atenção à vertente de oferta, promovendo uma comunidade profissional e especializada de agentes, e às instituições, redes e mercados que garantam a colocação de oportunidades de avaliação a concurso e a sua execução de acordo com elevados padrões de qualidade.
10. Entre os principais elementos de uma comunidade de avaliação profissional, encontram-se a educação e formação de avaliadores, bem como elevados padrões de qualidade e códigos de ética acordados e partilhados entre os requerentes e os profissionais da avaliação, incluindo consultores e académicos. Estas qualidades da oferta desenvolvem-se melhor, muitas vezes, através de sociedades de avaliação nacionais, abertas à participação, e comprometidas com elevados padrões de qualidade.

(Índice da Parte IV)

4. PARTE IV: A SELECÇÃO DE MÉTODOS E TÉCNICAS

4.1. Introdução

4.2. Factores que influenciam a escolha de métodos e técnicas

4.3. As opções para diferentes tipos de intervenção

4.4. As opções para diferentes objectivos de avaliação

4.5. As opções para diferentes fases de programas/políticas

4.6. A recolha e utilização de dados em avaliação

4.7. Criação de indicadores e sistemas de indicadores

4.8. Utilização de indicadores para melhorar a gestão

4.9. Regras de ouro

4.1. Introdução

Esta secção do GUIA descreve os métodos e as técnicas disponíveis para a avaliação do desenvolvimento socioeconómico. Os métodos e técnicas específicos são descritos no Manual Técnico II. Numa secção anterior do GUIA, *A Concepção e o Planeamento da Avaliação*, a concepção da avaliação foi considerada em termos gerais. Aqui, centrar-se-á nas escolhas que são necessárias adoptar, quer em termos de categorias abrangentes de métodos, quer em termos de técnicas específicas dentro dessas categorias. A aplicação de métodos e técnicas levanta inevitavelmente questões relativamente à origem de dados e de provas concretas nas quais se baseiam os avaliadores. Aqui, consideram-se essas fontes de dados, não só em termos de como poderão ser analisadas, mas também em termos de como poderão ser localizadas ou inclusivamente criadas, no caso de não existirem previamente. Esta secção do GUIA começa por considerar: o contexto em que decorre a selecção de métodos e técnicas, a distinção (algo difusa) entre dados quantitativos e dados qualitativos, e a forma de obtenção e utilização dos dados concretos. As subsecções seguintes descrevem:

- Os métodos e as técnicas aplicáveis aos diferentes tipos de intervencções socioeconómicas, incluindo as principais prioridades temáticas e as diferentes áreas de política;
- Os métodos e técnicas dos diferentes objectivos da avaliação (planeamento, responsabilização, implementação, produção de conhecimento, capacitação institucional);
- Os métodos e técnicas aplicáveis às diferentes fases dos programas e das políticas, desde a formulação de políticas até aos impactos;
- Os métodos e técnicas aplicados nas diferentes fases do processo de avaliação (o Manual Técnico II apresenta os métodos e técnicas desta forma)
- A recolha e a utilização de dados com diferentes características na avaliação;
- Criação de indicadores e sistemas de indicadores;
- Utilização de indicadores para melhorar a gestão;
- Regras de ouro.

4.2. Factores que influenciam a selecção de métodos e técnicas

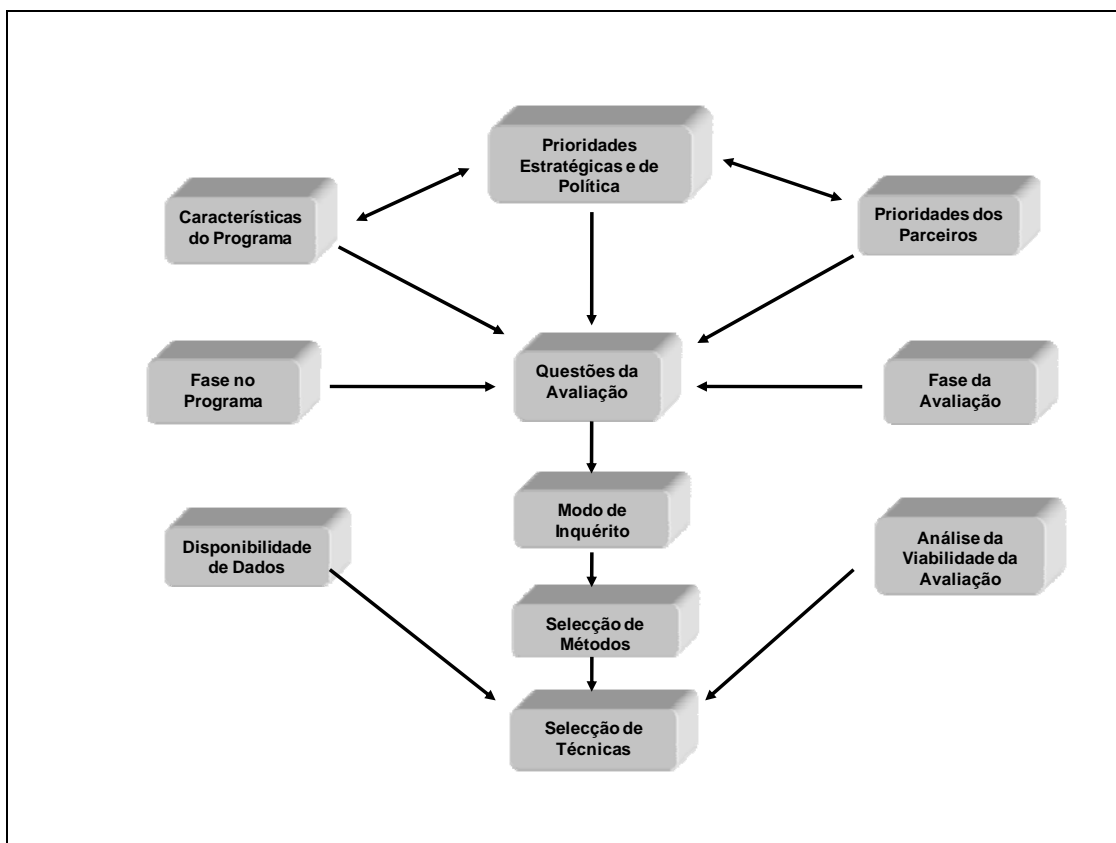
- A selecção de métodos e técnicas

- Quantitativo vs. qualitativo: uma falsa questão?
- A recolha e utilização de dados e de provas

A selecção de métodos e técnicas

Como já se realçou nas secções do GUIA sobre a concepção e a implementação da avaliação e a capacitação institucional para a avaliação, é necessário tomar uma série de decisões na concepção da avaliação. É necessário consultar os parceiros, mapear a lógica do programa, identificar as questões da avaliação, seleccionar critérios, e analisar a viabilidade da proposta. A selecção de métodos e técnicas consiste, por conseguinte, numa componente essencial em todo este processo. A Caixa A selecção de métodos num contexto mais alargado posiciona a selecção de métodos e técnicas neste contexto mais alargado.

Caixa: Selecção de métodos num contexto mais alargado



Esta secção do GUIA permite preencher muitas das células da caixa apresentada acima, uma vez que este é o único modo de garantir que a selecção de métodos e técnicas tenha em consideração todos os factores importantes.

Os métodos decorrem da escolha de uma concepção de avaliação ou modo de inquérito: estes devem responder a determinados tipos de questões, e só devem ser seleccionados se puderem responder a estas questões. Este aspecto poder parecer óbvio, mas um dos problemas da prática da avaliação reside na tendência que os avaliadores e os requerentes têm para privilegiar determinados métodos, independentemente de estes serem adequados em termos de proporcionarem respostas às questões colocadas. Referiu-se, na secção do GUIA sobre a capacitação para a avaliação, que é uma boa

prática os requerentes da avaliação proporcionarem espaço aos avaliadores para que estes possam determinar o seu método de abordagem preferido, e, de facto, para que as primeiras fases de uma avaliação permitam elaborar um relatório inicial, com a revisão e discussão da concepção, da metodologia, das técnicas, da recolha de dados, etc.. Infelizmente, esta flexibilidade nem sempre é assegurada. No entanto, presumimos, nesta secção do GUIA, que os avaliadores terão margem de manobra suficiente para seleccionar métodos e técnicas adequados, e que os requerentes da avaliação serão informados, de acordo com critérios e conhecimentos semelhantes, acerca de quais os métodos adequados a cada um dos objectivos.

Após a selecção de uma concepção de avaliação abrangente, a selecção de métodos e técnicas específicos terá que considerar, ainda assim, as realidades dos programas e das políticas, bem como uma série de factores contextuais. Por exemplo:

- A forma da intervenção socioeconómica. Referiu-se, na secção do GUIA relacionada com a avaliação e o desenvolvimento socioeconómico, que as características de um objecto de avaliação, neste caso um tipo de intervenção de desenvolvimento socioeconómico, são factores importantes que determinam a concepção da avaliação.
- Tipo de objectivos da avaliação. As avaliações podem ter objectivos bastante diferentes, da responsabilização à optimização da gestão e à explicação do que efectivamente funciona, e em que circunstâncias. Estes diferentes objectivos, associados a diferentes formas de estudo, também possuem implicações para os métodos e técnicas seleccionados.
- Diferentes programas/políticas. Uma avaliação possui requisitos diferentes numa fase de avaliação inicial, ex-ante, numa fase intercalar ou intermédia, durante a implementação de um programa, e numa fase ex-post, após a sua conclusão. Cada uma destas fases poderá exigir diferentes métodos e técnicas.
- Diferentes fases do processo de avaliação. Numa mesma avaliação é necessário conceber, recolher dados, analisar os resultados e elaborar as conclusões. Cada uma destas actividades está associada a diferentes métodos e técnicas.

Não se pretende com isto dizer que existe uma correlação directa entre estas diferentes circunstâncias e contextos com que os avaliadores irão debater-se. Porém, existem alguns *clusters*, constituídos por métodos e técnicas, associados aos contextos referidos acima, que podem funcionar como pontos orientadores úteis para os avaliadores e para os requerentes de avaliação.

É necessário, no entanto, prestar particular atenção aos dois aspectos que se seguem:

- Todos os métodos e técnicas possuem forças e fraquezas; muitas vezes, utilizam-se em circunstâncias que estão longe de ser as ideais para a sua aplicação. Em qualquer processo de avaliação, as técnicas devem ser seleccionadas e aplicadas de modo a explorar as suas virtudes e a reconhecer as suas limitações.
- Face ao exposto, o melhor é aplicar conjuntamente os métodos e técnicas, integrados em quaisquer trabalhos de avaliação específicos. Basear-se num único método ou técnica poderá ser mais frágil do que obter diversas perspectivas (o que se designa, muitas vezes, por triangulação).

Quantitativo vs. qualitativo: uma falsa questão?

É comum fazer uma distinção entre métodos, técnicas e dados quantitativos e qualitativos. De facto, esta distinção não é tão clara como pode parecer à primeira vista. Quando os dados qualitativos das pessoas são classificados e adicionados, tornam-se

quantitativos, como, por exemplo, “50% dos entrevistados afirmou ter usufruído do programa”. De facto, a base de sustentação de muitas avaliações quantitativas é qualitativa. A análise de dados administrativos exige, muitas vezes, considerações de ordem qualitativa relativamente à forma de classificar, por exemplo, grandes, pequenas ou médias empresas. Do mesmo modo, os inquéritos por correio integram dados qualitativos.

No entanto, não deveremos subestimar o rigor, muitas vezes integrado em procedimentos e protocolos analíticos de aceitação mais generalizada, necessário para converter contributos qualitativos em conclusões quantitativas. É por isso que a amostragem, a significância estatística e a distinção entre diferentes tipos de análise de dados, entre muitas outras convenções, são essenciais para fazer avaliações quantitativas genuínas.

Os limites entre os métodos quantitativos e qualitativos tornam-se ainda mais difusos quando se trata de distinguir entre métodos de recolha de dados e métodos de os analisar. Os dados recolhidos podem ser qualitativos (e.g., entrevistas, inquéritos e observações), e, ainda assim, analisados quantitativamente. Diversos modelos estatísticos, por exemplo, utilizam dados de entrevistas ou inquéritos como recursos. E só é possível compreender na íntegra as análises quantitativas depois de se fazer interpretações qualitativas do significado dos resultados.

A natureza do desenvolvimento socioeconómico num contexto europeu é em termos genéricos ascendente, com uma combinação de intervenções concebidas à medida das necessidades específicas dos territórios ou sectores, difíceis de descrever de acordo com categorias padronizadas. Isto coloca limites às avaliações quantitativas, que procuram fornecer medidas comparativas simples (normalmente, indicadores) ou contagens de resultados e efeitos. A aplicação de indicadores ou de outras medidas padronizadas não poderá proporcionar comparações entre os diversos programas locais e sectoriais. Devido à natureza altamente contextualizada do desenvolvimento socioeconómico, os métodos quantitativos mais eficazes são técnicas e modelos estatísticos que reconhecem a importância do contexto. Estes têm, por exemplo, que assumir como base de comparação uma configuração ao longo do tempo (i.e., o mesmo território em 2002 e, depois, em 2006), e uma configuração no seu contexto regional (ou, possivelmente, tendo em conta outras unidades territoriais correspondentes). Estas técnicas – geralmente, em forma de modelos prospectivos, modelos macroeconómicos ou análises de resultados multivariados – deverão ter a capacidade de avaliar as diferenças entre os resultados esperados e reais do desenvolvimento. Deverão oferecer respostas a questões como: as tendências ao nível do emprego ou da produtividade das empresas são diferentes em áreas visadas pelos programas comparativamente a outras áreas semelhantes? A utilização de dados comparativos também é importante em termos da avaliação da deslocação: os efeitos positivos de uma configuração decorrem à custa de outros. Por exemplo, o desenvolvimento num determinado território absorveu simplesmente a inovação e o crescimento ou novas oportunidades para grupos marginalizados em territórios circundantes?

Em suma, as forças da avaliação quantitativa residem no facto de poder:

- *Permitir que se retire conclusões agregadas.* Os decisores políticos pretendem saber, durante o processo, se uma política ou programa está a funcionar. Os resultados agregados, como por exemplo, se houve mais pessoas a obter emprego em, digamos, 100 locais do programa, proporcionam material de suporte para essas conclusões. Contudo, as considerações agregadas não são necessariamente capazes de demonstrar que os programas, ou alguns dos seus aspectos em particular, são responsáveis por estas alterações.

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- *Permitir modelações explicativas ou prospectivas*. Existem diversas técnicas estatísticas ou de modelação sofisticadas que são úteis na avaliação, sobretudo como forma de explicar ou prever, embora raramente para definir, as relações causais subjacentes às diferenças. Os métodos experimentais e os modelos macroeconómicos baseiam-se, por isso, em dados quantitativos; porém, conforme realçado anteriormente, os métodos mais adequados são geralmente os que têm em consideração o contexto.
- *Proporcionar uma perspectiva global*, que informa a análise qualitativa subsequente. Com base em medidas descritivas agregadas globais, torna-se mais claro onde ocorreram subcategorias ou casos inesperados. Este aspecto direcciona a atenção para uma segunda fase de análise, muitas vezes qualitativa.
- *Permitir que se façam estimativas de dimensão e de escala*. Sempre que estiverem disponíveis dados adequados, uma avaliação quantitativa permite calcular a dimensão das mudanças ocorridas no âmbito de uma intervenção. Este factor é particularmente importante sempre que for necessário avaliar se as vantagens compensam os custos inerentes aos recursos aplicados.
- *Permitir que se estabeleça um determinado grau de comparação entre os contextos*. Os decisores políticos necessitam de compreender se existem diferentes graus de eficácia das políticas, nos diferentes locais de intervenção. As comparações básicas tornam-se mais fáceis se estas puderem ser quantificadas, embora, no caso do desenvolvimento socioeconómico, se possa obter apenas formas de quantificação relativamente fracas, a não ser que sejam apoiadas por análises e modelações estatísticas.
- *Permitir que se realizem avaliações mais sólidas de intervenções específicas*. As avaliações quantitativas do desenvolvimento socioeconómico mais eficazes centram-se, em muitos casos, em intervenções específicas, vistas independentemente de um processo de desenvolvimento mais amplo que envolve múltiplos tipos de intervenção. Por conseguinte, as avaliações quantitativas de incentivos às empresas ou de intervenções no mercado de trabalho comportam fortes resultados em termos de efeitos dessas intervenções específicas.
- *Permitir que se façam análises de tendências ao longo do tempo*. As medidas quantitativas ao longo do tempo, por exemplo, recolhendo indicadores comuns numa base regular, poderão ajudar a monitorizar as alterações e permitir que se faça o acompanhamento do processo de desenvolvimento.

Alguns métodos e técnicas são mais obviamente qualitativos ou, pelo menos, estão normalmente associados à recolha de dados qualitativos. As entrevistas; a observação participante ou os estudos etnográficos; os diários de auto-relatório; a análise de discurso ou de conteúdo de textos; as descrições pormenorizadas de um contexto local, com o objectivo de comunicar uma determinada disposição ou *ethos*, fariam todos parte dessa categoria. O mesmo aconteceria com métodos compósitos, como os estudos de caso, que têm tendência para se basear na maioria das técnicas específicas referidas. Nesta secção do GUIA, faz-se referência a muitos destes métodos qualitativos, entre outros apresentados no Manual Técnico II. No entanto, a lógica dominante por detrás da selecção de métodos não reside na suposta superioridade de um tipo ou outro de técnica; antes, reside na sua adequação aos objectivos.

Os métodos qualitativos para a recolha e a análise de dados são importantes para o desenvolvimento socioeconómico, porque:

- *Estamos interessados em processos subtis*. A qualidade das oportunidades de trabalho, a experiência de discriminação, uma tendência para a inovação, e a

eficácia das parcerias constituem fenómenos qualitativos subtis que é necessário captar de forma detalhada.

