

O recente reconhecimento da importância e a conseqüente valorização da avaliação de políticas e programas sociais esconde algumas armadilhas. A mais preocupante delas é pensar a avaliação sob os efeitos da sedução tecnológica, uma vez que, à sua eficiência e eficácia imediatas, não corresponde necessariamente um compromisso efetivo com a real melhoria da qualidade de vida dos cidadãos beneficiários de programas e serviços públicos.

Os ensaios deste livro são fruto da reflexão de pesquisadores experientes na formulação e execução de projetos de avaliação de políticas e programas sociais, reunidos no **Seminário de Avaliação de Políticas Públicas e Programas Sociais**, ocorrido em 2000. Os textos foram agrupados em três grandes conjuntos, em função do tipo de abordagem realizada: em primeiro lugar, quadros teóricos das tendências na pesquisa avaliativa; em segundo lugar, reflexões com base em pesquisas avaliativas realizadas; e, por fim, desenhos de sistemas de monitoramento e avaliação.

Este livro interessará a todos aqueles que, de algum modo, estão envolvidos com a elaboração, execução, monitoramento e avaliação de políticas públicas, no âmbito da ação pública ou privada, e que se confrontam com o constante desafio da gestão da área social. Com mais esta publicação, o **IEE/PUC-SP** reafirma seu compromisso ético com a real melhoria da qualidade de vida dos cidadãos brasileiros.

Tendências e perspectivas na avaliação de políticas e programas sociais



María Cecilia Roxo Nobre Barreira e
Maria do Carmo Brant de Carvalho (Orgs.)

**TENDÊNCIAS E PERSPECTIVAS NA AVALIAÇÃO DE POLÍTICAS E
PROGRAMAS SOCIAIS**

Sônia Miriam Draibe
Marta Tereza da Silva Arretche
Judy Baker
Jean Hébrard
Elba Siqueira de Sá Barretto / Regina Pahim Pinto
Maria do Carmo Brant de Carvalho / Denise Blanes
Mariangela Belfiore Wanderley / Denise Blanes
Nabil Bonduki

**Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)**

María Cecilia Roxo Nobre Barreira, Maria do Carmo Brant de Carvalho (Orgs.)

Tendências e perspectivas na avaliação de políticas e programas sociais /
Orgs. Maria Cecilia Roxo Nobre Barreira, Maria do Carmo Brant de Carvalho. -
São Paulo : IEE/PUC-SP, 2001.
224 p. ; 23 cm

1. Política social - Avaliação 2. Renda Mínima - Programa - Avaliação 3.
Organizações não-governamentais 4. Política e educação 5. Política habitacional
- Avaliação I. Barreira, Maria Cecilia Roxo Nobre. II. Carvalho, Maria do Carmo
Brant de. III. Instituto de Estudos Especiais - Pontifícia Universidade Católica de
São Paulo.

ISBN: 85-86894-18-4

361.61

CDD

Créditos de produção e edição

Projeto gráfico e editoração eletrônica de capa e miolo
Homem de Melo & Troia Design

Revisão de originais
Tereza Maria Lourenço Pereira

Digitação e correções
Carla Marques Li-Volsi

Tiragem
2000 exemplares

Impressão
Cromosete Gráfica e Editora Ltda

Este livro teve o apoio do Banco Itaú S.A.

Sumário

Apresentação	5
<i>Maria do Carmo Brant de Carvalho e Maria Cecilia Roxo Nobre Barreira</i>	
Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas	13
<i>Sônia Miriam Draibe</i>	
Uma contribuição para fazermos avaliações menos ingênuas	43
<i>Marta Tereza da Silva Arretche</i>	
Avaliando o impacto de projetos em desenvolvimento voltados à pobreza	57
<i>Judy Baker</i>	
Avaliação dos sistemas educativos e avaliação dos resultados escolares dos alunos	77
<i>Jean Hébrard</i>	
Avaliação de programas educacionais: indagações metodológicas e disseminação de resultados	101
<i>Elba Siqueira de Sá Barretto e Regina Pahim Pinto</i>	
Sistema de monitoramento e avaliação implementado por uma organização do terceiro setor: o caso da Ação Comunitária do Brasil	127
<i>Maria do Carmo Brant de Carvalho e Denise Blanes</i>	
Monitorando a implantação de um programa social	163
<i>Mariangela Belfiore Wanderley e Denise Blanes</i>	
Uma metodologia para avaliar programas de habitação	183
<i>Nabil Bonduki</i>	