- *Estamos interessados em contextos.* Estes são constituídos por diversos factores diferentes (seja em termos geográficos, históricos, culturais, estruturas económicas, grupos sociais, disposições institucionais, clima, tendências do emprego, histórias de desenvolvimento do passado, etc.), e a forma como interagem em contextos de desenvolvimento específicos só podem ser descritos em termos qualitativos. Além disso, é necessário definir todo o processo de desenvolvimento em termos de contexto, com vista a usufruir de processos de aprendizagem transferíveis.
- *Estamos interessados em opiniões humanas.* Estas poderão ser constituídas pelas opiniões dos parceiros, cuja lógica de intervenção e teorias de programa os avaliadores possam pretender suscitar. Ou, então, poderão ser constituídas pelas opiniões e pelas experiências dos beneficiários-alvo do desenvolvimento socioeconómico.
- *Estamos interessados em percepções de tipo ascendente.* Entre estas, poderão incluir-se: as ambições de desenvolvimento de actores de base (pequenas empresas, autoridades municipais, associações profissionais) e as expectativas e experiências de intervenientes locais, num contexto de desenvolvimento local. Estas informações de tipo ascendente são difíceis de ajustar a categorias de tipo descendente, altamente diversificadas, e são necessários conhecimentos qualitativos.
- *Estamos interessados em explicar relações causais.* Para aprender com o desenvolvimento e para o replicar, é necessário compreender o que se passa dentro da caixa negra, para ir além dos recursos aplicados e das realizações. Caso contrário, poderemos ficar a saber o que funciona, mas não como ou porque funciona. Isto requer uma análise detalhada e, muitas vezes, qualitativa.
- *Estamos interessados nos impactos em diferentes grupos.* Os programas terão, muitas vezes, diferentes impactos de acordo com os diferentes grupos ou beneficiários-alvo. A divisão de populações agregadas em grupos que são normalmente de dimensões reduzidas permite-nos investigar esses impactos diferenciais.
- *Estamos interessados em categorias inovadoras.* O desenvolvimento é frequentemente incerto porque procura implementar algo de novo. Apenas ao examinar os detalhes específicos do que está a acontecer num contexto de desenvolvimento é que será possível identificar as categorias significativas em que os avaliadores têm de se concentrar. Mesmo que a avaliação possa recorrer a métodos quantitativos, esta fase deverá ser precedida por uma fase exploratória, que permite clarificar as categorias relevantes.

A recolha e utilização de dados e de provas

Como se deixou claro na secção do GUIA sobre a avaliação e o desenvolvimento socioeconómico, poderão existir, entre os avaliadores, diversas perspectivas relativamente à natureza dos dados e ao que constituem elementos de prova válidos. Os que acreditam na capacidade da observação objectiva (ex. os positivistas), terão uma perspectiva diferente dos dados e das provas do que os que estão mais preocupados com a forma como as percepções e a teoria influenciam as observações (ex. os construtivistas). Neste GUIA, adoptamos uma perspectiva pragmática. Já se assumiu uma tendência para a adopção de um quadro de referência realista que, embora valorize a observação, a investigação empírica e a medição desde que sejam passíveis de aplicação prática, também se centra nos diferentes contextos em que decorrem os fenómenos e nas teorias utilizadas para explicar estes fenómenos. Ao mesmo tempo, e em determinadas circunstâncias, também se reconhece a importância das filosofias

construtivistas, sobretudo em contextos de exclusão social e de desenvolvimento de tipo ascendente, em que deverá dar-se prioridade à experiência, aos interesses e às considerações dos participantes nos programas. Também não se elimina completamente algumas das lições que as grandes conquistas da ciência e da investigação positivistas alcançaram, com a sua ênfase no questionamento sistemático e na interpretação cuidadosa das provas.

Os cientistas gostam de utilizar o termo ‘dados’, fazendo frequentemente distinções entre os tipos de dados (i.e., a matéria prima da investigação) e as informações que têm de ser processadas para serem úteis. As provas levam-nos um passo mais à frente, ao ajustar as informações, transformando-as em formas que se podem considerar fiáveis, ou perspectivadas como sendo suficientemente fortes para serem tidas em consideração por utilizadores como os decisores políticos e os gestores de programas.

Nesta secção do GUIA, também se debatem diversas questões relacionadas com as provas e com os dados. Especificamente, os avaliadores utilizam tipos de dados altamente variáveis. Alguns dados existem previamente e são decorrentes de fontes administrativas (por exemplo, os registos de um serviço local público de emprego ou as receitas fiscais de uma determinada área). Um programa, através das suas actividades de monitorização, irá gerar outras fontes de dados. (De facto, a qualidade dos sistemas de monitorização, responsabilidade sobretudo dos gestores de programa, é essencial para uma avaliação bem sucedida). Porém, os próprios avaliadores terão que gerar alguns dados; por exemplo, ao modificar os sistemas de monitorização, entrevistar gestores de PME locais, ou analisar o conteúdo de anúncios publicitários em revistas empresariais.

A qualidade de muitas avaliações seria reforçada se fosse prestada maior atenção à utilização de todas as fontes de dados disponíveis. Contudo, os gestores de programas e os decisores políticos também devem estar conscientes da sua obrigação de implementar e disponibilizar fontes de dados que podem ser úteis aos avaliadores. Este aspecto também não pode ser deixado para o momento de requisição da avaliação. Como se referiu na secção do GUIA sobre a concepção e a implementação da avaliação, a implementação de sistemas básicos de dados deverá fazer parte do planeamento dos programas, bem como das avaliações.

4.3. As opções para diferentes tipos de intervenção

- Avaliação das prioridades temáticas
- Prioridades sectoriais e de política
- Desenvolvimento local e territorial

Avaliação das prioridades temáticas

Um tipo de intervenção específico, característico do desenvolvimento socioeconómico e incluído de forma transversal nos Fundos Estruturais e de Coesão Europeus, é descrito como uma prioridade temática. Isto ocorre sempre que existe uma estratégia global, que se espera inserir ou mais formalmente integrar numa diversidade de intervenções.

A avaliação de temáticas requer a agregação de provas provenientes de diversas intervenções específicas, para compreender as formas como as prioridades estratégicas estão a ser cumpridas. Isto poderá implicar a construção de índices que são posteriormente acrescentados em diversos contextos. Porém, é normalmente difícil encontrar indicadores significativos que possam funcionar em diferentes contextos. As avaliações temáticas possuem, por isso, uma grande probabilidade de procurar diferentes medidas em diferentes contextos, para depois avaliar qualitativamente o seu resultado

global. As avaliações qualitativas da medida em que as estratégias temáticas se adequam aos critérios de política também são úteis, incluindo, por exemplo, inquéritos junto dos beneficiários, entrevistas aos parceiros e estudos de caso das alterações verificadas.

Alguns exemplos destas prioridades temáticas, no contexto dos Fundos Estruturais, incluem:

- Igualdade de oportunidades. Este aspecto é, por norma, visto em termos de igualdade de oportunidades entre homens e mulheres, mas é aplicado cada vez mais a outras áreas de discriminação ou desigualdade, como sejam as que afectam pessoas portadoras de deficiência ou outros grupos minoritários, incluindo comunidades étnicas e refugiados. Os métodos adequados a esta área temática são geralmente diferentes a um nível macro e a um nível micro. Consideremos, por exemplo, as circunstâncias de emprego das mulheres. A um nível macro, as técnicas quantitativas utilizam-se normalmente para registar a evolução das mulheres no mercado de trabalho, por exemplo, em termos de participação (percentagem de mulheres activas no mercado de trabalho), níveis salariais, níveis de desemprego, etc.. Estes métodos poderão incluir inquéritos postais, a análise de dados administrativos realizada por centros de emprego, e a reanálise de dados disponíveis, através de inquéritos junto dos empregadores sobre as taxas salariais. É necessário, no entanto, abordar determinados aspectos da experiência no mercado de trabalho, de uma forma qualitativa. Por exemplo, para estudar as condições de trabalho, a entrevista a pessoas empregadas é um complemento necessário para obter métodos de estudo mais integrados. A um nível micro, a análise da experiência dos que são alvo de discriminação no mercado de trabalho ou as condições de vida e de trabalho (ou casos em que estas não sejam equilibradas) exige métodos qualitativos como as entrevistas, estudos de caso, *focus groups*, etc.. Esta será a única forma de descrever convenientemente as opiniões das pessoas que foram alvo de discriminação.
- Capacitação institucional e administrativa. Como se referiu na secção do GUIA sobre a capacitação para a avaliação, a capacidade institucional e administrativa constitui um aspecto fundamental, não só da capacidade de avaliação, mas também da forma como o sector público e o Estado respondem a novas exigências de responsabilização e de transparência. No âmbito do desenvolvimento socioeconómico, já se constatou como as parcerias constituem um elemento dominante da forma como esses programas são implementados. A avaliação da forma como as instituições e as administrações se adaptam, bem como da eficácia das condições de parceria, constituem conseqüentemente outra prioridade temática. Embora seja possível monitorizar os resultados das diferentes instituições e administrações em termos de indicadores de desempenho e eficiência, estes só fazem sentido quando contextualizados na análise das disposições organizacionais. Para o efeito, é necessário utilizar métodos de campo descritivos, tais como estudos de caso, técnicas de observação, entrevistas e análise documental.
- Desenvolvimento sustentável. Este constitui uma temática comum, transversal a praticamente todas as áreas de política, que implica diversos factores em interacção, e que exige conseqüentemente diversos métodos e técnicas em qualquer avaliação. Isto poderá implicar técnicas económicas, tais como análises custo-benefício, ou modelos de alocação de recursos e as respectivas conseqüências. Numa avaliação temática recente do desenvolvimento sustentável e os Fundos Estruturais, recorreu-se a um

grande número de estudos de caso para testar o modelo dos Quatro Capitais do desenvolvimento sustentável, e investigar as suas implicações como técnica de planeamento e análise. Deu-se particular destaque à opinião especializada dos avaliadores, que possuem uma sólida formação em aspectos de desenvolvimento sustentável, bem como capacidades de avaliação adequadas. Considerando o enfoque prospectivo do desenvolvimento sustentável em relação às consequências para as “gerações futuras”, são particularmente adequados métodos prospectivos como os inquéritos Delphi (uma técnica de consulta a peritos acerca das suas opiniões) ou os métodos de previsão. Anteriormente, identificou-se a categoria Prospectiva como uma classe importante ou “modo de inquérito” em investigação e estudos de avaliação. O desenvolvimento sustentável constitui um bom exemplo de uma situação em que este modo de inquérito é importante. Contudo, o âmbito do que se pode incluir no desenvolvimento sustentável levanta problemas familiares de agregação e de síntese. Como será possível encontrar um método de descrição ou um indicador compósito, que resuma tudo o que se concretizou em tantas intervenções diferentes? Este é um dilema típico noutras áreas temáticas de grande abrangência.

- Promoção da inclusão social. A redução das disparidades em toda a UE tem sido um tema consistente das políticas da UE. Este aspecto foi reforçado com a Estratégia de Lisboa, o que se encontra devidamente exemplificado através da Estratégia Europeia de Emprego, que destaca a inclusão social e uma maior coesão social. Utiliza-se, normalmente, grandes conjuntos de métodos, tais como a definição de metas e técnicas específicas, como indicadores para medir e orientar a coesão social. Estes utilizam-se nos Planos de Acção Nacional sobre Inclusão Social, e correspondentes relatórios que os Estados-membros têm vindo a preparar desde 2001. Para além disso, existe aqui espaço para métodos mais qualitativos e participativos. Como referido relativamente à igualdade de oportunidades, a opinião dos beneficiários-alvo das intervenções, como forma de evitar a exclusão social, constitui um elemento essencial do seu êxito. Entrevistas, inquéritos, *focus groups*, estudos de caso e técnicas de observação são importantes a este nível mais micro.
- Sociedade da informação. Uma vez mais, na sequência de Lisboa, e com o reforço da Cimeira de Estocolmo, o desenvolvimento de uma sociedade da informação e de uma economia do conhecimento é considerado uma questão essencial, quer em termos de competitividade europeia, quer em termos de coesão social. Este é conseqüentemente um tema transversal na Europa, em termos de diversos aspectos de desenvolvimento socioeconómico. Como é comum em áreas onde a concorrência europeia é uma questão, o *benchmarking* com outras partes do mundo, sobretudo com a América do Norte e com o Japão, é uma técnica habitual. Sempre que são planeadas intervenções em particular, como, por exemplo, as que estão incluídas nas iniciativas eEurope, recorre-se normalmente a métodos explicativos que procuram demonstrar o antes e o depois dos efeitos das intervenções. Porém, o âmbito das intervenções possíveis sob a bandeira da sociedade da informação é extraordinariamente vasto. Incluem-se intervenções ao nível das infra-estruturas regionais, educação, telemedicina e acesso de grupos desfavorecidos à tecnologia. Levantam-se conseqüentemente problemas de agregação e de síntese que não são atípicos. O desafio nesta área, como em outras áreas temáticas, consiste em procurar formas de acrescentar um grande conjunto de diferentes intervenções para informar uma apreciação coerente do que se alcançou. Isto tentar-se-ia fazer, muitas vezes, através de indicadores compósitos, embora sujeitos a críticas, por misturarem “alhos”

com “bugalhos”. Em alternativa, os avaliadores poderão segmentar este tema em áreas específicas de intervenção, medir os resultados alcançados nesses segmentos específicos, e, posteriormente, retirar conclusões qualitativas relativamente à evolução global.

Prioridades sectoriais e de política

Como referido na secção do GUIA sobre a avaliação e o desenvolvimento socioeconómico, determinadas abordagens da avaliação têm tendência para estar associadas a determinadas áreas de política. O mesmo acontece com a selecção de métodos e técnicas. Isto decorre da natureza da área de política correspondente. No sector agrícola na Europa, as avaliações dos impactos dos regimes de subsídio à produção tenderiam a ser uma abordagem mais comum do que a avaliação das intervenções de recursos humanos. Neste último caso, seria necessário incluir na discussão formas de medir a acumulação de capital humano e impactos no mercado de trabalho. Em certa medida, por conseguinte, determinados métodos e técnicas estão devidamente associados a determinadas áreas de política. Por exemplo:

- A avaliação das intervenções a nível dos transportes poderá incluir investimentos em infra-estruturas, em que seria adequado recorrer a análises do tipo custo-benefício e a outras técnicas, que possam modelar ou descrever os resultados das diferentes alocações de recursos. A utilização de sistemas de transportes poderá ser coberta, de forma mais adequada, através da análise de dados administrativos detidos pelas autoridades ou operadores de transportes. A satisfação dos passageiros, por outro lado, poderá ser captada, de forma mais adequada, por inquéritos por amostragem junto dos utilizadores de transportes;
- Avaliação das intervenções de gestão a nível do ambiente e da energia. Uma vez mais, poderá recorrer-se aqui a análises de custo-benefício, na fase de investimento. Esta é, também, uma área onde existem normalmente relações de compromisso entre as diferentes prioridades, como, por exemplo, as vantagens ambientais e as vantagens para o emprego. A descrição e medição desta situação de compromisso poderiam por norma ser abrangidas pelo que se designa de estudo de impacte ambiental. Uma vez que muitos dos aspectos da política ambiental possuem uma forte dimensão internacional (metas da ONU e de Kyoto), também é comum proceder-se ao *benchmarking* de acordo com estas normas internacionais.
- A avaliação das intervenções da justiça criminal e dos direitos humanos (conforme previsto em diversas intervenções da agenda política em relação à justiça, liberdade e segurança da Comissão, no período pós-Tampere) tem tendência para acompanhar as alterações a nível de resultados ao longo do tempo (ex. em termos de migração e asilo), analisando dados administrativos, e realizando estudos de cidadãos, tanto na UE, como em países terceiros (ex. as potenciais vítimas de tráfico humano).
- A avaliação das intervenções activas de políticas de mercado de trabalho e de educação e formação recorre a inquéritos aos beneficiários e estudos por painéis ou por coorte, que permitem acompanhar os impactos, a curto e longo prazo, dos programas. As avaliações nestas áreas de política também utilizam, em grande medida, métodos experimentais.

No Manual Técnico I, descreve-se uma gama mais completa de temas e prioridades de políticas, em termos de tipos de métodos de avaliação, dados e indicadores, que se aplicam geralmente. É necessário, no entanto, repetir o alerta lançado anteriormente. Os avaliadores que trabalham intensivamente em determinadas áreas políticas (tais como o ambiente, transportes, recursos humanos, ciência e tecnologia, e desenvolvimento de PMEs) têm tendência a ficarem presos a determinados métodos e técnicas. Muitas

vezes, isto poderá dever-se a motivos sólidos, como por exemplo, quando este processo decorre da natureza do conteúdo da política correspondente. Porém, também poderá dever-se ao facto de estes avaliadores trabalharem exclusivamente numa área política específica, com determinados colegas de avaliação e de políticas, e têm tendência para se isolarem da comunidade de avaliação mais alargada, e ignorar conjuntos de métodos que podem ser úteis, mas com os quais não estão familiarizados.

Desenvolvimento local e territorial

O desenvolvimento socioeconómico sempre teve uma forte dimensão territorial. Ao longo dos últimos anos, o desenvolvimento local (influenciado, como referido na secção do GUIA sobre a avaliação e o desenvolvimento socioeconómico, por teorias decorrentes da Nova Geografia Económica) tem vindo a ser perspectivado como sendo cada vez mais importante em termos de realização do potencial socioeconómico.

Não é habitual recorrer-se a métodos econométricos na avaliação do desenvolvimento socioeconómico, sobretudo porque as intervenções e os programas são geralmente responsáveis por uma parte relativamente pequena de recursos efectivos, e muitos outros recursos podem representar a sua maioria (transferência de pagamentos, outros programas, etc.).

Os modelos econométricos são adequados sempre que os programas abrangem todo um território, perante situações que satisfaçam os critérios que se seguem:

- Os fundos aplicados no âmbito do programa são significativos, comparativamente à economia do território em causa, e
- O território deve ser suficientemente grande (um país ou uma grande região), ou suficientemente fechado (uma ilha) para que o funcionamento da sua economia seja considerado isoladamente.

Sempre que estas condições estão reunidas, a equipa de avaliação pode utilizar diversas técnicas aplicadas no estudo de economias regionais ou em macroeconomia. Entre estas, incluem-se:

- A análise das componentes de variação (*análise shift-share*), que consiste em projectar as tendências da economia nacional na economia de uma região. Esta técnica utiliza-se para fazer a estimativa de uma situação em que a política não tenha sido implementada. Comparando-a com uma situação de aplicação da política, é possível avaliar o impacto global do programa.
- O modelo input-output e os modelos econométricos, que se aplicam para estimar o desenvolvimento económico de uma região. Em termos gerais, estas técnicas aplicam-se numa fase ex-ante a dois tipos de pressupostos (com/sem o programa), e permitem obter uma perspectiva global dos prováveis impactos macroeconómicos.