**Maria do Carmo Brant de Carvalho e
Maria Cecília Roxo Nobre Barreira**

Apresentação

MARIA DO CARMO BRANT DE CARVALHO é doutora em Serviço Social pela Pontifícia Universidade Católica de São Paulo e pós-doutorada em Serviço Social Aplicado pela École des Hautes Études en Sciences Sociales de Paris. Atua como professora no Programa de Pós-Graduação em Serviço Social da Pontifícia Universidade Católica de São Paulo. Membro do Conselho Nacional da Comunidade Solidária da Casa Civil da Presidência da República do Brasil, é autora de vários trabalhos publicados.

MARIA CECÍLIA ROXO NOBRE BARREIRA é doutora em Serviço Social pela Pontifícia Universidade Católica de São Paulo, com aperfeiçoamento em Avaliação de Políticas e Programas Sociais no University Center for Social and Urban Research da University of Pittsburgh na Pensilvânia — Estados Unidos. Consultora de organizações sociais na elaboração de sistemas de monitoramento e avaliação, é autora de vários trabalhos publicados.

O Instituto de Estudos Especiais — IEE vem buscando, em sua história recente, executar, refletir e publicizar pesquisas, estudos e debates sobre a gestão de políticas públicas e, particularmente, sobre a questão de avaliação de políticas e programas sociais.

O IEE, neste empreendimento, envolve outros centros de pesquisas e intervenção no âmbito das políticas sociais. Assim, para estudar e disseminar o conhecimento acumulado no campo da avaliação, nosso tema em foco, realizou em 1996 um primeiro seminário, em conjunto com o NEPP — Núcleo de Estudos de Políticas Públicas —, da Universidade de Campinas, e a Fundap — Fundação do Desenvolvimento Administrativo do Governo do Estado de São Paulo.

O sucesso deste primeiro debate (que foi publicado e já está na sua segunda edição) entusiasmou-nos a realizar, em 2000, um segundo seminário sobre avaliação de políticas e programas sociais na Pontifícia Universidade Católica de São Paulo, em que se aliaram duas organizações de pesquisa e intervenção: o Cenpec — Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária e Cedac — Centro de Estudos e Documentação para Ação Comunitária.

Na primeira fase, ocorrida em junho de 1996, partiu-se da seguinte hipótese: “Não há ainda no Brasil uma sistematização das concepções, tendências e metodologias de avaliação para aferir processos, resultados e impactos no campo das políticas e programas sociais”. Em 2000, a ênfase manteve-se na busca de apreender e refletir sobre a pauta da pesquisa avaliativa e os conceitos metodológicos aí presentes. O foco foi no desvelamento das inovações ocorridas nestes últimos anos na área da avaliação.

A hipervalorização recente da avaliação de políticas e programas sociais pode nos levar a cair em armadilhas. A primeira e a mais preocupante delas é de pensar aquela avaliação centrada na sedução tecnológica, que acaba por expressar a ação na sua eficiência e eficácia imediatas, sem um compromisso intencional com a real transformação e melhoria da qualidade de vida dos cidadãos beneficiários de programas e serviços públicos.

Nesse patamar, a avaliação se espelha na reificação de indicadores e índices de eficiência e eficácia, despolitizando a própria ação. Em outras palavras, os indicadores apresentam-se como medidas reificadas de valor, instrumentos de controle dos agentes decisórios, quase como cartões de “apresentação de excelência técnica” das instituições, mas pouco penetráveis e suscetíveis de se transformarem em ferramentas de poder e controle pelos cidadãos comuns usuários de programas e serviços.

É preciso insistir no interesse primeiro deste debate:

- A avaliação não tem um valor em si, não substitui a política ou programa nos processos e resultados que move e persegue.
- A avaliação é oportunidade de reflexão crítica da ação e possibilidade de disputa e conquista dos resultados idealizados.
- A avaliação é imperativo ético, desconstrói e reconstrói a política ou programa na sua intencionalidade, resultados e efeitos na alteração da qualidade de vida — enquanto proteção, inclusão e cidadania — dos envolvidos na ação pública.