Sempre que uma avaliação se encontra relacionada com intervenções com um âmbito mais definido, é possível efectuar uma análise aprofundada das relações causais entre a intervenção e os seus efeitos. Neste contexto, podem utilizar-se diversas técnicas:

- Utiliza-se a análise de variância, a análise factorial e a análise de *clusters* para desvendar as semelhanças e as diferenças entre uma amostra de observações, para criar tipologias e para identificar factores exógenos relacionados com a produção dos impactos.
- O inquérito Delphi foi concebido para fazer uma estimativa dos impactos. É devidamente adequado às avaliações ex-ante que se baseiam em dados secundários. A técnica mobiliza e analisa os dados através da intervenção dos peritos. É, por isso, semelhante ao painel de peritos que também é adequado à análise de dados.

- Os grupos de comparação utilizam-se para fazer a estimativa dos resultados efectivos, assinalando a diferença entre um grupo de beneficiários e um grupo de não beneficiários.
- A análise de regressão utiliza-se para fazer a estimativa dos resultados efectivos, e para determinar se as relações causais entre a intervenção e os seus efeitos são estatisticamente significativas.
- Os estudos de caso, as entrevistas de grupo e a observação participante constituem técnicas de observação, mas também podem ser utilizadas de forma flexível para realizar a análise de conteúdo e comparar e analisar dados. É possível fazer uma estimativa dos efeitos (dentro de um intervalo, com mínimos e máximos de efeitos) através de estudos de caso executados entre uma selecção de intervenções a nível de projecto.

Uma vez que o desenvolvimento local começa com a análise do potencial, da capacidade e das necessidades locais, a sua avaliação é particularmente adequada a métodos participativos que suscitam aos parceiros e aos cidadãos locais as suas prioridades, atitudes e comportamentos. É nestas condições de desenvolvimento local que o envolvimento activo dos parceiros locais numa avaliação, incluindo a auto-avaliação participativa e avaliações orientadas para a capacitação, é mais útil. Estas abordagens estão cuidadosamente alinhadas com estratégias de desenvolvimento comunitário, elas próprias muitas vezes desenvolvidas em contextos de desenvolvimento local. Obviamente, a análise das bases de referência socioeconómicas actuais permite recorrer à gama tradicional de análises económicas e estatísticas. Além disso, as comparações transversais a diversas áreas muitas vezes privilegiadas (por exemplo, o benchmarking), requerem a aplicação de índices e medidas padrão, de modo a avaliar as realizações e os resultados.

O desenvolvimento local em particular é em muitos casos entendido como um processo que, ao longo do tempo, atravessa uma série de fases, em que cada uma das fases irá influenciar a fase seguinte. Por essa razão, os métodos de avaliação dos processos que acompanham o desenvolvimento ao longo do tempo são particularmente úteis. Entre estes, podem incluir-se, por exemplo: acompanhar os incidentes críticos ao longo do tempo; incentivar os residentes locais a manter registos diários; criar estudos de painel (inquéritos longitudinais junto dos mesmos inquiridos) que são revistos ao longo das diferentes fases do processo de desenvolvimento, e críticas à avaliação por parceiros locais (ver avaliação local, estudos de caso, abordagem participativa).

Uma das características do desenvolvimento local e territorial é a importância que atribuem à interacção de muitos factores diferentes, que contribuem para o processo de desenvolvimento, ou, pelo contrário, são responsáveis pelo subdesenvolvimento. Por isso, é necessário que a avaliação do desenvolvimento local e territorial considere métodos que identifiquem, descrevam e avaliem as interacções entre as diferentes intervenções e as suas contribuições relativas, bem como as suas sinergias (a sua matriz de impactos cruzados). Embora isto possa realizar-se recorrendo a indicadores, também é necessário desenvolver modelos que mostrem a relação entre os diferentes factores e as diferentes intervenções.

4.4. Opções para diferentes objectivos de avaliação

Outro conjunto de considerações necessárias para informar a selecção de métodos e técnicas é constituído pelos diferentes objectivos de avaliação identificados na secção do GUIA sobre a avaliação e o desenvolvimento socioeconómico. Os objectivos são:

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

- *Planeamento/eficiência* que garanta a existência de uma justificação para uma política/programa e que os recursos se desenvolvem eficientemente;
- *Responsabilização* - demonstrar em que medida é que determinado programa alcançou os seus objectivos, e de que forma utilizou os seus recursos;
- *Implementação* - aumentar a eficiência dos programas e a eficácia da sua implementação e gestão;
- *Produção de conhecimento* - aumentar o nosso entendimento do que efectivamente funciona e em que circunstâncias, e de que modo é que se pode tornar as diferentes medidas e intervenções mais eficazes.
- *Reforço institucional e de rede* - melhorar e desenvolver as capacidades entre os participantes no programa e as suas redes e instituições.

Em certa medida, a estes diferentes objectivos estão associados métodos e técnicas específicos. Por exemplo:

- No que diz respeito ao *planeamento e eficiência*, os métodos estão, sobretudo, relacionados com a atribuição de recursos e com a economia de custos. Serão adequadas diversas formas de análise de impacto, bem como diferentes formas de análise custo-benefício. Num sentido de gestão mais lato, também serão utilizadas técnicas orientadas para os objectivos, como as características de algumas abordagens de quadros lógicos. O Manual Técnico II descreve uma série destes métodos e técnicas. Entre eles, incluem-se, por exemplo, a análise input-output e a análise de eficiência.
- Relativamente à *responsabilização*, os métodos consistem, sobretudo, em avaliar o desempenho de acordo com determinados níveis ou metas, e aplicar critérios relevantes para o êxito e desempenho. Na sua forma mais directa, este aspecto aproxima-se do que se considera, em termos clássicos, o trabalho dos auditores. É possível fazer comparações com níveis de execução de diversas formas. Por exemplo, é possível utilizar indicadores para comparar os resultados reais com os esperados. Também é possível fazer comparações com exemplos externos, através do *benchmarking*. Embora não haja uma forma fácil de fazer as comparações a um nível externo, como acontece muitas vezes no âmbito específico do desenvolvimento socioeconómico, é possível efectuar comparações numa base de antes e depois, mostrando as alterações ao longo do tempo. Regra geral, as avaliações maioritariamente relacionadas com a responsabilização tendem a realçar medidas financeiras e monetárias e técnicas quantitativas. Porém, isto nem sempre acontece, uma vez que os decisores políticos geralmente consideram útil dispor de materiais ilustrativos dos casos e descrições qualitativas dos resultados do desenvolvimento, para apoiar descrições mais abstractas em termos financeiros ou, mesmo, monetários.
- No que diz respeito à *implementação*, alguns métodos comuns procurarão descrever os processos e os resultados internos, para proporcionar *feedback* aos responsáveis pela implementação dos programas. Muitos destes métodos e técnicas serão baseados em estudos organizacionais e de política. Poderá fazer-se comparações entre o desempenho das diferentes unidades administrativas, por exemplo, no sentido de saber se as diferentes regiões ou municípios estão a evoluir mais ou menos. Os estudos de caso das disposições organizacionais e de parceria irão ajudar a compreender as forças e as fraquezas das diferentes abordagens de implementação. Geralmente, estes tipos de métodos implicam o que se designa de métodos e técnicas de avaliação formativa. Estes colocam um destaque particular sobre o avaliador, no sentido de proporcionar *feedback* de formas que serão úteis, e que irão ajudar os gestores de programa a transformar os dados emergentes em acções práticas.

- No que diz respeito à *produção de conhecimento*, os métodos estarão mais próximos dos utilizados pelos investigadores académicos. Estes estarão sujeitos a exigências de maior rigor, representatividade e interpretação cuidada dos resultados, sobretudo nos casos em que estes possam ser inconsistentes. Normalmente, para efeitos de produção de conhecimento, os avaliadores têm que responder à questão: O que é que funciona? Numa perspectiva positivista, esta seria uma área onde os métodos experimentais são considerados importantes. No entanto, a natureza diversificada e de abordagem ascendente das intervenções socioeconómicas, a forma como estas se combinam em determinada configuração, e os diferentes locais e contextos onde decorrem os programas dificultam a aplicação das experiências tradicionais, excepto em circunstâncias muito pouco habituais. É por isso que o raciocínio realista, com destaque para a influência do contexto sobre os resultados, se tornou mais comum nestes tipos de avaliação. Coloca-se, assim, algumas questões mais complexas: o que é que funciona, para quem, como e em que circunstâncias? Os métodos e técnicas adequados para o efeito implicam normalmente uma comparação entre os diferentes casos seleccionados, para evidenciar intervenções alternativas e contextos alternativos. Estas comparações podem basear-se em estudos de caso, bases de dados que estruturam configurações de intervenções/resultados/contextos ou uma série de outras técnicas, com capacidade para captar e descrever estes diferentes aspectos de desenvolvimento socioeconómico.

Entre a comunidade de avaliação, é geralmente aceite que os conhecimentos fiáveis raramente decorrem de uma única avaliação. Existe, por isso, um interesse crescente em realizar estudos de síntese e diversos tipos de meta-análise, que procuram reunir as informações de um grande número de avaliações, à medida que estas vão estando disponíveis. Uma vez que a produção de conhecimento se tornou mais importante com o empenhamento dos decisores políticos na decisão política factual, alargou-se a utilização de diversos tipos de meta-análise. Esta forma de análise é reforçada sempre que, ao conceber avaliações que possam subsequentemente incluir-se em meta-análises, se recolha algumas estruturas comuns e dados, em todos os casos.

No que diz respeito ao *reforço institucional e de rede*, é reconhecido actualmente que a avaliação não se destina exclusivamente a satisfazer as necessidades dos gestores de programas e dos financiadores, mas que deve pertencer também ao vasto grupo de parceiros. Além disso, a implementação eficaz dos programas depende, muitas vezes, das capacidades das instituições e organizações de onde provêm estes parceiros, bem como das redes mais alargadas da sociedade civil. Muitas vezes, os métodos mais adequados nestes casos serão os participativos - com destaque particular para o trabalho de cooperação entre os avaliadores e as instituições e redes envolvidas. Estas abordagens participativas são importantes, não só para formular questões de avaliação, como também para gerar dados e utilizar os resultados da avaliação. Por exemplo, num contexto de comunidade onde existem diversos interesses e, talvez, uma falta de perspectivas partilhadas, os avaliadores poderão ter necessidade de trabalhar em conjunto com representantes da comunidade, no sentido de chegar a um consenso relativamente à utilização dos resultados de uma avaliação. Obviamente, as abordagens ao reforço institucional e de redes poderão ser adoptadas de forma muito mais directa. Por exemplo, é possível efectuar estudos da capacidade administrativa de organizações parceiras específicas, para as ajudar a adoptar processos de gestão e sistemas de informação mais adequados.

4.5. As opções para diferentes fases de programas/políticas

- Formulação: identificação de necessidades e prioridades
- Concepção: intervenções e organização
- Implementação: feedback e resultados intermédios
- Conclusões/resultados: efeitos e impactos

Introdução

A importância do ciclo temporal de implementação dos programas e das políticas tem sido um tema central deste GUIA. Nos Fundos de Coesão e nos Fundos Estruturais Europeus, este aspecto é formalizado em termos de avaliação ex-ante, intercalar e ex-post. Além destas etiquetas específicas, as fases subjacentes de um programa, desde a formulação de política e concepção de programas até à implementação, realização e conclusão ou obtenção de resultados, coloca algumas exigências à avaliação da maioria dos principais programas:

- Na fase de **formulação**, o enfoque reside na identificação de necessidades e na clarificação dos objectivos;
- Na fase de **concepção**, o enfoque reside na identificação de intervenções adequadas, e nos procedimentos de gestão da organização com capacidade de os realizar;
- Na fase de **implementação**, o enfoque reside nos processos de *feedback*, nos resultados intermédios e em proporcionar *feedback* capaz de incentivar a aprendizagem;
- Na fase de **conclusões ou resultados**, o enfoque incide nos efeitos ou nos impactos nos territórios ou beneficiários-alvo, em relação às intenções (por exemplo, decorrentes dos objectivos), bem como nas consequências não esperadas.

Formulação: identificação de necessidades e prioridades

Em termos gerais, o desenvolvimento socioeconómico decorre numa de duas perspectivas: uma perspectiva positiva, relativamente ao potencial de desenvolvimento, e uma perspectiva negativa, relacionada com as necessidades e os problemas a ultrapassar. Os métodos deverão descrever as circunstâncias de base, provavelmente em comparação com as circunstâncias de outros locais de desenvolvimento elegíveis. Em termos de medidas concretas, é necessário adoptar técnicas como o *benchmarking* e a análise de dados administrativos sobre níveis de rendimento, qualificações, participação no mercado de trabalho e características de mercado (tais como a proporção de actividades económicas dedicadas às exportações externas). As análises económicas do perfil do território ou região podem ser úteis, mas não só para identificar os fossos existentes em determinadas actividades. Também podem utilizar-se para demonstrar a existência de potencial de novas actividades, sobretudo nos casos em que a base de comparação reside noutra território ou região com algumas características idênticas ao local onde está a ser planeado o desenvolvimento. Modelos estatísticos e macroeconómicos mais sofisticados também permitirão estabelecer comparações com contextos idênticos (territórios ou sectores semelhantes, talvez noutros países), que podem utilizar-se para efeitos explicativos, numa fase posterior do desenvolvimento.

Considerando a importância dos parceiros no planeamento e oferta de programas de desenvolvimento socioeconómico, são úteis métodos que permitam o envolvimento dos parceiros numa fase inicial. Estes poderão incluir desde *focus groups* com métodos consultivos ou participativos, consultas locais, reuniões públicas, etc., até técnicas mais

formais, como uma análise SWOT, realizada junto de diferentes grupos de parceiros, para mostrar o que consideram ser importante alterar, e com que objectivos.

Apesar de muitos dos conteúdos do estabelecimento de prioridades decorrerem do sistema político, também é possível utilizar métodos que permitem aprofundar as considerações do que é necessário, incluindo, por exemplo, o mapeamento de questões ou o mapeamento de conceitos, que permitem proporcionar uma base de identificação, agrupamento e estabelecimento de prioridades das potenciais intervenções.

Concepção: intervenções e organização

A formulação de programas e de políticas inclui normalmente a identificação de condições de partida e das metas e objectivos pretendidos. No entanto, não é possível especificar as relações na cadeia que ligam as circunstâncias actuais ao futuro que se pretende. Este processo é o que decorre na fase de concepção. A criação de teorias de programas ou de modelos lógicos de programas socioeconómicos, que mostrem as cadeias de implementação relacionadas com as intervenções específicas, constitui uma forma útil de preencher as fases entre as circunstâncias de base e as metas a longo prazo. É possível complementar a utilização destes tipos de modelos com outras técnicas, como a avaliação da capacidade de avaliação, que vão muito além dos quadros lógicos, envolvendo activamente os gestores de programa e os decisores políticos na avaliação do que é possível proporcionar, de forma viável.

Uma vez mais, é necessário estar consciente do papel de muitos dos parceiros, incluindo cidadãos locais, em programas de desenvolvimento socioeconómico. Torna-se, assim, útil combinar a concepção de programas com métodos participativos, que também permitem começar a delinear as fases posteriores de uma avaliação. O envolvimento activo de grupos de parceiros na agregação das suas próprias teorias de programas, em vez de se basear num único exercício com decisores políticos, pode constituir uma abordagem. Algumas formas de criação de teorias de programa, como por exemplo a designada abordagem de teoria de mudança, são concebidas com o intuito de serem participativas, e para revelar as perspectivas e as teorias implícitas dos parceiros como actores, e não como meros destinatários dos recursos dos programas.

É necessário aplicar técnicas de avaliação a nível programático, e não em termos de intervenções individuais. Por exemplo, pode utilizar-se técnicas de apreciação dos projectos, incluindo análises de custo-benefício, para determinar as escolhas entre as diferentes intervenções, com vista a alcançar os mesmos objectivos. Também pode ser útil avaliar as relações de compromisso entre as diferentes medidas e intervenções.

Nesta fase, também é possível realizar estudos de síntese de mecanismos de implementação anteriores. Por exemplo, o que é que se sabe acerca dos procedimentos administrativos e organizacionais adequados? Quais os tipos de estrutura de tomada de decisão e de parcerias mais eficazes? Estes tipos de questões poderão certamente ter uma resposta mais adequada através de estudos de caso comparativos e de revisões da literatura relativamente a avaliações existentes.

Implementação: *feedback* e resultados intermédios

Ao longo da implementação de um programa, é necessário obter *feedback*, para que os gestores de programa possam identificar problemas e adoptar medidas correctivas. Os sistemas de monitorização proporcionam grande parte deste *feedback*. No entanto, os próprios sistemas de monitorização poderão ajudar a identificar as áreas problemáticas, que merecem uma investigação mais pormenorizada. Por exemplo, o início lento de determinados projectos e a subsequente subutilização do orçamento disponível, ou a

retirada de apoio de um importante grupo de parceiros, poderá justificar estes tipos de actividades de avaliação.

Sempre que a inovação planeada for particularmente inovadora ou experimental, poderá justificar-se o acompanhamento ou o seguimento do processo de implementação com algum pormenor. Estas actividades de avaliação formativas podem implicar técnicas como a observação participante, que poderiam ter que ser reforçadas através de *feedback* sistemático. Nesta fase do processo, o *feedback* pode ser muito bem recebido, mas também pode representar uma grande ameaça. São necessários diversos tipos de capacidades de comunicação e consultoria para gerir este tipo de *feedback* de forma construtiva. Este tipo de avaliação também pode exigir aos gestores de programa determinadas capacidades. Por exemplo, poderá ser necessário executar a sua própria auto-avaliação como parte de um processo global de avaliação.

Os sistemas de monitorização também permitem registar os resultados intermédios. Partindo do princípio que foram criados pormenorizadamente quadros lógicos e teorias de programas na fase de concepção, deverá existir um modelo que permita descrever as metas que se espera alcançar nas diferentes fases do programa. É possível utilizar-se indicadores como parte integrante deste processo de acompanhamento.

Conclusões: resultados e impactos

Os decisores políticos (por motivos de responsabilização) e os principais parceiros (devido às suas necessidades e compromissos) procuram na avaliação uma forma de fornecer informações sobre os resultados e os impactos, no final do ciclo de um programa. Os métodos de avaliação procuram comparar os resultados alcançados com os resultados esperados, e os aspectos finais com a base de partida. É possível implementar uma vasta gama de técnicas, incluindo:

- Inquéritos aos beneficiários-alvo;
- Modelos econométricos ou estatísticos, que demonstrem alterações do desempenho económico, comparativamente aos resultados previstos (talvez comparando tendências, num contexto de desenvolvimento, com outros contextos, e utilizando modelos desenvolvidos no início de um ciclo de desenvolvimento); e
- Indicadores baseados em dados contextuais ou administrativos, fornecidos por autoridades públicas.

Nos programas de desenvolvimento local, em particular, a dimensão participativa constitui uma parte importante dos métodos de avaliação. O objectivo não consiste apenas em assegurar a consideração da opinião dos beneficiários; consiste, também, em assegurar que os grupos locais, os cidadãos, e as associações, entre outros, possam tecer as suas próprias considerações relativamente ao êxito dos programas. Afinal, são as suas considerações, juntamente com as técnicas de auto-avaliação, que são importantes, devido à orientação ascendente destes programas – mais do que as considerações dos avaliadores ou decisores políticos, baseadas em critérios mais abstractos.

4.6. Recolha e utilização de dados em avaliação

- Todos os dados são “produzidos”
- Acesso aos dados como actividade planeada
- Dados quantitativos e qualitativos

Todos os dados são “produzidos”

Os avaliadores dependem do que se designa por dados: a matéria-prima que, depois de recolhida, é organizada, descrita, agrupada, contada e manipulada através de diversos métodos e técnicas. Faz-se normalmente a distinção entre dados primários (gerados como consequência directa de um programa ou intervenção) e dados secundários (gerados para outras finalidades, já existentes antes do programa ou intervenção). Por exemplo, entre as fontes de dados secundários poderão incluir-se:

- Fontes estatísticas, tais como estatísticas nacionais e regionais, EUROSTAT e outras bases de dados, com manutenção a cargo da DG REGIO;
- Relatórios anuais de autoridades de desenvolvimento ou federações de empresas; e
- Registos administrativos de agências de emprego públicas, reembolso de impostos, dados de qualificações e formação.

A existência de todos estes dados exige um esforço considerável, devendo os avaliadores saber de que forma se reuniu os dados secundários, antes de os poderem utilizar. Quais foram as amostras utilizadas? Como se definiram os resultados? Qual o prazo abrangido? Qual é a unidade de análise? Será apenas colocando estas questões que se pode fazer uma apreciação da viabilidade da sua utilização numa avaliação em particular. Normalmente, por exemplo, a unidade geográfica relativamente à qual se reúne dados administrativos ou estatísticos não se adequa aos limites do programa de desenvolvimento socioeconómico em questão.

É mais fácil para um avaliador compreender a origem dos dados primários. Entre estes, poderão incluir-se:

- Monitorização dos dados produzidos por um programa, como parte das suas obrigações de relatório perante as autoridades de financiamento;
- Dados de utilização gerados pela utilização ou retenção de serviços, fundos ou instalações proporcionados por um programa; e
- Dados recolhidos de centros de desenvolvimento e beneficiários-alvo pelos avaliadores, através de inquéritos junto de beneficiários, contabilização das pessoas que utilizam fundos de consultoria, *focus groups* e consulta aos parceiros.

No entanto, também neste caso os dados não emergem totalmente processados. A sua recolha também implicou a aplicação de protocolos e técnicas que determinam os fins para os quais podem ser utilizados. Os dados de utilização variam de acordo com os diferentes tipos de utilizadores? As informações de monitorização estão limitadas aos dados financeiros? Em que medida é que as amostras de beneficiários são representativas?

Uma vez que todos os dados são processados e são resultantes de decisões adoptadas no decurso da recolha, os avaliadores necessitam de compreender em que consistiam estas decisões, sobretudo quando adoptadas por terceiros – tal como acontece com os dados secundários. Este passo nem sempre é fácil, mas é necessário. No que diz respeito aos dados primários, que são gerados por ou próximo de um programa e da sua avaliação, a equipa de avaliação encontra-se na melhor posição para saber quais foram as decisões adoptadas. Mesmo neste caso, deve fazer-se uma distinção entre as formas de dados produzidas directamente por uma avaliação e as formas de dados geradas pelo programa (por exemplo através de monitorização), sobre as quais a equipa de avaliação possui um controlo mais reduzido. No entanto, mesmo quando os dados são recolhidos directamente pela equipa de avaliação, as suas forças, os seus limites, o seu âmbito e a sua relevância têm que ser pensados em termos dos tipos de análises futuras que serão adoptadas, e dos tipos de argumentos que os dados terão que apoiar.

Este é mais um argumento para que a avaliação seja pensada cuidadosamente na fase de concepção. É necessário projectar a recolha de dados em consonância com a selecção de métodos de análise.

Acesso aos dados como actividade planeada

A existência de múltiplas intervenções, envolvendo diversos parceiros, significa que os dados correspondentes a um programa de desenvolvimento socioeconómico serão mantidos em diversos locais diferentes. Fazer simplesmente um mapeamento do local de manutenção dos dados e do que se encontra disponível constitui um sério risco. A negociação e acordo dos termos ao abrigo dos quais os dados serão fornecidos ou colocados à disposição poderão ser mais complexos. Os dados administrativos especificamente poderão estar sujeitos a diversas regras de confidencialidade ou protecção de dados. Por vezes, os dados administrativos, por exemplo, só poderão ser divulgados após a eliminação das etiquetas de identificação (nomes e códigos postais). Mesmo nos casos em que tal não constitua um problema, os organismos administrativos são em muitos casos zelosos das suas fontes de informação. A negociação do acesso aos dados constitui uma tarefa que requer algum tempo.

Por vezes, o processo de aceder às fontes de dados poderá, ele próprio, gerar dados úteis para uma avaliação. Por exemplo, a disponibilidade dos parceiros para partilhar informação poderá ser vista como um indicador da coerência e da força de uma parceria. A recusa constante de partilhar informações, por exemplo, sugere que a parceria não se caracteriza por elevados níveis de confiança.

Sempre que os avaliadores estão directamente envolvidos na geração de dados, como acontece com diversas fontes de dados primários, ainda existem problemas de acesso, que requerem uma atenção cuidada. Alguns exemplos de questões de acesso incluem:

Garantia de elevadas taxas de resposta em inquéritos por amostragem. As reduzidas taxas de resposta são extremamente frequentes na avaliação, e isto poderá enfraquecer a base de factos disponíveis para retirar conclusões e tecer considerações. Existem diversas formas possíveis de aumentar as taxas de resposta, como, por exemplo:

- Informar claramente (possivelmente através da imprensa local ou especializada) qual o objectivo da avaliação, em geral, e dos inquéritos, em particular.
- Conceber os instrumentos do inquérito em linguagem simples, e não em termos técnicos.
- Dedicar algum tempo às actividades de acompanhamento, tais como cartas, telefonemas e re-emissão dos instrumentos do inquérito, para lembrar os não respondentes, depois de decorrido um determinado período de tempo.

Obter acesso a grupos desfavorecidos. Todos os que se designam, muitas vezes, de grupos de difícil acesso são por norma essenciais numa avaliação. Estes grupos podem desconfiar da acção oficial, e esta situação pode transpor-se para uma avaliação. Entre as formas existentes para ultrapassar estes problemas, encontram-se:

- Estabelecer a ligação com “guardiões” da comunidade, no sentido de poderem agir como advogados locais de uma avaliação.
- Produzir instrumentos em diferentes línguas (sempre que se utilizar línguas minoritárias) ou em diversos formatos, como Braille ou áudio, para as pessoas portadoras de deficiência.
- Empregar elementos locais, pertencentes a estes grupos, para recolher informações, gerir *focus groups* e explicar a avaliação nas suas próprias redes.

Beneficiários e parceiros que pretendem retirar benefícios de uma avaliação. Uma das principais razões da falta de cooperação (ou de uma cooperação menos entusiástica) é a percepção de que os que são solicitados para cooperar não irão retirar quaisquer benefícios deste exercício. Este facto pode ser ultrapassado, ou, pelo menos, minimizado, se:

- Existir um envolvimento adequado dos beneficiários e dos parceiros nas fases de concepção da avaliação global, e na concepção e implementação piloto de instrumentos em particular. Esta medida irá garantir que, por exemplo, os gestores de PME vejam as suas questões de avaliação como estando incluídas na agenda da avaliação e, por conseguinte, percepcionem os resultados como sendo importantes para eles mesmos.
- São dadas garantias de que, todos os que colaboram, irão obter *feedback*. Poderá assumir a forma de um relatório disponível publicamente, de uma carta de *feedback* com um resumo ou um convite para uma reunião de discussão de *feedback*, após a conclusão da avaliação.

A qualidade dos dados e a disponibilidade dos “guardiões”, dos parceiros e dos beneficiários para cooperar será um factor determinante para a qualidade dos dados, e, em última instância, da avaliação como um todo. Vale a pena prestar atenção e proceder ao planeamento como parte integrante da recolha dos dados e selecção dos métodos.

Dados quantitativos e qualitativos

Já se destacou a importância de nos basearmos num conjunto completo de métodos e técnicas de avaliação, incluindo os quantitativos e qualitativos, em simultâneo. Agora, descrevem-se resumidamente algumas das características dos dados, em termos dos seus factores diferenciadores qualitativos e quantitativos.

Em primeiro lugar, é essencial reiterar a importância de se estabelecer uma distinção entre dados recolhidos e dados analisados. Como já foi referido, é necessário que praticamente todos os dados sejam produzidos e processados para poderem ser utilizados, e, em muitos casos, o que começa por ser informação qualitativa transforma-se em dados quantitativos, através de diversos métodos analíticos. Porém, mesmo após a aplicação de diversos métodos de análise, as características e a ponderação dos dados quantitativos são diferentes.

O que se designa por dados quantitativos pode assumir diferentes formas, como, por exemplo:

- Pode não ser mais do que uma forma de descrever categorias. Os dados categoriais ou nominais não têm valor numérico; antes, os números são utilizados para fazer a distinção entre as diferentes categorias. Assim, Grupo 1, 2, 3 e 4 poderão ser rótulos aplicados a quatro conjuntos diferentes de PME, distinguidos pelos sectores em que operam.
- Ligeiramente mais fortes em termos de quantificação poderão ser os dados ordinais, relativamente aos quais se sabe que alguns itens são mais ou menos, maiores ou mais pequenos, do que outros. Por exemplo, algumas empresas poderão estar em crescimento, enquanto outras poderão estar em declínio, e este factor pode ser utilizado como fonte de dados, mesmo que não seja possível calibrar ou ter acesso às diferenças precisas entre o crescimento e o declínio de cada empresa.
- Um modo de quantificação ainda mais forte poderá decorrer quando se coloca diferenças relativas numa escala, onde se poderá conhecer os intervalos na escala. Diversos itens e sistemas de classificação e pontuação nos questionários correspondem a este tipo de quantificação. Por exemplo,

determinado perito poderá classificar o impacto ambiental de um projecto algures entre -3 e +3, ou, então, poderá pedir-se a um inquirido para se posicionar numa escala de satisfação entre 1 e 5. Embora estas formas de quantificação sejam mais fortes do que as descritas anteriormente, são relativamente fracas em termos das suas possibilidades numéricas e de cálculo.

- Os dados de relação (*ratio data*) constituem a forma de quantificação mais forte. Esta situação decorre sempre que for conhecido um ponto zero numa escala. Assim, não se trata de uma série de intervalos inventada, mas sim de algo que se pode medir independentemente. Por exemplo, os valores monetários, os perfis etários das populações, os fluxos de exportação e os índices de produtividade baseados em resultados anuais constituem normalmente uma forma de dados de relação. Poderia argumentar-se que é isso a que normalmente se refere quando se fala de dados quantitativos, embora seja menos comum em avaliação do que por vezes se possa imaginar.

O motivo desta descrição relativamente técnica das diferentes características dos dados quantitativos reside na necessidade de realçar que, no desenvolvimento socioeconómico, a maior parte do que se designa por dados quantitativos é, de facto, relativamente frágil. A redução da exclusão social, o aumento da qualidade dos recursos humanos e a diversificação de uma economia rural depende geralmente de dados categoriais, ordinais e de intervalo. Estas medidas não podem ser consideradas objectivas ou tão sólidas como por vezes se assume que sejam os dados quantitativos. Determinados tipos de dados (como, por exemplo, dados relativos à competitividade de empresas locais ou à participação no mercado de trabalho) podem estar sujeitos a uma análise mais rigorosa, utilizando o que se descreveu acima como dados de relação. Porém, estes níveis de rigor serão menos frequentes na maioria das avaliações nesta área.

Como parecem mostrar as distinções anteriores, é possível compreender melhor os dados quantitativos/qualitativos como um *continuum* dos mais quantitativos para os mais qualitativos. Os dados categoriais e ordinais podem ser vistos, em diversos aspectos, como sendo relativamente qualitativos. O que se poderia considerar dados puramente qualitativos é altamente discutível, talvez constituído por relatórios ou instâncias únicas, podendo descrever-se apenas numa base casuística. Um estudo de caso ou uma história de vida reunidos sem qualquer comparação ou categorização prévia poderão corresponder a este ideal qualitativo. Porém, assim que forem recolhidos diversos exemplos destes casos ou biografias, torna-se possível proceder ao mesmo processo de categorização, conforme descrito anteriormente, no título sobre dados quantitativos.

Em termos qualitativos, os dados podem ser comparados e, eventualmente, analisados de acordo com uma série de dimensões. Sempre que estiver em estudo uma população de fenómenos relacionados, a categorização ou classificação numa escala é o mais adequado. Por exemplo, uma série de pessoas, uma amostra de empresas, ou uma amostra de agregados familiares. Porém, sempre que estiverem em estudo exemplos únicos ou individuais (como sejam uma comunidade rural ou um sector de produção em particular), poderão ocorrer mais comparações ao longo do tempo (antes e após as medidas), ou em relação a níveis de desempenho ou critérios externos.

O objectivo último deste tipo de perspectivas consiste em baralhar a distinção entre dados quantitativos/qualitativos. A distinção é mais forte em termos de intenção analítica. As opiniões e as perspectivas das pessoas podem ser únicas e qualitativas ou tipificadas, no âmbito de uma classificação mais vasta; mas os dados brutos permanecem os mesmos. É uma questão de apurar a forma de processamento dos dados em bruto, para efeitos de análise. Os dados quantitativos serão mais utilizados quando for necessário

proceder a uma agregação e generalização; os dados qualitativos serão utilizados sempre que for necessário descrever a complexidade e os pormenores de uma experiência. A escolha entre estas estratégias deve, em última instância, depender das questões a que determinada avaliação pretende responder. Os inquéritos de satisfação junto dos beneficiários dos programas e as avaliações participativas junto dos grupos excluídos utilizarão os dados de opinião de forma bastante diferente.

4.7. Criação de indicadores e sistemas de indicadores

- O que são indicadores?
- Os indicadores e a avaliação
- Selecção de indicadores
- Avaliação da qualidade
- Aplicação dos indicadores

O que são indicadores?

Definição e características de um indicador

Poderá definir-se um indicador como a medição de um objectivo que se pretende cumprir, um recurso que se pretende mobilizar, um efeito que se pretende obter, um medidor de qualidade ou uma variável de um contexto. Um indicador produz informações quantificadas, com vista a ajudar os actores relacionados com as intervenções públicas a comunicar, negociar ou tomar decisões. No âmbito da avaliação, os indicadores mais importantes estão relacionados com os critérios de sucesso das intervenções públicas.

Para ser útil, um indicador deve possuir preferencialmente as seguintes características:

- A definição de indicador está estreitamente relacionada com metas, objectivos e/ou fins políticos. (De facto, os indicadores são de maior utilidade quando os objectivos forem especificados em termos de fins ou marcos aplicáveis, de acordo com a definição do indicador).
- O indicador é medido regularmente. É útil possuir informações de situações onde as definições precisas de indicador foram aplicadas consistentemente. Numa situação ideal, os dados devem estar disponíveis desde antes da adopção ou implementação da intervenção. Porém, as próprias intervenções exigem, muitas vezes, a recolha de novos dados.
- É medido numa base independente. É preferível que as informações sejam recolhidas por agências que não sejam directamente responsáveis pela intervenção ou legislação.
- A medida baseia-se em dados fiáveis.

Na prática, os indicadores raramente apresentam todas estas características, e pode ser necessário reunir provas de uma série de fontes distintas, incluindo:

- Os recursos e a calendarização do processo de programação;
- As fontes secundárias;
- As fontes primárias, incluindo inquéritos aos parceiros;
- Informações administrativas.

Muitas destas informações poderão ter sido recolhidas para outros efeitos que não o da avaliação.

Um indicador quantifica um elemento considerado relevante para a monitorização ou avaliação de um programa. Por exemplo: "1.200 desempregados de longa duração

fizeram formação financiada pelo programa" ou "75% dos participantes nos cursos de formação estão satisfeitos ou altamente satisfeitos".

Um bom indicador deve proporcionar informações simples que, tanto o prestador do serviço, como o utilizador, possam transmitir e compreender. Esta é, no entanto, uma qualidade necessária, mas insuficiente. Os indicadores que se seguem são exemplos de indicadores facilmente compreensíveis: taxa de absorção orçamental; percentagem de empresas regionais assistidas; número de empregos efectivos criados; e número de desempregados na área elegível.

Um indicador poderá possuir diversos valores ao longo do tempo. A taxa de desemprego, por exemplo, poderá ter um valor, à partida, diferente do valor intercalar através da implementação de um programa, e assim sucessivamente. As variações ao longo do tempo constituem tendências.

Tipos de indicadores

Existem diversas tipologias de indicadores:

- Em relação às variáveis: indicadores completos, parciais e complexos;
- Em relação ao processamento da informação: indicadores elementares, derivados e compostos;
- Em relação à comparabilidade da informação: indicadores específicos, genéricos e essenciais;
- Em relação ao âmbito da informação: indicadores de contexto e de programa;
- Em relação às fases de conclusão do programa: indicadores de recurso, realização, resultados e impactos;
- Em relação aos critérios de avaliação: indicadores de relevância, eficiência, eficácia e desempenho;
- Em relação ao modo de quantificação e utilização da informação: indicadores de monitorização e de avaliação.

As tipologias mais úteis para os programas socioeconómicos são constituídas pela distinção entre: indicadores de recursos, efeitos, resultados e impactos. Os indicadores contextuais, que podem coincidir, em alguns casos, com os indicadores de impacto, constituem mais uma categoria útil.

Os **indicadores de recurso** transmitem informações sobre os meios financeiros, humanos, materiais, organizacionais ou regulamentares utilizados para implementar os programas. Os recursos constituem uma responsabilidade conjunta das autoridades de financiamento, que os distribuem, bem como os operadores que os implementam. Os indicadores de recurso são, na sua maioria, regularmente quantificados por sistemas de monitorização. Entre os exemplos de indicadores de recurso incluem-se: o orçamento total (quantidade de recursos); a absorção orçamental anual (taxa de absorção de recursos); a percentagem de subutilização/sobreutilização esperada; a percentagem de financiamento europeu, face ao financiamento público total; o número de pessoas a trabalhar na implementação do programa; o número de organizações envolvidas na implementação.