Neste livro, publicamos textos elaborados pelos pesquisadores, que os apresentaram no Seminário de Avaliação de Políticas e Programas

Sociais, realizado em 17 e 18 de agosto de 2000. A seqüência em que se apresentam nos leva a um aprofundamento do tema avaliação por três perspectivas: uma primeira, teórica/conceitual/metodológica das tendências na pesquisa avaliativa; uma segunda, que revela estudos e conclusões reflexivas com base em pesquisas avaliativas realizadas; e uma terceira, que apresenta o desenho de sistemas de monitoramento e avaliação nas fases de elaboração e implementação.

Sônia Miriam Draibe apresenta, em seu texto, conceitos instrumentais e uma metodologia de trabalho construída na sua trajetória de “quase duas décadas e algumas dezenas de pesquisas de avaliação de programas sociais” no Núcleo de Estudos de Políticas Públicas da Unicamp. Enfoca três temas centrais: as preferências e decisões do avaliador na avaliação de políticas públicas quanto aos objetivos, a natureza e o tipo de avaliação; os indicadores básicos das pesquisas de avaliação de processos e de impactos; e, finalmente, a busca de modelos analíticos para as avaliações de processos, com a proposta, pela autora, de uma particular alternativa de avaliação destes.

Marta Arretche, com sua experiência acumulada como avaliadora de políticas e programas sociais e como professora de Ciência Política da Unesp — campus de Araraquara, traz para o livro “uma contribuição para fazermos avaliações menos ingênuas”, centrando-se na avaliação da implementação que corresponde a uma fase da “vida” de um programa. Segundo a autora, para o avaliador “é fundamental ter em mente que a implementação modifica o desenho original das políticas, pois esta ocorre em um ambiente caracterizado por contínua mutação; mais do que isto, devem levar em conta que os implementadores é que fazem a política, e a fazem segundo suas próprias referências”. Daí a importância do tema abordado.

Tivemos a preocupação de buscar a interlocução com profissionais na área da avaliação em outros países. Do Banco Mundial, convidamos Judy Baker, economista responsável pela área da América Latina e Caribe e autora de um livro publicado recentemente, *Avaliando o impacto de projetos em desenvolvimento voltados à pobreza*. Em seu texto, Judy Baker apresenta um resumo dos temas principais de seu livro, incluindo:

conceitos e métodos, quantitativos e qualitativos, para avaliação de impacto; etapas do projeto e sua execução; lições aprendidas e exemplos de estudos de caso de avaliações de efeitos de “procedimentos ótimos”, inclui também a questão de custos e financiamentos.

Jean Hébrard, inspetor-geral da educação nacional do Ministério da Educação Nacional da França e professor associado ao Centro de Pesquisa sobre o Brasil Contemporâneo na École des Hautes Études en Sciences Sociales de Paris, analisa dois casos particularmente contrastantes que colocam, “de maneira evidente, a questão do estatuto das avaliações nos sistemas educativos”. O primeiro é o caso brasileiro do Programa Escola que Vale, do qual o autor é um avaliador externo. Tal Programa tem por objetivo a melhoria do desempenho de escolas públicas em quatro Estados do Brasil, envolvendo a Fundação Vale do Rio Doce e o Centro de Estudos e Documentação para Ação Comunitária — Cedac. O segundo exemplo refere-se à implementação, na França, de uma lei que orienta a avaliação permanente de políticas públicas de educação. O autor ressalta a importância da avaliação como fator que regula de modo permanente as ações das políticas públicas de educação, especialmente a avaliação da eficácia das escolas, envolvendo alunos e professores. A avaliação deve ser colocada a serviço da consolidação da real eficácia do investimento educativo.