Os **indicadores de realização** representam o produto das actividades dos programas. Mais precisamente, considera-se uma realização o que se obtém em troca da despesa pública. As realizações encontram-se normalmente sob a responsabilidade dos operadores que as relatam através do sistema de monitorização. Entre os exemplos de indicadores de realização, incluem-se: quilómetros de estradas construídos; taxa de evolução de construção de uma estrada; hectares de terrenos urbanos devolutos reabilitados; capacidade de centros de tratamento de resíduos construídos; número de

formandos cuja formação foi financiada pelo programa; e percentagem desta formação cuja qualidade se encontra certificada.

Os **indicadores de resultado** representam as vantagens imediatas do programa (ou, excepcionalmente, as desvantagens imediatas) para os beneficiários directos. Considera-se uma vantagem imediata se esta surgir enquanto o beneficiário se encontra em contacto directo com o programa. É possível observar os resultados completos quando o operador concluir a acção, e encerrar os pagamentos. Em termos gerais, os indicadores de resultado são facilmente reconhecidos pelos operadores, pelo que são geralmente quantificados durante a monitorização.

Os indicadores de resultado proporcionam informações sobre as mudanças que ocorrem, e que afectam os beneficiários directos, como, por exemplo, o tempo economizado pelos utilizadores de uma estrada, o custo reduzido das chamadas telefónicas, as qualificações obtidas pelos formandos, as novas actividades turísticas geradas por um agricultor, a utilização de novas capacidades de produção utilizadas por uma empresa, e a satisfação das empresas que receberam a prestação de serviços de consultoria.

É no momento em que os beneficiários recebem apoio ou serviços do programa que se pode quantificar os resultados. Executa-se medições directas (por exemplo, contabilizando o número de formandos recrutados no decorrer da formação), ou, então, pede-se aos beneficiários directos para especificar as vantagens obtidas (por exemplo, através de um inquérito, distribuído no final da consultoria).

Os **indicadores de impacto** representam as consequências do programa, para além da sua interacção directa e imediata com os beneficiários. Uma categoria inicial de impactos agrupa as consequências do programa para os beneficiários directos, que surgem ou duram até médio prazo (impactos específicos), como, por exemplo, o tráfego numa estrada um ano após a sua abertura, a taxa de formandos colocados depois de decorridos doze meses, os empregos sustentáveis criados numa unidade industrial construída com o apoio do programa, e a taxa de sobrevivência das empresas criadas com o apoio do programa. Alguns impactos não são previsíveis (*spin-offs*), mas raramente são criados indicadores de impactos imprevisíveis.

Uma segunda categoria de impactos é constituída por todas as consequências que afectam, a curto ou médio prazo, as pessoas ou organizações que não são beneficiários directos. Estes impactos podem ser semelhantes (por exemplo, o aumento da qualidade de vida das pessoas que vivem próximo de uma zona industrial devoluta reabilitada, o aumento da qualidade das praias junto a um centro de tratamento de resíduos) ou poderão alastrar a pessoas ou organizações que não se encontram ao abrigo do programa, como é o caso dos impactos macroeconómicos.

Pode dividir-se os mecanismos de propagação de impactos em duas categorias: efeitos de mercado (por exemplo, impacto sobre os prestadores ou subcontratantes das empresas assistidas) e efeitos não mercantis (por exemplo, o impacto positivo da imagem melhorada da região ou o impacto negativo da deterioração do ambiente). Uma vez que os efeitos não mercantis – ou externalidades – não se reflectem no sistema de preços sobre o qual os actores socioeconómicos individualmente baseiam as suas próprias decisões, e que estas decisões representam consequências económicas para outros actores, será particularmente útil considerar estes efeitos, no contexto de um programa público.

Devido ao fosso temporal ou à sua natureza indirecta, os operadores não poderão reconhecer facilmente os impactos durante a sua gestão quotidiana do programa. Os indicadores de impacto são, por isso, quantificados apenas de vez em quando, normalmente no decorrer de avaliações. Uma forma de determinar os impactos é realizar um inquérito junto de beneficiários directos, por exemplo um ano após abandonarem o programa. As questões colocadas poderão ser relativas a factos (por exemplo, quantos

empregos novos foram criados desde que foi recebido o apoio?) ou a opiniões (por exemplo, quantos empregos se poderia ter perdido se não tivesse sido atribuído o apoio?)

Os indicadores e a avaliação

Os indicadores desempenham uma série de papéis úteis na avaliação. A sua utilização é comum, no que diz respeito à avaliação do programa, sobretudo sempre que os objectivos são expressos em termos operacionais claros. A utilização de indicadores faz normalmente parte de uma avaliação. As informações proporcionadas devem ser cuidadosamente interpretadas à luz de outros factos, de modo a poder retirar-se conclusões de avaliação. Os indicadores possuem potencial para contribuir para a avaliação dos programas socioeconómicos de diversas formas:

- A análise da pontuação dos indicadores pode utilizar-se para sustentar uma filosofia de intervenção e uma atribuição de recursos.
- Os indicadores podem ser utilizados para comparar recursos e resultados, de modo a avaliar a eficiência.
- Os indicadores podem ser utilizados para comparar resultados reais com resultados esperados, de modo a avaliar a eficácia.
- Os indicadores podem ser utilizados para comparar recursos relativos a impactos e, por conseguinte, permitir a avaliação do valor (valor acrescentado) de políticas, legislação ou iniciativas.

O sistema de indicadores e o ciclo de programas

Os indicadores utilizam-se no início de um ciclo de programas para ajudar a definir territórios elegíveis para apoio, a analisar o contexto regional, a diagnosticar problemas económicos e sociais que é necessário solucionar, e a avaliar as necessidades que o programa terá que satisfazer. Nesta fase, os indicadores como a taxa de desemprego ou as disparidades entre infra-estruturas desempenham, muitas vezes, um papel decisivo.

A escolha e validação da estratégia de intervenção integram a segunda fase do ciclo de programação. Nesta fase, os responsáveis pela concepção do programa devem definir os objectivos de forma precisa, e quantificá-los. Os indicadores dependem da quantificação, e também são muito úteis para esclarecer os objectivos. A Caixa *A definição de indicadores permite esclarecer os objectivos* apresenta um exemplo.

Caixa: A definição de indicadores permite esclarecer os objectivos

Entre as estratégias utilizadas para sustentar a competitividade das PME's, inclui-se em muitos casos a prestação de serviços de consultoria. Os objectivos destas medidas poderão ser expressos em termos do número de empresas que usufruem de determinados tipos de serviços de consultoria. O indicador permite, não só quantificar o objectivo, mas também definir o serviço esperado. Por exemplo, a definição de recepção de um serviço por uma PME poderá incluir apenas trabalhos de consultoria significativos, num total superior a 5 dias de consultoria.

Depois de definido e adoptado, o programa é implementado. A sua monitorização e avaliação são efectuadas numa base contínua, e em fases intercalares. Nesta fase, os indicadores são indispensáveis para divulgar, de uma forma simples e condensada, informações exigidas pelos gestores de programa. Normalmente, os indicadores permitem monitorizar o ritmo de aplicação dos orçamentos, o grau de cumprimento da calendarização prevista, a proporção de população elegível alcançada, a taxa de satisfação dos beneficiários, e o número de empregos criados.

O ciclo de programação termina com uma avaliação ex-post, que possui como uma das suas principais funções reportar os resultados do programa, em termos de quais as metas alcançadas. Recomenda-se vivamente a utilização de indicadores nesta fase, na medida em que permite a transmissão de informações simples, apreendidas imediatamente pelo público em geral, como, por exemplo, o custo por cada emprego criado ou a taxa de colocação de desempregados apoiados.

Indicadores para programas integrados

A maioria dos programas socioeconómicos adopta estratégias integradas, ou seja, procura seleccionar todos os problemas que afectam um determinado território, e utiliza todos os instrumentos disponíveis para intervir nesse território. Esta característica implica necessariamente uma multiplicação das necessidades em termos de indicadores, que poderia causar confusão, se os programas não fossem devidamente estruturados. Os programas financiados pelos Fundos de Coesão e pelos Fundos Estruturais Europeus encontram-se normalmente estruturados de acordo com três níveis:

- o nível de programa global, com o qual está relacionado o objectivo geral, como, por exemplo, o emprego ou o desenvolvimento económico. Este nível é constituído por um pequeno número de eixos prioritários (normalmente, menos de seis), que dividem o objectivo global de acordo com as suas dimensões estratégicas;
- o nível de medida (de um até várias dezenas) existiu até 2006 e correspondeu à unidade básica de gestão do programa. Cada medida possuía um aparelho de gestão específico. Embora as medidas já não façam parte do programa de Fundos de Coesão e Estruturais *per se*, as autoridades de gestão continuam a operar "medidas" para efeitos de gerir os seus programas;
- o nível de projecto (muitas vezes, algumas centenas), que é constituído pela unidade de implementação do programa, uma vez que cada um dos projectos constitui um ponto de articulação entre o programa e os seus beneficiários.

Aspectos organizacionais: envolver utilizadores e fornecedores de informação

Um sistema de indicadores possui uma maior probabilidade de funcionar quando os fornecedores e utilizadores das informações foram envolvidos na sua criação. Em contrapartida, um grupo fechado de especialistas será tentado a construir um sistema dispendioso, tecnicamente ideal, que poderá nunca funcionar de forma satisfatória.

No que diz respeito aos utilizadores, deverá assegurar-se o apoio explícito do mais elevado nível de autoridade de gestão do programa. É, por isso, aconselhável criar um grupo de futuros utilizadores do sistema, e atribuir-lhe a função de definir os indicadores.

Deverá, então, nomear-se uma equipa para apoiar o grupo e desempenhar as funções de secretariado. Por norma, os membros da equipa pertencem à autoridade de gestão do programa. Estes deverão possuir os recursos humanos e financeiros necessários. A equipa deve, sobretudo, garantir que o sistema de indicadores reflecte claramente os objectivos do programa, e beneficia a comparabilidade. É preferível que a mesma equipa que é responsável pela criação de indicadores seja posteriormente responsável pela implementação do sistema.

O público também pode estar envolvido na concepção do sistema de indicadores. Um exemplo de envolvimento dos beneficiários na escolha de indicadores da experiência americana (região de Benton Harbor, ver a Caixa *Envolvimento dos beneficiários na escolha de indicadores*), começou com uma série de *focus groups*, envolvendo

representantes de empresas regionais. O trabalho destes grupos permitiu seleccionar indicadores com maior capacidade de atrair a atenção do público, e serem assim mais facilmente compreendidos.

Caixa: Envolvimento dos beneficiários na escolha dos indicadores

As autoridades políticas da região de Benton Harbor (USA) criaram um sistema de indicadores de contexto, com o objectivo de avaliar o desenvolvimento económico da área metropolitana de Berrien. Os principais elementos da descrição do desenvolvimento económico basearam-se numa série de entrevistas de *focus groups*, que envolveram os principais empreendedores da região. Os indicadores escolhidos foram, por exemplo: - Em termos de disponibilidade de capital humano qualificado: despesas de educação por criança; rácio professor/estudante; taxa de abandono escolar. - Em termos de crescimento e diversificação da economia: rendimento *per capita*; taxa de emprego no sector industrial; valor acrescentado no sector do comércio; índice de diversidade sectorial; taxa de emprego em PME; valor e número de residências relativamente aos quais foram emitidas autorizações de construção. Em termos de qualidade de vida: o custo de vida (relativo), a taxa de crimes contra a propriedade e de crimes contra a integridade das pessoas.

Erickcek G.A. (1996) The Benton Harbor Area Benchmarking Data System, Michigan: WE Upjohn Institute

As principais fontes de informações são os operadores que implementam o programa em campo. A sua participação poderá assegurar que o sistema é pragmático, uma vez que estão familiarizados com as possibilidades práticas e com os limites da recolha de dados.

Também é aconselhável envolver os operadores num teste preliminar do sistema de indicadores. O procedimento recomendado começa com a selecção de alguns operadores voluntários, que participarão na concepção do sistema. Estes voluntários devem representar todos os componentes do programa. Eles ajudam a seleccionar os indicadores, no sentido de os definir e de planear o processo de recolha de informações. Expressam as suas necessidades em termos de *feedback* de informações (frequência e forma como as informações lhes são transmitidas). O teste inclui uma quantificação inicial de todos os indicadores, por operadores voluntários. A duração normal destes testes é de um ano. Dependendo das conclusões do teste, e depois de efectuar as modificações necessárias, o sistema é validado. As definições e os procedimentos de recolha e restituição de dados são claramente determinados, e é preparado o manual.

As informações relacionadas com o contexto são retiradas dos dados estatísticos. É, por isso, aconselhável envolver um perito, com conhecimentos exaustivos e recentes de dados estatísticos capazes de exploração, na concepção do sistema. Conforme o caso, este perito pertence a um instituto de estatística ou a uma universidade, ou, se possível, à instituição responsável pela execução do programa.

Seleção dos indicadores

Seleção dos indicadores mais relevantes

Cada um dos actores do programa tem as suas próprias responsabilidades, as suas próprias áreas de tomada de decisão, e, por conseguinte, as suas próprias necessidades de informação. Consequentemente, nem todos os indicadores são úteis a todos os níveis. Pelo contrário, considera-se normalmente que cada um dos actores necessita de um relatório operacional, com um pequeno número de indicadores, escolhidos por se considerarem os mais adequados à natureza das decisões a tomar. Provou-se que, numa situação de tomada de decisão, não se pode ter em consideração mais do que cerca de

dez indicadores de cada vez (estas tipologias são consistentes com as especificadas no documento *Common Guidelines for Monitoring and Evaluation*, Comissão Europeia, 1995, Luxemburgo: OPOCE). Quando existem demasiados indicadores, os decisores ficam “inundados” com demasiada informação.

A heterogeneidade dos programas

A experiência obtida com os Fundos Estruturais e com os Fundos de Coesão tem vindo a revelar que é difícil seleccionar os indicadores que são absolutamente necessários para a monitorização e avaliação de um programa. Uma vez que os programas são multi-sectoriais e possuem múltiplos objectivos, existe uma tendência para tentar medir tudo e conceber sistemas de indicadores tão pesados, que se torna impossível pô-los em funcionamento.

Na prática, é impossível produzir e utilizar regularmente um volume tão grande de informações. Por exemplo, a proposta inicial de um Programa de Fundos Estruturais no Burgundy (França) era constituída por mais de 200 indicadores. No final, destes quantificaram-se apenas cerca de 50.

Em diversas regiões europeias, as avaliações mostraram que algumas dezenas de indicadores são suficientes para satisfazer as necessidades de informação das autoridades responsáveis pela execução do programa (como, por exemplo, a Irlanda do Norte – Ver a Caixa *As recomendações para uma avaliação: reduzir o número de indicadores de 330 para 52*). Isto não significa, contudo, que não possam ser necessários indicadores adicionais para satisfazer as necessidades de informações dos operadores.

Caixa: Recomendações para uma avaliação: reduzir o número de indicadores de 330 para 52

O FSE financiou um programa socioeconómico, entre 1994 e 1999, na Irlanda do Norte. Fez-se uma avaliação intercalar deste programa, sintetizando seis avaliações individuais dos subprogramas. Os responsáveis pela concepção do programa seleccionaram 330 indicadores de monitorização. Estes indicadores foram agrupados numa única base de dados, que se revelou de difícil utilização, e criou problemas de manipulação e disponibilidade de informações. Estes problemas reduziram consideravelmente a utilização dos indicadores. A equipa de avaliação propôs uma solução, que consistia em:

- Seleccionar uma série de indicadores contextuais, situados entre os indicadores macroeconómicos e os indicadores dos subprogramas. Estes indicadores, que têm como objectivo reflectir o impacto global do programa, dividem-se em três categorias: crescimento económico, coesão interna e coesão externa.
- Seleccionar um pequeno número de indicadores do programa, limitando-os aos principais resultados e impactos;
- Delegar nos operadores a quantificação e utilização dos outros indicadores.

Desta forma, a dimensão do sistema seria reduzida a 52 indicadores directamente relacionados com os principais objectivos do programa. As recomendações foram aplicadas com sucesso.

Fonte: Colin Stutt Consulting, (1997) Northern Ireland Single Programme 1994-99; Relatório intercalar da avaliação externa.

Uma possível abordagem para limitar a dimensão dos sistemas de indicadores sem perder informações importantes consiste em identificar indicadores genéricos, ou agrupar os indicadores por categorias de beneficiários.

Sugestões para limitar a dimensão dos sistemas de indicadores

As sugestões para se limitar a dimensão dos sistemas de indicadores baseiam-se geralmente na utilização de indicadores genéricos, ou no agrupamento de indicadores por categoria de beneficiários. Um sistema mais leve limita a recolha e a circulação de informações aos elementos mais essenciais, a nível de programa. Por outro lado, isto significa que a evolução e os resultados de cada uma das acções não serão monitorizados de forma detalhada e centralizada. Também significa que o sistema se centra menos nas decisões a adoptar pelos operadores, e mais nas decisões a adoptar pelas autoridades de gestão do programa.

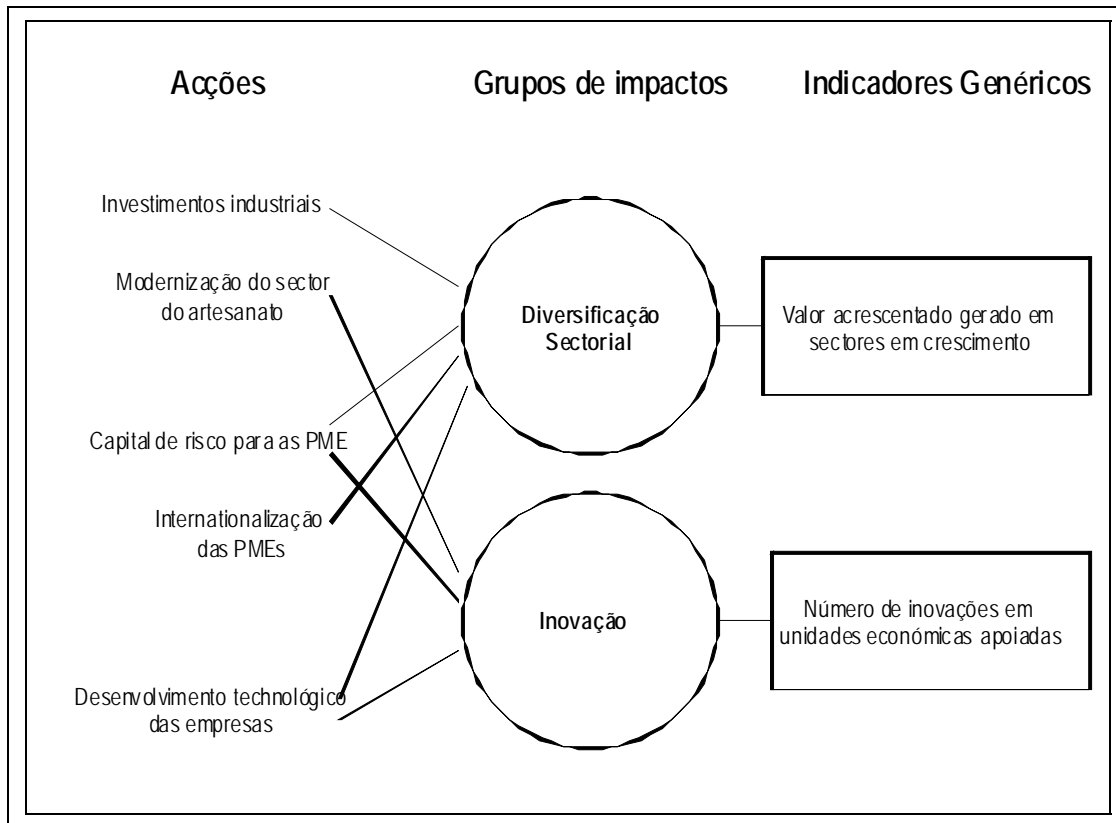
Indicadores nucleares

Na fase de preparação para a actualização da Avaliação Intercalar de 2005, a Comissão Europeia introduziu o conceito de "indicadores nucleares", na documentação de orientação correspondente. A ideia consistia em identificar alguns indicadores básicos (sobretudo resultados) que poderiam ser agregados e comparados (com precaução) entre diversas regiões e Estados-membros. Os indicadores nucleares foram revistos e desenvolvidos nas orientações da Comissão sobre os indicadores, no período entre 2007 e 2013. Incentivando a utilização de indicadores nucleares, o documento de orientação também destaca a necessidade de as autoridades responsáveis desenvolverem indicadores de resultados de melhor qualidade, em conformidade com os objectivos do programa e com o enfoque estratégico mais forte do novo período de programação.

Procurar indicadores de impacto genéricos

Os indicadores de impacto são indispensáveis na avaliação, mas são difíceis de conceber e de quantificar. À excepção do número de empregos criados, é particularmente raro encontrarem-se indicadores de impacto genéricos na programação de documentos (apresentam-se alguns exemplos na Caixa *Exemplos de indicadores de impacto genéricos*).

Caixa: Exemplos de indicadores de impacto genéricos



Avaliação da qualidade

A utilidade de indicadores é muito maior se a sua qualidade aumentar constantemente. A avaliação desempenha um papel importante na avaliação da qualidade dos sistemas de indicadores e na recomendação de formas de os melhorar. Embora não haja um método normalizado para realizar este controlo de qualidade, propõe-se uma abordagem com base nos critérios que se seguem, divididos em dois grupos: critérios de qualidade aplicáveis a cada um dos indicadores e critérios de qualidade a todo o sistema.

Critérios de qualidade aplicáveis a cada um dos indicadores

O primeiro critério de qualidade de um indicador é constituído pela capacidade de o quantificar a intervalos regulares. Por vezes, um ou mais indicadores incluídos nos documentos de programação não terão, ainda, sido quantificados, e, por conseguinte, não podem determinar a avaliação ou a monitorização do programa. A disponibilidade dos dados para permitir a quantificação constitui o factor primário a considerar. Os indicadores de monitorização (recursos, realizações e alguns resultados) deverão ser quantificados em cada uma das reuniões de monitorização, ou seja, a cada seis ou doze meses. Os indicadores de avaliação (determinados resultados e impactos, bem como indicadores de contexto) são quantificados com menor frequência, por norma anualmente ou a cada três ou seis anos.

Depois de um indicador ser quantificado, poderá levar vários meses – ou mesmo anos – até que as informações sejam realmente utilizadas para monitorização e avaliação. É muito particularmente o caso de alguns indicadores de contexto decorrentes de publicações estatísticas. A actualidade das informações constitui um importante critério de qualidade. Por vezes, as estatísticas são publicadas dois ou mais anos após a recolha dos dados.

Durante a avaliação dos efeitos do programa, os indicadores devem ser seleccionados de forma que o programa possa introduzir alterações ao valor do indicador. A capacidade das intervenções para influenciar determinado indicador é conhecida como sensibilidade. Consideremos como exemplo uma intervenção de apoio à exportação: os resultados efectivos das empresas não constituem um indicador suficientemente sensível. Um indicador melhor seria os resultados efectivos relacionados apenas com novos clientes, contactados com o apoio do programa.

Os resultados produzidos na sequência da aplicação dos indicadores devem ser fiáveis e credíveis. A fiabilidade costuma aplicar-se a factos e números, e pode definir-se como o facto de duas medidas, adoptadas por duas pessoas diferentes, sob circunstâncias semelhantes, produzirem o mesmo valor para o indicador. Nos casos em que os indicadores são quantificados com base nas questões colocadas por uma pessoa a outra, a fiabilidade deixa de poder definir-se em termos tão mecânicos, embora os testes tenham que ser credíveis. A credibilidade depende geralmente da solidez do método, embora a independência e a reputação da equipa de avaliação também sejam importantes.

A utilidade de um indicador depende, em grande parte, do facto de permitir fazer comparações internas entre diferentes medidas do programa, ou comparações externas inter-regionais. A comparabilidade do indicador constitui, por isso, um critério importante.

Um outro critério de qualidade de um indicador é a normatividade. Os indicadores devem estar relacionados com os resultados que podem ser considerados satisfatórios ou não. Os indicadores devem evitar a ambiguidade. Todos os valores dos indicadores devem, assim, ser comparados com uma norma, por exemplo: objectivos a cumprir; norma a ultrapassar; ou média europeia a atingir.

Um bom indicador deve ser compreendido por todos os que o utilizam. Para os decisores e para o público, o significado do indicador deve ser o mesmo para os gestores de programa e para os promotores do projecto. Deve reflectir de forma precisa o conceito que se pretende medir. Isto designa-se, muitas vezes, por validade.

Critérios de qualidade aplicáveis à totalidade do sistema de indicadores

Propõe-se os critérios que se seguem para avaliar os sistemas de indicadores:

- Os indicadores seleccionados devem abranger uma proporção suficientemente grande das medidas do programa. Esta cobertura deve ser igual ou superior a três quartos das despesas previstas.
- O sistema deve ser constituído por um bom equilíbrio entre os indicadores das diferentes categorias.
- O sistema de indicadores deve ser simples. O critério de selecção exige que se respeite a capacidade dos gestores de programas para absorver informações. As informações devem, por isso, limitar-se a um máximo de algumas dezenas de indicadores.
- A relevância do sistema implica que os indicadores sejam desenvolvidos, em primeiro lugar, pelas medidas ou temas com implicações significativas em termos de tomada de decisão. Por exemplo, medidas com um orçamento muito elevado, medidas inovadoras, e temas considerados estratégicos.

Em muitos casos, a configuração dos indicadores não se faz de raiz, e, sempre que possível, os sistemas e os indicadores devem ser consistentes com os que já se encontram em funcionamento.

Aplicação dos indicadores

Utilização dos indicadores para fazer comparações entre programas

A utilização de indicadores para fazer comparações válidas entre programas é importante, mas também é difícil. Isto deve-se a diversos factores, como a diversidade das intervenções num programa, a diversidade de contextos regionais, ou a incompatibilidade de definições. Por exemplo, dependendo das regiões e dos programas, as viagens turísticas podem ser contabilizadas em termos do número de visitas ou do número de noites de estadia; os formandos podem ser contabilizados em termos de número de participantes no curso de formação ou em horas de formação efectivas; e a protecção ambiental pode ser medida em termos de número de projectos, número de sítios ou número de hectares protegidos.

É possível procurar e assegurar a comparabilidade através de intercâmbios entre os gestores em diferentes regiões ou países. O programa INTERREG proporcionou diversas oportunidades de fazer este tipo de comparação (ver a Caixa *A cooperação como forma de gerar comparações transnacionais*).

Caixa: A cooperação como forma de gerar comparações transnacionais

As regiões de Pays-de-la-Loire (França) e Emilie-Romagne (Itália) organizaram reuniões bilaterais em 1997-1998 para discutir o seu trabalho de avaliação, no domínio do turismo rural. Durante dois seminários de dois dias, as regiões analisaram as suas respectivas acções, e definiram, em conjunto, um diagrama lógico de efeitos, resultados e impactos. Propuseram-se indicadores comuns, com o objectivo de permitir comparações inter-regionais.

A comparabilidade é, muitas vezes, mais fácil de alcançar, e mais útil a todos os níveis, se for resultante de um esforço colectivo coordenado, a um nível geográfico superior. Esta abordagem tem a vantagem de multiplicar as possibilidades de comparação, como também de permitir a agregação de indicadores, a nível regional ou nacional. O Gabinete Escocês criou uma série de indicadores padrão, aplicáveis a sete programas escoceses (ver a Caixa *Um conjunto de indicadores padrão para diversos programas?*). Apesar de se ocupar sobretudo dos indicadores contextuais e menos com a avaliação, a Urban Audit representa um exemplo dos esforços adoptados para conseguir a comparabilidade.

Caixa: Um conjunto de indicadores-padrão para diversos programas?

No âmbito da sua responsabilidade de monitorização e de avaliação de sete programas co-financiados pela União Europeia, a *Scottish Executive* desenvolveu uma série de indicadores padrão. Alguns destes indicadores estão relacionados com o contexto: taxa de emprego e de desemprego, produtividade, mão-de-obra, rendimento médio, etc. Alguns indicadores de efeitos e de resultados também foram normalizados, como ilustram os exemplos que se seguem, aplicáveis às medidas de infra-estruturas das empresas:

- Construção de novos edifícios (metros quadrados)
- Edifícios renovados (metros quadrados)
- Desenvolvimento de novos sítios (hectares)
- Melhoria dos locais existentes (hectares)
- Acessos rodoviários novos/melhorados (quilómetros)
- Área de superfície em que foram construídas ou melhoradas as rodovias de acesso (hectares)
- Taxa de ocupação dos novos edifícios (percentagem após um ano/percentagem após três anos)
- Número de centros de formação criados (número)

Para obter um elevado nível de comparabilidade, os gestores de programa escoceses organizaram diversas reuniões, nas quais participaram representantes de todos os programas. A normalização foi o resultado de um longo processo de discussão colectiva.

Comunicação pública

Os sistemas de indicadores devem ser úteis à tomada de decisão. Também são importantes para efeitos de responsabilização, por exemplo para o Parlamento Europeu ou para os parlamentos nacionais, para os representantes eleitos a nível regional ou local, para os parceiros socioeconómicos, para os jornalistas, e, através deles, para os cidadãos e para os contribuintes.

Se o objectivo dos sistemas de indicadores for servir de base à comunicação pública, é necessário seleccionar, quantificar e publicar um pequeno número de indicadores que possam ser facilmente compreendidos por leigos. A publicação destes indicadores encontra-se geralmente organizada sob a forma de quadros simples, acompanhados de comentários, por exemplo num relatório anual. Também podem ser colocadas à disposição informações mais detalhadas, através de um "observatório" aberto ao público, ou através de um sistema de consulta na Internet.

Ao definir estes indicadores publicamente acessíveis, deverá dar-se prioridade a indicadores genéricos (aplicáveis a várias acções diferentes, no mesmo programa) e a indicadores padrão (que permitam fazer comparações entre programas em diferentes regiões ou países). Além disso, estes indicadores devem ser compreendidos por todos os parceiros sem explicações extensas, e sem quaisquer falhas na compreensão do seu significado.

Propostas de indicadores-chave de acesso público

A Caixa *Propostas de indicadores-chave* propõe uma série de indicadores que podem ser utilizados como indicadores-chave de acesso público. Esta lista não é fechada, obrigatória, ou estável, e encontra-se limitada aos indicadores do programa. Os

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

indicadores foram pontuados, de *** a *, por ordem decrescente de interesse. Os indicadores são correspondentes a: recursos, efeitos, resultados e impactos.

Caixa: Propostas de indicadores-chave

Recursos

Interesse	Indicador
	Recursos humanos
**	Emprego temporário nas empresas que executam o trabalho durante a implementação (empregos x anos)
*	Número de operadores (organizações públicas e privadas, responsáveis pela prestação de assistência aos beneficiários)
*	Número de conselheiros (FTE) mobilizados para aconselhar os beneficiários
	Recursos financeiros
***	Taxa de absorção orçamental (% de fundos alocados)
**	% projectos (em termos financeiros) particularmente orientados para as mulheres
**	% projectos (em termos financeiros) em mercados/sectores em rápido crescimento
*	% de orçamento reservada para medidas de mitigação ambiental
*	% projectos (em termos financeiros) orientados para as áreas mais desfavorecidas

Realizações

Interesse	Indicador
	Progresso dos trabalhos
***	Taxa de conclusão (% do objectivo)
**	Conformidade com a duração do projecto
	Capacidade dos trabalhos terminados
**	Número de potenciais ligações (empresas/habitações) à rede de serviços básicos (distribuição por serviços)
	Actividade dos operadores em termos de atracção e selecção de participantes
**	Taxa de selecção (% de projectos aceites, proporcionalmente aos projectos elegíveis)
**	Taxa de cobertura (penetração): % da população-alvo que participou (deveria participar) no programa
*	% de beneficiários pertencentes aos grupos prioritários (por exemplo, desempregados de longa duração, crianças com abandono escolar precoce)
*	% de beneficiários localizados nas áreas mais desfavorecidas
**	% de beneficiários envolvidos em mercados/sectores em rápido desenvolvimento
***	% de mulheres entre os beneficiários
***	% de PME entre os beneficiários
	Serviços financiados pelo programa
***	Número de beneficiários individuais que usufruíram de serviços, aconselhamento, formação
***	Número de unidades económicas (empresas, unidades agrícolas,

	armadores, viveiros, turismo profissional) que usufruíram de serviços, aconselhamento, formação
**	Número de horas de formação/aconselhamento concedidas aos beneficiários

Resultados

Interesse	Indicador
	Satisfação dos beneficiários
*	Taxa de satisfação (% de beneficiários que estão satisfeitos ou altamente satisfeitos)
	Vantagens obtidas pelos beneficiários
**	Velocidade média entre os principais centros económicos
	Investimentos facilitados aos beneficiários
**	Efeito de alavanca (gastos do sector privado decorrentes como contrapartida do apoio financeiro recebido)

Impactos

Interesse	Indicador
	Sucesso sustentável
**	Taxa de colocação (por exemplo, % de beneficiários individuais a trabalhar após 12 meses, inc. % num emprego estável a longo prazo)
*	Taxa de sobrevivência (por exemplo, % de unidades económicas apoiadas, ainda activas após 12/36 meses)
	Impacto percebido pelos beneficiários
***	Valor acrescentado gerado (por exemplo, após 12 meses em termos de euros/ano/colaborador,)
***	Emprego criado ou salvaguardado (por exemplo, após 12 meses, equivalente a tempo inteiro)
	Impacto globalmente percebido na área
**	Atractividade residencial (por exemplo, % de habitantes que pretendem permanecer na área)
	Impacto indirecto
*	Efeitos de arrastamento regionais (por exemplo, % de empresas regionais entre os prestadores de serviços às empresas apoiadas após 12 meses)

4.8. Utilização de indicadores para melhorar a gestão

- Gestão do desempenho, em vez dos recursos
- Interpretação e comparação de indicadores
- Evitar os efeitos adversos dos indicadores
- A utilidade de um sistema de indicadores

Gestão do desempenho, em vez dos recursos

Um número crescente de indicadores tem sido criado para efeitos de "gestão de desempenho". Estes sistemas são uma forma da Nova Gestão Pública, que realça os

resultados e os impactos obtidos, faces às formas de gestão mais antigas, baseadas na alocação de recursos e no controlo de resultados.

Em consonância com o espírito de gestão de desempenho, é concedida aos operadores uma maior autonomia na utilização dos seus recursos. Em contrapartida, comprometem-se, eles próprios, com objectivos mais claros, no que diz respeito aos resultados e aos impactos a obter. Terão que medir o seu desempenho para procederem à auto-avaliação, e enviar relatórios periódicos. Este novo equilíbrio entre a descentralização e a medição do desempenho encontra-se na base de muitos desenvolvimentos na administração pública (por exemplo, o Performance and Results Act, GPRA, de 1993, do Governo dos EUA, tornou obrigatória a utilização de indicadores em toda a administração governamental).

Em muitas regiões europeias, a cultura administrativa permaneceu impermeável à descentralização e à gestão de desempenho, e o desenvolvimento dos indicadores de resultado e de impacto é geralmente considerado difícil. Os gestores dos programas estão mais familiarizados com os indicadores de recursos e de resultados. As mudanças culturais são, porém, lentas, mas ocorrem seguramente em determinados países, sob a pressão das reformas administrativas iniciadas pelos governos nacionais. A monitorização e avaliação de programas co-financiados pela União Europeia constituíram um factor que incentivou a gestão de desempenho em termos de resultados e impactos.

Interpretação e comparação de indicadores

Em determinadas situações, os indicadores falam por si mesmos, mas isto é a excepção. Regra geral, os indicadores têm que ser interpretados através da comparação ou descrição correspondente. Num exemplo, a comparação de três indicadores mostrou que a formação apoiada por financiamento da UE não chegou aos desempregados de longa duração, como deveria ter acontecido (ver a Caixa *Comparação de diversos indicadores como forma de revelar um fenómeno*).