Ainda na área da educação, Elba Siqueira de Sá Barretto, professora da Faculdade de Educação da USP e pesquisadora da Fundação Carlos Chagas, e Regina Pahim Pinto, também da Fundação Carlos Chagas, apresentam um estudo sobre avaliação de programas educacionais que se baseia em um levantamento dos artigos encontrados em dez das principais publicações científicas da área da educação no período de 1990 a 1998. Uma das constatações das autoras é de que a avaliação intencional de algumas ações governamentais, no campo da educação, tem se intensificado a partir dos anos 80, ainda prevalecendo, no entanto, “grandes vazios de informação e reflexão” sobre os programas educacionais brasileiros. As políticas educacionais do país são elaboradas sem que se tenha uma visão dos resultados acumulados de experiências anteriores e dos seus impactos.

Maria do Carmo Brant de Carvalho, professora do Programa de Pós-

Graduação em Serviço Social da Pontifícia Universidade Católica de São Paulo e coordenadora-geral do Centro de Estudos e Pesquisas em Educação Cultura e Ação Comunitária — Cenpec, e Denise Blanes, pesquisadora do Instituto de Estudos Especiais — IEE da PUC-SP, apresentam a construção de um “sistema de monitoramento e avaliação implementado por uma organização do terceiro setor: o caso da Ação Comunitária do Brasil — ACB”. As autoras destacam a “iniciativa inovadora e corajosa” da ACB na implantação de avaliação de resultados e impactos. “A ACB, ao se propor a avaliar suas ações, está introduzindo na agenda do Terceiro Setor um novo modo de gestão da ação privada com fins públicos.” Contribui para o que hoje se denomina *accountability*, ou seja, a necessária prestação de contas à sociedade, aos sujeitos envolvidos e ao beneficiário sobre as prioridades políticas e resultados atingidos.

No contexto da gestão social municipal, Mariângela Belfiore Wanderley, vice-coordenadora do Programa de Pós-Graduação em Serviço Social da PUC-SP e diretora do IEE, e Denise Blanes, pesquisadora do IEE da PUC-SP, montaram, em conjunto com a Prefeitura de Santo André, um sistema de avaliação do Programa de Renda Mínima desenvolvido por esse município paulista. O processo de elaboração do sistema enfrenta aspectos desafiantes por dificuldades na definição dos indicadores avaliativos, escassez de dados, especialmente aqueles anteriores à implementação do programa, bem como ausência de um sistema informatizado que dê agilidade às informações necessárias ao monitoramento do programa. O enfrentamento de tais dificuldades e a divulgação dessas pesquisas avaliativas geram um saber que se acumula na avaliação de programas sociais.

Nabil Bonduki, arquiteto, urbanista e professor do Departamento de Arquitetura e Planejamento da Escola de Engenharia de São Carlos — USP, é o coordenador de um trabalho articulado da academia com o poder público e com vários segmentos organizados da sociedade civil, o qual visa avaliar e aperfeiçoar os programas de habitação e saneamento. Em seu texto, apresenta a proposta preliminar de “uma metodologia de avaliação inédita no país, uma vez que os programas de habitação e saneamento, ao contrário dos programas sociais (saúde, edu-

cação e assistência), não têm sido objeto de avaliações globais, que ficam restritas, se tanto, a aspectos específicos ou a estudos de caso". Ressalta-se, mais uma vez, a preocupação dos formuladores de políticas e de programas sociais em aprimorar, valorizar e divulgar suas ações, pela implantação de um sistema de planejamento e avaliação competente.

Pela leitura dos textos, podemos verificar o quanto caminhamos com referência à sistematização de metodologias avaliativas e à construção de desenhos de sistemas de monitoramento e avaliação radicados em nossa realidade, nesses últimos quatro anos que separam a realização dos dois Seminários de Avaliação. O sentido de nosso empenho na publicação deste livro é, portanto, dar continuidade e incentivar a reflexão e o debate no campo da avaliação de políticas e programas sociais.

Sônia Miriam Draibe

Avaliação de
implementação:
esboço de uma
metodologia de
trabalho em
políticas públicas

INTRODUÇÃO

Registro, neste texto, alguns conceitos instrumentais e, sobretudo, uma metodologia de trabalho decantada em quase duas décadas e algumas dezenas de pesquisas de avaliação de programas sociais que venho conduzindo no Núcleo de Estudos de Políticas Públicas — NEPP, da Unicamp.¹ Uso deliberadamente a expressão *metodologia de trabalho* por dois motivos. Em primeiro lugar, porque, mais do que examinar e confrontar conceitos e alternativas metodológicas, trato aqui de indicar, com muita liberdade teórica e conceitual, um instrumental analítico simples, mas que, no entanto, tem mostrado razoável eficácia e consistência sobretudo quando utilizado como ferramenta auxiliar de pesquisas de alguma robustez. Em segundo lugar, porque deliberadamente considero no texto pelo menos três etapas ou planos do trabalho de pesquisa, quais sejam: o das preferências e escolhas iniciais do investigador; o dos indicadores básicos das pesquisas de avaliação de impactos e de processos; finalmente, uma particular alternativa de avaliação de processos por mim elaborada. Ora, esses planos remetem muito mais a um modo de trabalhar do que a uma reflexão sistemática sobre as alternativas conceituais e metodológicas neles implicadas.

SÔNIA MIRIAM DRAIBE é livre-docente em Ciência Política, professora-adjunta e pesquisadora do Instituto de Economia da Universidade de Campinas — Unicamp e do Núcleo de Estudos de Políticas Públicas — NEPP, da Unicamp; seus campos principais de investigação têm sido as Políticas Públicas e as Políticas Sociais. É autora de vários trabalhos publicados.

O objetivo do trabalho é também modesto. Visa auxiliar todos aqueles que, como eu, têm enfrentado as recorrentes dificuldades de formulação e elaboração dos projetos de avaliação de políticas e programas sociais. Como se sabe, no campo da análise de políticas públicas há pouco consenso acerca de conceitos básicos, o que obriga o analista a declarar, de início, suas preferências e escolhas, até mesmo para ser entendido. Ora, esse entendimento inicial é decisivo em qualquer pesquisa, seja para definir o terreno comum sobre o qual trabalhará a equipe, seja para garantir uma boa compreensão quando os resultados são divulgados.

Sintetizando leituras, elaborações próprias e experiências de longos anos, os conceitos, os modelos de análise e os parâmetros aqui definidos amparam-se na ampla literatura nacional e internacional relativa ao tema. Beneficiaram-se, sobretudo, da colaboração dos vários pesquisadores com quem tive a sorte de trabalhar ao longo desses anos.

AVALIAÇÕES DE POLÍTICAS PÚBLICAS: PREFERÊNCIAS E DECISÕES DO AVALIADOR

Estudos, pesquisas e atividades de avaliação de políticas apóiam-se em um conjunto de decisões, conscientes ou não, tomadas pelo avaliador no início e ao longo do seu trabalho. É o conjunto dessas opções e preferências que definirá os objetivos, a natureza e o tipo de avaliação, assim como o plano privilegiado de estudo do objeto eleito. Tais escolhas integram o que, com alguma liberdade, se pode denominar *estratégia de avaliação*, conceito que, entretanto, abrange também as decisões metodológicas encaminhadas pelo avaliador e apontadas no desenho da investigação.

OBJETIVOS, NATUREZA E TIPOS DE AVALIAÇÕES

O avaliador de políticas públicas enfrenta, de início, as decisões acerca dos objetivos e do tipo de pesquisa que realizará, bem como do recorte programático da intervenção que pretende avaliar.

Políticas, Programas, Projetos: Recortes Programáticos “Práticos”

A pesquisa de avaliação aqui tratada é aquela que tem como objeto uma dada intervenção na realidade social, uma *política pública*, entendida então como a que se desenvolve em esferas públicas da sociedade — e não no plano privado e interno das instituições ou organizações da sociedade. Políticas dessa natureza não se restringem, portanto, apenas às políticas estatais ou de governo, podendo abarcar, por exemplo, políticas de organizações privadas ou não-governamentais de quaisquer tipos, sempre e quando preservado o caráter público acima referido.

O senso comum e algumas velhas lições de administração, teoria de organização ou técnicas de planejamento distinguem, em graus decrescentes de abrangência, políticas, programas e projetos. Tomemos um exemplo da política pública governamental brasileira. Se nos referimos à política educacional como o conceito mais abrangente da intervenção na educação, o programa da merenda escolar constitui uma entre as tantas ações em que se desdobra aquela política, enquanto o projeto de descentralização do programa da merenda escolar constituiria uma unidade ainda menor de ação, entre as que integram o programa da merenda.