Caixa: Comparação de diversos indicadores como forma de revelar um fenómeno

No âmbito da sua avaliação intercalar do FSE nas regiões italianas de Objectivo 1, a ISFOL reuniu e interpretou uma série de indicadores de monitorização. Em determinados casos, as comparações feitas entre estes indicadores conduziu a conclusões fortes e claras. Por exemplo, a parte do programa orientada para os desempregados de longa duração pode ser medida através de três indicadores:

- encargos planeados com desempregados de longa duração: 18% dos encargos totais planeados
- financiamento reservado aos desempregados de longa duração: 11% do total de financiamento alocado
- financiamento realmente despendido com os desempregados de longa duração: 8% do total de financiamento utilizado

A comparação destes números mostra claramente que os desempregados de longa duração foram desclassificados com êxito ao longo de cada uma das fases de implementação do programa. Para um leitor comum, que visse o quadro de indicadores e estes três números isoladamente, estas informações não teriam qualquer significado. Para identificar o problema de o financiamento não chegar aos desempregados de longa duração, foi necessário dividir estes três indicadores em termos de duração da situação de desemprego do beneficiário directo, e, para a ISFOL, comparar os três tipos de informações entre si.

Sempre que possível, é útil comparar os indicadores do programa (tais como quilómetros de auto-estradas construídos) com os indicadores de contexto correspondentes (por exemplo, quilómetros auto-estrada existentes na região). Este tipo de informações pode ser apresentado de forma concisa, através de indicadores daí decorrentes, que mostram os resultados ou os efeitos dos programas como uma percentagem do indicador de contexto.

Para serem úteis em avaliação, os indicadores de contexto devem ser utilizados juntamente com conclusões qualitativas. Para interpretar os indicadores, é necessário considerar o contexto como um todo, os factores que ajudam a promover ou a retrainir o desempenho do programa, a filosofia do programa, e o processo de implementação.

Uma técnica consiste em pedir a um painel de peritos para analisar os elementos quantitativos e qualitativos integrados da situação, para interpretar as medidas de desempenho (ver a Caixa *Painéis de peritos para avaliar o desempenho dos Centros de Investigação Universitários*).

Caixa: Painéis de peritos para avaliar o desempenho dos Centros de Investigação Universitários

Os painéis de peritos foram utilizados durante algum tempo para avaliar o desempenho de centros de investigação universitários no Reino Unido. O trabalho de todos os centros de investigação no Reino Unido é monitorizado em termos quantitativos, e avaliado periodicamente, por painéis de peritos. Na sua avaliação, os peritos consideram indicadores (por exemplo, o número de publicações), mas também elementos puramente qualitativos (por exemplo, qualidade de cada uma das quatro melhores publicações de cada investigador). A avaliação tem como resultado a classificação das equipas de investigação (numa escala de 1 a 5, com a classificação de uma estrela). A alocação de financiamento público está directamente relacionada com esta classificação (Bone & Carter, 1997).

A interpretação dos indicadores utiliza, muitas vezes, comparações entre o desempenho de duas intervenções. Este tipo de comparação diz respeito, sobretudo, aos indicadores de eficiência, por exemplo: custos de formação de um formando; custos de criação de um emprego; custos de uma habitação ligada ao sistema de saneamento. Porém, as comparações nem sempre são aceitáveis para os operadores e gestores de um programa. Em particular, as comparações de desempenho parecerão injustas aos gestores que trabalham nos contextos mais difíceis. Para tornar as comparações aceitáveis, os termos de comparação devem ser consideravelmente refinados, com o intuito de ter em conta os desafios adicionais das intervenções a decorrer em ambientes mais difíceis.

Apresentamos um exemplo da forma como esta questão foi processada, no domínio da educação, na Caixa Exemplo de uma comparação de desempenho aceite pelos operadores. A técnica consiste em reconstituir, em cada uma das intervenções a comparar, um conjunto de diversas intervenções semelhantes, implementadas noutras regiões, num ambiente com graus de dificuldade semelhantes. As comparações foram feitas de acordo com a média das intervenções deste "grupo de referência".

Caixa: Exemplo de uma comparação de desempenho aceite pelos operadores

A comparação de desempenho entre organizações de formação tropeça, muitas vezes, no facto de os diferentes grupos de beneficiários não terem as mesmas dificuldades. A questão reside em saber como fazer uma comparação em termos semelhantes. Há alguns anos, uma equipa de avaliação teve que comparar escolas de 133 zonas, em diferentes estados dos Estados Unidos. A equipa decidiu comparar cada uma das zonas com um "grupo de referência" de outras 14 escolas com características semelhantes em termos de alunos (pobreza, língua materna, nível de educação dos pais, pertença a um grupo minoritário, etc.). Para constituir os grupos-padrão, a avaliação combinou uma análise factorial com painéis de peritos. No final da avaliação, a validade das comparações de desempenho foi aceite sem quaisquer problemas.

Henry et al, (1992)

Evitar os efeitos adversos dos indicadores

A utilização de indicadores é, muitas vezes, afectada pelo receio de provocar efeitos adversos. Existem diversos tipos de efeitos adversos:

- O efeito de "favorecimento" ou de "filtragem";
- Convergência para a média;

- Efeitos inesperados, cujos resultados estão subordinados às pontuações dos indicadores.

Os efeitos de “filtragem” poderão ocorrer sempre que o desempenho de organizações de formação e de emprego for medida de acordo com a taxa de colocação de beneficiários. Para obter uma melhor taxa de colocação para os seus beneficiários, é do interesse das organizações recrutar pessoas na melhor situação possível, que também satisfaçam os critérios de elegibilidade. Por conseguinte, os operadores tendem a “filtrar” ou favorecer os beneficiários cuja empregabilidade é superior. Este efeito é indesejável porque concentra o apoio nos que têm relativamente menos necessidades.

Apresentamos um exemplo da forma como os indicadores provocaram uma redução das diferenças, através de uma convergência para a média, na Caixa *Efeitos adversos de um sistema de indicadores de gestão de desempenho: convergência para a média, em vez da convergência para a excelência*. Um indicador também pode incentivar um comportamento conducente a um desempenho inferior à norma. Este processo decorre sempre que os indicadores premiam resultados indesejados, ou sempre que o sistema dê origem a que os operadores trabalhem para os indicadores, em vez de trabalharem para os resultados.

Caixa: Efeitos adversos de um sistema de indicadores de gestão do desempenho: convergência para a média, em vez da convergência para a excelência

O sistema de indicadores da British Audit Commission (Comissão de Auditoria Britânica) organiza a quantificação de cerca de duas centenas de indicadores de efeitos e de resultados, relacionados com os serviços oferecidos pelos municípios. Os indicadores são comparáveis entre autarquias e publicados anualmente na imprensa local. Este sistema permite criar uma imagem muito forte da cidade quando um destes serviços regista um desempenho insuficiente. Consequentemente, muitas autarquias aumentaram o orçamento dos serviços que revelaram um desempenho mais reduzido, na esperança de elevarem o nível de eficácia. Nestes casos, os recursos financeiros retiravam-se por vezes aos serviços mais eficazes. A utilização dos indicadores provocou, assim, uma redução nas diferenças através de uma convergência para a média. Constituiu um efeito adverso porque se esperava que os indicadores provocassem a convergência do desempenho para a excelência.

Os efeitos adversos surgem inevitavelmente após dois ou três anos de funcionamento de um sistema de indicadores, independentemente de ter sido bem concebido. Estes efeitos indesejáveis não são normalmente previsíveis.

A possível emergência de efeitos adversos não deverá constituir um argumento para recusar a medição do desempenho. É possível minimizar os efeitos adversos, seja corrigindo os indicadores que estão na origem do problema, seja criando um procedimento de interpretação dos indicadores, por painéis de peritos. É, pois, importante observar o aparecimento de efeitos adversos e corrigir o sistema, sempre que surgirem estes efeitos.

Criação de incentivos de desempenho

Existem diversas formas de utilizar os indicadores para incentivar a melhoria do desempenho dos operadores. Entre estas, incluem-se:

- Os operadores com níveis de desempenho reduzidos recebem apoio técnico específico, para os ajudar a progredir. Se a situação não melhorar, o

orçamento será restringido. Este método baseia-se no princípio de que não é o erro que deve ser penalizado, mas sim a incapacidade de corrigir os erros.

- Aos operadores com o melhor desempenho, é concedida uma maior autonomia e um menor controlo.
- Os operadores com o melhor desempenho receberão apoio para apresentar os seus resultados e efeitos ao público em geral.
- Os operadores que não tenham obtido um desempenho suficiente são desqualificados nos procedimentos de selecção para projectos futuros (apresentamos um exemplo deste facto na Caixa Seleccção de projectos com base na sua eficácia).
- Aos operadores com o melhor desempenho, é concedido apoio adicional.

Caixa: Seleccção de projectos com base na sua eficácia

No âmbito do Objectivo 2 do programa correspondente ao período entre 1989 e 1993 na região do Pays-de-la-Loire (F), as acções do FSE foram geridas pelo Conselho Regional. Para fazer a gestão do FSE, aplicaram um método de selecção de projectos baseado na gestão de desempenho.

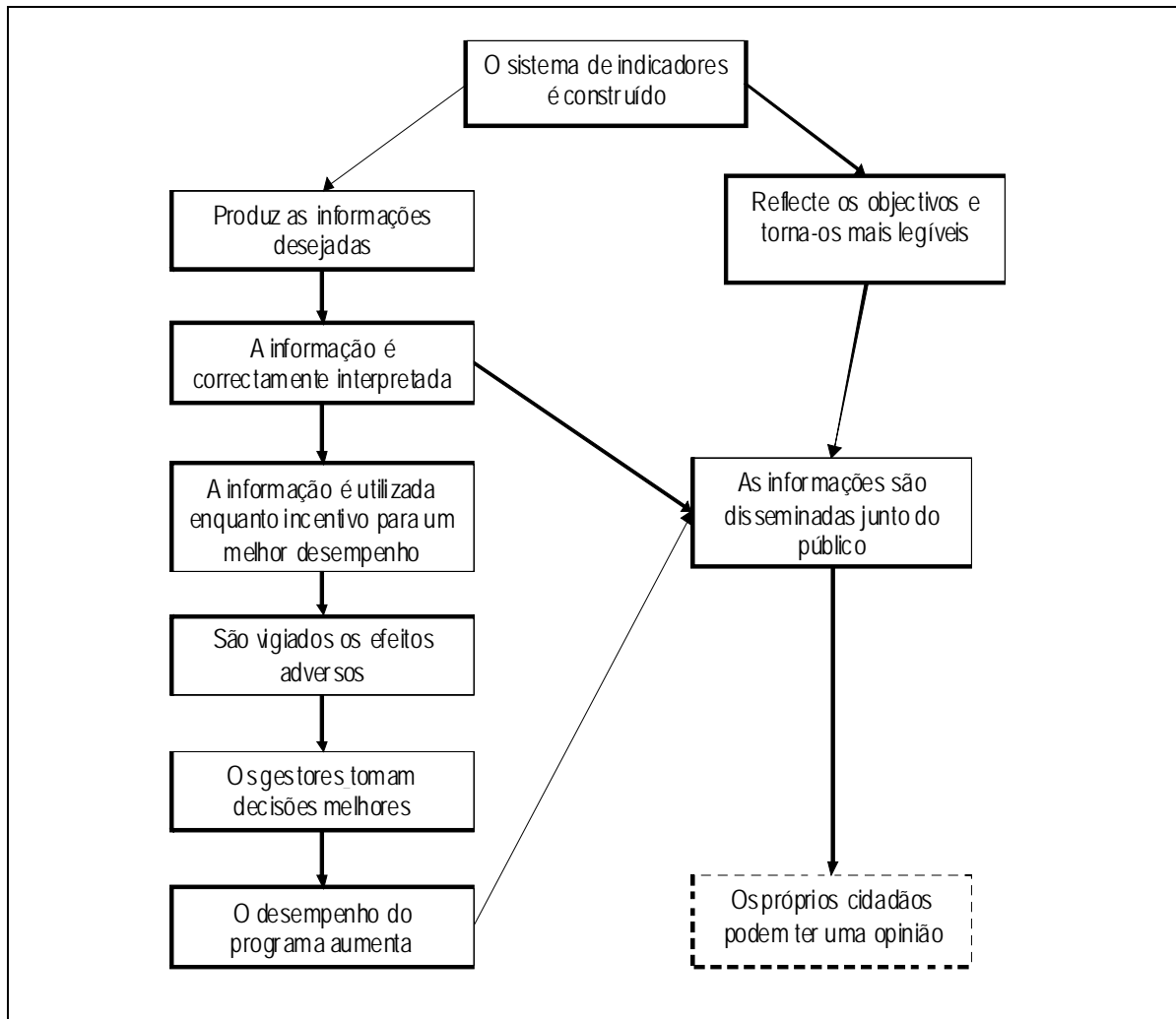
O sistema de monitorização obriga todas as organizações que promovem formação subsidiada a proporcionar ao Conselho Regional uma avaliação da taxa de colocação dos seus formandos, depois de decorrido um período de seis meses. A candidatura a financiamento secundário/adicional é assegurada, em primeiro lugar, às organizações com as melhores taxas de colocação. O objectivo deste mecanismo é beneficiar as organizações que estão mais sintonizadas com as necessidades das empresas e que possuem os melhores impactos.

A utilidade de um sistema de indicadores

A Caixa Um sistema de indicadores, apresentada a seguir, resume as principais características descritas acima. Apresenta um conjunto de pressupostos constituídos, muitas vezes implicitamente, após a criação de um sistema de indicadores, com a intenção de melhorar o desempenho de um programa.

A apresentação realça o facto de a construção de um sistema de indicadores não ser efectivamente um primeiro passo. Decorrem pelo menos seis outros passos antes de o sistema ser realmente útil.

Caixa: Um sistema de indicadores



4.9. Regras de ouro

1. A selecção de métodos e técnicas decorre directamente do tipo de questões que se pretende colocar, e estas questões fazem parte de um vasto exercício de concepção, que inclui a consulta dos parceiros e a avaliação das características do programa. A selecção, em primeiro lugar, de métodos e técnicas, e a tentativa, depois, de os adaptar às questões para as quais não foram especificamente escolhidos, irá criar sempre problemas. As técnicas seleccionadas devem reflectir o objectivo e o enfoque da avaliação.
2. A maior parte das técnicas possui forças e fraquezas, que é necessário reconhecer. E, sempre que possível, devem aplicar-se as diferentes técnicas em conjunto, para reforçar a análise e tornar os resultados e as conclusões da avaliação mais fiáveis.
3. Devido ao carácter distintivo do desenvolvimento socioeconómico: ascendente, utilizando diferentes combinações de intervenções, e com adequação às necessidades territoriais e sectoriais, é difícil medir e comparar resultados entre contextos de programas de desenvolvimento socioeconómico. O que não significa que a medição, quantificação e estatísticas não sejam relevantes. Podem ser ferramentas capazes, sempre que se faça comparações ao nível do programa de desenvolvimento particular, e não tentem comparar contextos incomparáveis.
4. Os métodos e técnicas qualitativos são adequados ao desenvolvimento socioeconómico, devido à subtilidade e natureza holística do que se procura

- realizar, e às diferenças entre contextos, que é necessário descrever em termos qualitativos. A natureza participativa da capacidade de desenvolvimento local relativamente ao potencial e às ambições dos cidadãos e dos parceiros locais é particularmente adequada aos métodos qualitativos e aos métodos participativos.
5. As prioridades temáticas, que são muito comuns em programas europeus, colocam verdadeiras dificuldades aos avaliadores. Uma vez que os decisores políticos pretendem compreender a forma como as suas políticas, como um todo, têm êxito, faz-se, muitas vezes, pressão para integrar resultados e encontrar uma forma comum de descrever ou, inclusivamente, avaliar o que está a acontecer. Isto não pode, muitas vezes, fazer-se. Por vezes, apenas as descrições qualitativas é que funcionam. É necessário ter cuidado para não misturar “alhos” e “bugalhos”.
 6. Existe, muitas vezes, uma tensão entre seleccionar avaliadores que têm vastos conhecimentos acerca de uma área de política específica e avaliadores cujas capacidades de avaliação são mais genéricas. Numa situação ideal, numa equipa de avaliação procura-se incluir estes dois conjuntos de conhecimentos e de experiência. Os requerentes devem estar conscientes dos perigos inerentes à contratação de avaliadores que viveram a maior parte da sua vida profissional a trabalhar numa área especializada e a utilizar um conjunto limitado de métodos e técnicas. Este é outro argumento a ter em conta para obter um equilíbrio de capacidades na equipa de avaliação.
 7. É importante fazer a distinção entre métodos e técnicas para recolher dados, para analisar dados e para transmitir conclusões da avaliação. Esta distinção nem sempre é feita, em parte porque os que realizam avaliações podem estar mais preocupados com uma fase do processo do que com outra. Como em tudo, é necessário haver um equilíbrio. Não vale a pena investir em métodos sofisticados de recolha de dados se depois a análise desses mesmos dados for relativamente simplista.
 8. Os dados nunca são puros, nem nunca ocorrem naturalmente; necessitam de ser produzidos. Por isso, os avaliadores têm que saber de onde provêm os dados e quais as decisões tomadas no decurso da sua produção. No final, a força dos argumentos e as conclusões que se poderá retirar dependem das forças e das características dos dados utilizados.
 9. Um aspecto importante da qualidade da avaliação dos dados decorre da forma de acesso aos dados, e da forma de negociação desse acesso. Sobretudo nos programas de desenvolvimento socioeconómico, existe uma série de problemas que é necessário resolver. É necessário que os diferentes parceiros estejam dispostos a partilhar informações; os grupos excluídos não confiam, muitas vezes, na avaliação como um exemplo adicional do comportamento oficial, pelo que é necessário suscitá-los. Além disso, é necessário convencer todos os parceiros de que irão retirar algo da avaliação, antes de fornecerem acesso, com qualquer entusiasmo, a quaisquer informações que retenham. O investimento nestes tipos de negociação fará a diferença em termos de qualidade e de avaliação como um todo.
 10. A relação entre quantitativo/qualitativo é sobrevalorizada. Os dados constituem, muitas vezes, parte de um contínuo, começando como sendo qualitativos, e tornando-se, após a análise, quantitativos. As formas mais fracas de dados quantitativos (por exemplo, categorizações ou *rankings*) estão mais próximas dos dados qualitativos. O que é necessário, aquando da avaliação dos programas de desenvolvimento socioeconómico, é que os dados qualitativos possam captar as nuances, a experiência das pessoas e as opiniões, e que os dados quantitativos forneçam perspectivas, por exemplo no sentido de agregar resultados de uma intervenção e proporcionar uma perspectiva comparativa.
 11. Os sistemas de monitorização e os indicadores bem concebidos são um complemento capaz da avaliação. Muitas vezes, os avaliadores estão

dependentes dos sistemas de monitorização, que são baseados em indicadores. Se estes não forem aplicados numa fase precoce do ciclo de concepção do programa, mais tarde poderá ser demasiado tarde para criar estes sistemas.