Entre as decisões prévias do avaliador, está a de identificar o recorte programático do seu objeto. Seria a política? Um ou mais programas, no interior de uma dada política? Um projeto? Teoricamente, pode-se avaliar tudo, desde a mais restrita até a mais abrangente das políticas, contanto que se disponha dos recursos intelectuais, materiais e metodológicos para fazê-lo. Mas é sempre bom ter em mente que avaliações rigorosas de intervenções complexas e abrangentes são difíceis. Em geral, os conceitos referidos neste trabalho mostram-se muito mais adequados a programas e projetos do que a complexas políticas públicas.

Objetivos da Avaliação

Avaliações de políticas e programas respondem a distintos objetivos. Desde logo, há objetivos de conhecimento: levantadas certas hipóteses, pesquisa-se um dado programa para conhecer a relação entre condições, meios, resultados e impactos da intervenção. É por isso, aliás, que

utilizamos aqui a expressão *pesquisa de avaliação*, e não simplesmente *avaliação*, preservando e enfatizando o suposto de que é um novo conhecimento que será produzido por meio de investigação que observa os cânones científicos mínimos.

As pesquisas de avaliações respondem, contudo, também a objetivos de verificação de eficácia, de eficiência e de *accountability* das ações. Com efeito, é legítimo, embora não tão freqüentemente quanto se almejava, que os responsáveis pelo programa — ou seus beneficiários, ou os demais grupos de interesse a ele relacionados — perguntem-se acerca da pertinência do seu custo ou da eficácia de suas ações. Nisso, podem estar movidos pelos mais distintos interesses: reduzir custos, usar mais adequadamente os recursos ou, enfim, prestar contas à sociedade do uso dos recursos. Objetivos mais imediatos podem ainda impor-se à avaliação, tais como os de detectar dificuldades e obstáculos e produzir recomendações, as quais possibilitem, por exemplo, corrigir os rumos do programa ou disseminar lições e aprendizagens.

São objetivos dessa natureza que fazem da pesquisa de avaliação de políticas públicas uma pesquisa interessada ou, como se diz no jargão da área, *policy oriented*, já que também busca detectar obstáculos e propor medidas de correção e alteração de programas, visando à melhoria da qualidade do seu processo de implementação e do desempenho da política.

Os dois grupos de objetivos não são excludentes. Ao contrário, é muito freqüente que apareçam de forma combinada, em cada pesquisa de avaliação, predominando um ou outro grupo, de acordo com as perspectivas e decisões do avaliador ou do proponente da avaliação. Mas esses objetivos também não se alcançam automaticamente, mesmo quando conscientemente escolhidos pelo avaliador. Por exemplo, objetivos de conhecimento dificilmente são alcançados se a investigação não conta com base teórica e hipóteses adequadas ou não se assente em metodologias capazes de garantir a confiabilidade dos resultados. Tampouco se pode pensar, por exemplo, que recomendações de políticas constituam um desdobramento automático de qualquer pesquisa de avaliação: exigem pelo menos o exame sistemático e o confronto das alternativas, não necessariamente contemplados na pesquisa inicial.

Tipo e Natureza das Avaliações

O tipo e a natureza de uma dada pesquisa de avaliação são definidos em um campo bastante complexo de alternativas, referentes, cada uma delas, a distintas dimensões, momentos e etapas do programa ou da política que se pretende avaliar. Sem nenhuma pretensão de abarcar todas as possibilidades, vamos nos referir aqui às distinções mais comuns.

A relação temporal entre o programa a ser avaliado e a pesquisa de avaliação propriamente dita permite fazer a clássica distinção entre dois tipos: avaliações *ex ante* e avaliações *ex post*.

Avaliações *ex ante* precedem o início do programa, ocorrendo em geral durante as fases de sua preparação e formulação; também referidas como avaliações-diagnóstico, são realizadas para apoiar decisões finais da formulação, atendendo a um ou dois de seus objetivos:

- produzir orientações, parâmetros e indicadores que se incorporem ao projeto, melhorando seu desenho e suas estratégias metodológicas e de implementação;
- fixar um ponto de partida que permita comparações futuras (linha de base ou tempo zero).