12. Os sistemas de indicadores demasiado elaborados podem ser contraproducentes. Embora nos programas multi-sectoriais e multi-objectivos haja uma tendência para medir todos os pormenores, esta tendência deverá ser contrariada para limitar os seus custos e permitir usar os seus resultados.
13. Os indicadores são utilizados, muitas vezes, para efeitos de gestão e de responsabilização. Pode ser difícil reutilizar os indicadores exclusivamente desenvolvidos para estes objectivos como parte de uma avaliação. Poderá haver demasiada pressão para delinear as informações de formas positivas ou, pelo menos, no sentido de se certificar que as más notícias não passam. Por outro lado, numa cultura de gestão de desempenho bem desenvolvida, estes tipos de indicadores podem ajudar a melhorar os conteúdos dos programas, bem como a sua gestão.

ANEXO A AO GUIA – AS PRINCIPAIS ETAPAS DA AVALIAÇÃO

- **Avaliação ex-ante**
- **Avaliação intercalar**
- **Avaliação ex-post**
- **Ciclos de programação sucessivos**

Avaliação ex-ante

A avaliação ex-ante decorre no início do ciclo, antes da adopção de um programa. Esta forma de avaliação ajuda a garantir que o programa final é tão relevante e coerente quanto possível. As suas conclusões têm como objectivo serem integradas no programa aquando da tomada de decisões.

A avaliação ex-ante centra-se, em primeiro lugar, numa análise das forças, fraquezas e potencial do Estado-Membro, região ou sector correspondente. Proporciona às autoridades em causa uma apreciação sobre se as questões de desenvolvimento foram correctamente diagnosticadas, se a estratégia e os objectivos propostos são relevantes, e se existe algum tipo de incoerência em relação às políticas e orientações comunitárias, se os impactos esperados são realistas, e sucessivamente. Também proporciona os fundamentos necessários para a monitorização e avaliações futuras, assegurando a existência de objectivos explícitos e, sempre que possível, quantificados. Ajuda a especificar os critérios de selecção relativos à selecção de projectos, e a garantir o respeito das prioridades comunitárias. Finalmente, ajuda a assegurar a transparência das decisões, proporcionando uma explicação clara das escolhas adoptadas e dos seus efeitos esperados.

As avaliações ex-ante são executadas no momento em que as autoridades públicas se encontram envolvidas em discussões e negociações acerca do programa futuro. Estão, por conseguinte, sujeitos a fortes constrangimentos: pressão dos prazos, formalização vaga do programa proposto para avaliação, emendas a esta proposta no decorrer dos trabalhos, exigências de confidencialidade, etc. A equipa de avaliação deve, por conseguinte, poder intervir de forma flexível e rápida, bem como aplicar técnicas de análise de necessidades e simulação de efeitos socioeconómicos.

Avaliação intercalar

A avaliação intercalar é executada no decorrer da segunda fase do ciclo de programação, durante a implementação das intervenções.

Dependendo das conclusões da avaliação intercalar, poderá fazer-se ajustes durante o ciclo. Esta avaliação analisa criticamente os primeiros efeitos e resultados das intervenções. Também avalia a gestão financeira do programa, e a qualidade da monitorização e da sua implementação. Mostra se e de que modo as intenções originais foram executadas, e, quando aplicável, verifica se foram efectuadas alterações *de facto* aos objectivos iniciais. Comparativamente à situação inicial, realça as alterações ao contexto económico e social geral, e avalia se os objectivos continuam a ser relevantes. A avaliação intercalar também analisa se a evolução das prioridades e políticas da Comunidade coloca problemas de coerência, e ajuda a preparar ajustes e a fazer a reprogramação, e a argumentá-los de forma transparente.

A avaliação intercalar baseia-se, em grande parte, nas informações retiradas do sistema de monitorização, mas também na avaliação ex-ante e nas informações sobre o contexto

e sobre a sua evolução. Geralmente, consiste em exercícios breves e exaustivos, centrados sobretudo nos resultados do programa avaliado, sem tentar realizar uma análise aprofundada dos impactos que ainda não tiveram tempo de emergir. É, no entanto, possível e aconselhável ter em conta avaliações aprofundadas ou temáticas de programas anteriores, sempre que disponíveis. A avaliação intercalar possui uma natureza "formativa", ou seja, produz *feedback* directo sobre o programa que está a ajudar a melhorar, em termos da sua gestão.

No período 2007-2013, a avaliação intercalar foi substituída pela avaliação intermédia ou contínua, de preferência baseada num plano de avaliação plurianual. Esta nova abordagem tem como objectivo ultrapassar os aspectos de rigidez da avaliação intercalar de 2003, e facilitar a avaliação de todas as partes dos programas numa altura em que os responsáveis pelos programas necessitam dessas conclusões.

Avaliação ex-post

A avaliação ex-post recapitula e tece considerações acerca de todo o programa, sobretudo em termos dos seus impactos. O seu objectivo consiste em justificar a utilização de recursos e reportar a eficácia e a eficiência das intervenções, e saber em que medida se alcançou os efeitos esperados. Centra-se em factores de sucesso ou de fracasso, e na sustentabilidade dos resultados e dos impactos. Procura retirar conclusões que se pode generalizar e aplicar a outros programas ou regiões.

Numa situação ideal, os resultados desta avaliação devem estar disponíveis quando se planear o programa seguinte, isto é, pelo menos um ano antes do final do programa. Porém, para que os impactos tenham sido produzidos, a avaliação ex-post teria que ser realizada dois a três anos após o final do período de programação. Enquanto se aguarda que este período passe, é pedido, muitas vezes, um relatório provisório pouco antes do final do ciclo de programação, em relação com a avaliação ex-ante do ciclo seguinte.

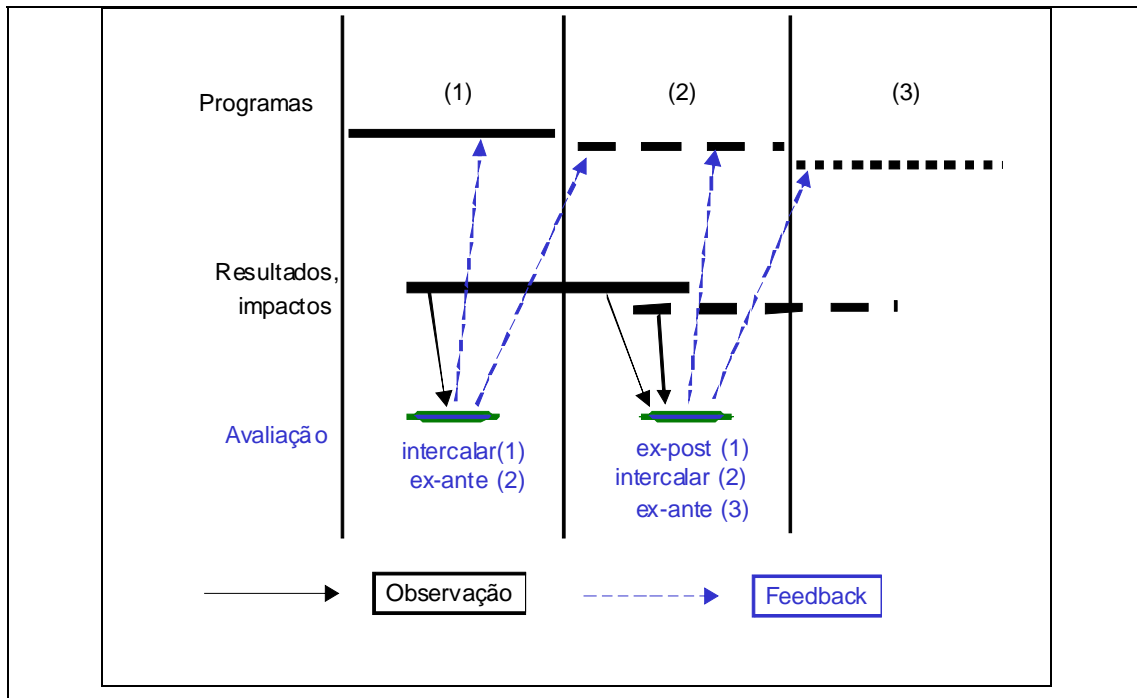
A análise de impacto constitui, sempre, um exercício a larga escala, se executada sistematicamente. As avaliações ex-post têm tendência, por isso, para envolver inquéritos de campo, e decorrer por longos períodos, entre doze e dezoito meses.

Ciclos de programação sucessivos

A sequência de três fases de avaliação em ciclos sucessivos cria sobreposições que é necessário organizar da forma mais eficiente possível, para evitar a sua duplicação. O princípio básico consiste na combinação do trabalho de avaliação, durante um programa, com a utilização de conclusões de avaliação executadas sobre o programa anterior. A continuidade relativa das acções programadas de um período para o seguinte permite utilizar conclusões recentes para considerar a relevância das novas medidas propostas. O esquema que se segue mostra que poderão decorrer interacções ao nível das avaliações executadas nas diferentes fases dos diversos programas sucessivos.

Caixa Articulação dos ciclos de programação da avaliação

Caixa: Articulação dos ciclos de programação da avaliação



Assim, uma avaliação ex-ante que prepare a adopção de um futuro programa deverá poder tirar partido dos resultados do trabalho anterior, i.e.:

- a avaliação intercalar do período que está a chegar ao fim. Esta avaliação deverá produzir conclusões durante os primeiros anos de actividade, e relativamente aos programas seguintes. Poderá ser concluída com um relatório final sobre os efeitos e os resultados do programa actual, com base em informações do sistema de monitorização.
- a avaliação ex-post do período anterior ao período actual, possivelmente complementada por avaliações temáticas e análises aprofundadas. Estas avaliações terão possibilitado a observação e análise dos impactos das intervenções anteriores, que são semelhantes às intervenções planeadas, e ocorreram num contexto parcialmente semelhante.

Uma vez que a função da avaliação intercalar consiste em verificar se os objectivos ainda são relevantes, e se estão em processo de serem alcançados, será necessário basear-se, em primeiro lugar, nos dados do sistema de monitorização, e ainda:

- na avaliação ex-ante, e, sobretudo, no diagnóstico realizado relativamente ao contexto socioeconómico prevaiente, antes do início do programa, e que é necessário actualizar;
- na avaliação ex-post do programa anterior, cujas conclusões relativamente às mesmas áreas de intervenção podem servir como base de referência.

A avaliação ex-post baseia-se em dados de gestão e de monitorização, e em estudos de campo, que ajudam a observar e a analisar os impactos reais e sustentáveis das intervenções. Baseia-se na avaliação ex-ante, na medida em que tem que reportar o alcance dos objectivos. Baseia-se na avaliação intermédia, sobretudo para identificar o sucesso ou os fracassos identificados nesta fase.

No que diz respeito ao facto de a avaliação necessitar de retirar conclusões da experiência de programas anteriores para melhorar programas futuros, uma solução interessante consiste em definir um plano de avaliação plurianual, conforme recomendado pela Comissão Europeia, para o período 2007-2013 (e para os novos

EVALSED: A Avaliação do Desenvolvimento Socioeconómico – O GUIA

Estados-membros, em 2004-2006). A ideia consiste em identificar as diferentes avaliações possíveis, e em definir a sua data e o seu conteúdo relativamente aos prazos e às agendas políticas para efeitos de tomada de decisão dos "clientes" da avaliação, a diversos níveis.

ANEXO B AO GUIA – EVOLUÇÃO DAS REGRAS DOS FUNDOS ESTRUTURAIS

- Regulamentos aplicáveis entre 1988-93
- Regulamentos aplicáveis entre 1994-99
- Regulamentos aplicáveis entre 2000-2006
- Regulamentos aplicáveis no período entre 2007-2013

Regulamentos aplicáveis entre 1988-93

Na Comissão, as exigências de transparência para com os contribuintes, relativamente à utilização sólida de fundos dos Estados-membros, foi progressivamente ganhando terreno. Esta aspiração reflecte-se numa expressão simples: "lutar pela eficácia dos gastos Europeus".

O Acto Único Europeu, adoptado no início de 1986, introduziu no Tratado da CEE uma nova secção V, incluindo o artigo 130D. Este artigo anunciou a reforma dos Fundos Estruturais, com o objectivo de aumentar a sua eficácia, no sentido de melhorar a coesão económica e social na Comunidade. Do mesmo modo, em Fevereiro de 1992, o Tratado de Maastricht integrou esta necessidade imperativa de coesão, e o Artigo 130B estipula que o Conselho deve definir medidas necessárias para garantir a utilização eficaz dos Fundos.

Em conformidade com estes requisitos, os regulamentos relacionados com os Fundos Estruturais (os de 1988, no caso da primeira geração de Fundos Estruturais, e os de 1993, no caso da geração actual) passaram a integrar artigos específicos sobre a avaliação, sobretudo o Artigo 6.º do regulamento-quadro, e o Artigo 26.º do regulamento de coordenação.

O Artigo 6.º, Parágrafo 2, dos regulamentos de 1988, estipulou que "para avaliar a eficácia das intervenções estruturais, os impactos das acções Comunitárias devem ser avaliados ex-ante e ex-post. Por isso, as candidaturas a apoio pelas autoridades competentes devem incluir as informações necessárias, de modo a poderem ser avaliadas pela Comissão".

Estes primeiros regulamentos, adoptados pelo Conselho em 19 de Dezembro de 1988, diziam respeito a planos de desenvolvimento regional, que tinham que ser entregues até 1 de Janeiro de 1989. Para muitas instituições de administração nacional e regional, a necessidade de estabelecer programas plurianuais constituiu, em si, uma espécie de revolução. Além disso, a maioria das autoridades competentes e os serviços da Comissão não possuíam as capacidades de avaliação necessárias, pelo menos no que diz respeito à avaliação dos Fundos Estruturais. Consequentemente, as avaliações executadas no período entre 1989 e 1993 não apresentavam uma qualidade muito elevada, apesar dos esforços de muitos actores, sobretudo na Comissão. Observou-se especificamente que o quadro definido por esta primeira versão dos regulamentos não permitiu criar condições adequadas à implementação da fase de avaliação ex-ante de forma regular.

Regulamentos aplicáveis entre 1994-99

Reconhecendo este fracasso parcial, por insistência da Comissão, mas também através do forte ímpeto de determinados Estados-membros (sobretudo, da Holanda e do Reino

Unido), os novos regulamentos de Julho de 1993 reforçaram consideravelmente os requisitos em termos de avaliação.

Este novo regulamento (Artigo 6.º do regulamento-quadro) determina os conceitos de avaliação prévia, monitorização e avaliação ex-post. Os Artigos que descrevem pormenorizadamente o conteúdo dos documentos de programação, entregues pelos Estados-membros, introduzem o conceito de "objectivos específicos quantificados, sempre que possível". Do mesmo modo, o Artigo 26.º dos regulamentos de coordenação apresenta uma cláusula prévia de que "será concedido apoio quando uma avaliação prévia mostrar a obtenção de vantagens socioeconómicas a médio prazo, face aos recursos mobilizados".

Na prática, esta exigência muito forte de uma avaliação ex-ante, uma variante Comunitária do conceito Britânico de *Value for Money* (ou seja, proveitos proporcionais ao investimento), não foi aplicada de forma muito precisa. Em contrapartida, a introdução da avaliação intercalar, que não existia antes, foi particularmente bem sucedida. O reforço da legislação europeia no domínio da avaliação foi claramente um forte incentivo para a disseminação da ideia de que a avaliação era inevitável para obter fundos comunitários.

Regulamentos aplicáveis entre 2000-2006

Com a terceira geração de Fundos Estruturais, a avaliação da sua eficácia será reforçada. As acções comunitárias serão, daqui em diante, objecto de avaliação ex-ante pelos próprios Estados-membros, de avaliação intermédia pelos gestores dos programas, e de avaliação ex-post, por iniciativa da Comissão, para avaliar os seus efeitos, face aos objectivos (os seus objectivos particulares e os seus objectivos de coesão), e para analisar o seu impacto sobre problemas estruturais específicos.

A eficácia dos Fundos deverá medir-se em três níveis: o efeito geral sobre os objectivos do Artigo 130.º A do Tratado (particularmente, o reforço da coesão económica e social), o efeito sobre as prioridades propostas nos Planos e fornecidas a cada QCA, e os efeitos sobre as prioridades específicas seleccionadas para as intervenções. Poderão ser lançadas avaliações complementares, sempre que aplicável de natureza temática, com o objectivo de identificar experiências que são transferíveis de um programa para outro.

Para aumentar a transparência, deverá colocar-se à disposição do público relatórios de avaliação, mais extensivamente do que no passado.

Regulamentos aplicáveis no período entre 2007-2013

Em termos de avaliação, a nova organização de Fundos Estruturais para o período entre 2007-2013 representa uma mudança de uma abordagem obrigatória para uma abordagem contínua, mais flexível.

A avaliação ex-ante de todos os programas operacionais ao abrigo do Objectivo de Convergência (antigo Objectivo 1) continua a ser uma obrigação, enquanto os Programas Operacionais ao abrigo do Objectivo de Competitividade Regional (antigo Objectivo 2) e do Objectivo de Emprego (antigo Objectivo 3) devem ser avaliados por programa, por Fundo ou por eixo prioritário. Os Programas Operacionais ao abrigo do Objectivo de Cooperação Territorial devem ser avaliados individualmente ou em grupos. A avaliação ex-ante dos novos Quadros Nacionais de Referência Estratégica (que substituem o Plano de Desenvolvimento Nacional e o QCA) não é obrigatória, mas altamente recomendada pela Comissão Europeia, sobretudo aos Estados-membros com Objectivo de Convergência.

Com base na experiência de avaliação intercalar do período anterior, que mostrou que os prazos impostos eram demasiado rígidos, os requisitos demasiado abrangentes, e os TdR adaptados de forma insuficiente às necessidades específicas, a nova avaliação contínua deverá ser mais orientada por necessidades específicas de cada país, do que por Regulamentos. Prevê-se que os Estados-membros executem avaliações contínuas, relacionadas com a monitorização dos Programas Operacionais, o que só é necessário em casos de afastamento significativo dos objectivos iniciais dos programas e das propostas de revisão dos programas.

Globalmente, os Programas de Fundos Estruturais para 2007-2013 devem ser mais estratégicos, centrados nos objectivos globais de longo prazo que, conseqüentemente, se traduzirão pelos Estados-membros em medidas e objectivos mais operacionais e de curto prazo. Do mesmo modo, a avaliação também deverá centrar-se em contextos específicos dos países, e na satisfação das necessidades de decisores políticos, bem como do público em geral.