As avaliações *ex post* são feitas concomitantemente ou após a realização do programa (Cepal, 1998a, p.16), também com duplo objetivo:

- verificar os graus de eficiência e eficácia com que o programa está atendendo a seus objetivos;
- avaliar a efetividade do programa, ou seja, seus resultados, impactos e efeitos.²

Quanto à natureza, as avaliações distinguem-se entre avaliação *de resultados* e avaliação *de processo*.

É comum e freqüente que se queira saber, dos programas, se estes cumpriram seus objetivos. Na verdade, é mais do que isso: quer-se saber o quanto e com que qualidade seus objetivos foram cumpridos. A pesquisa de avaliação que poderá responder adequadamente a essa pergunta é a *avaliação de resultados*, no sentido amplo.

Já as *avaliações de processo* têm como foco o desenho, as características

organizacionais e de desenvolvimento dos programas. Seu objetivo é fundamentalmente detectar os fatores que, ao longo da implementação, facilitam ou impedem que um dado programa atinja seus resultados da melhor maneira possível.

Exploremos um pouco mais esta distinção entre avaliação de resultados e avaliação de processos.

AVALIAÇÕES DE RESULTADOS OU IMPACTOS

Há uma certa confusão provocada, na língua portuguesa, pelo uso pouco preciso da noção de resultado. Com efeito, tomados na sua acepção ampla, os resultados de um dado programa abrangem coisas diversas, que aqui denomino, na falta de melhor alternativa, como resultados *propriamente ditos*, *impactos* e *efeitos*. Estou certa da conveniência de distinguir tais noções, até porque são muito distintos os indicadores mais adequados para aferir cada um desses níveis de realização dos programas.

Para facilitar a compreensão, vejamos, no Quadro 1, alguns exemplos dessa diferenciação.

QUADRO 1 • DISCRIMINAÇÃO DOS TIPOS DE RESULTADOS DE PROGRAMAS: EXEMPLOS

Programas	Tipos de Resultados (sentido amplo)		
	Desempenho ou Resultados (sentido estrito)	Impactos	Efeitos
Vacinação antipólio	<ul style="list-style-type: none"> • % de cumprimento das metas • Nº vacinas aplicadas • Cobertura (% da população-alvo vacinada) 	Variação da taxa de incidência de poliomielite	<ul style="list-style-type: none"> • Melhora do nível de informação em saúde da população • Graus de satisfação dos pais
Leite para crianças desnutridas	<ul style="list-style-type: none"> • Quantidade de litros de leite distribuída / mês • Cobertura (% crianças que receberam o leite) 	Variação da taxa de desnutrição infantil	<ul style="list-style-type: none"> • Variação da taxa de mortalidade infantil • Melhora da capacidade institucional em tecnologia de distribuição de alimentos
Merenda escolar	<ul style="list-style-type: none"> • Cobertura (% de alunos que recebem a merenda) • Valores calóricos e protéicos ingeridos por aluno • Quantidade de refeições/dia distribuídas 	Melhora do perfil nutricional dos alunos	<ul style="list-style-type: none"> • Melhora do desempenho escolar dos alunos

Três programas, na primeira coluna, foram tomados como exemplo. Alguns de seus "resultados" estão exemplificados na segunda coluna. Muito sumariamente, pode-se dizer que os resultados (no sentido estrito) são os "produtos" do programa, previstos em suas metas e derivados do seu processo particular de "produção". Costuma-se também utilizar o termo *desempenho* para indicar resultados em sentido estrito (*outcomes*). No caso da merenda escolar, são exemplos de seus resultados a quantidade de refeições distribuídas, os valores calóricos e protéicos adicionados à dieta diária dos alunos ou, enfim, a quantidade de alunos que a receberam.

Já os impactos referem-se às alterações ou mudanças efetivas na realidade sobre a qual o programa intervém e por ele são provocadas. A terceira coluna do quadro nos diz que, de um programa de vacinação antipólio, por exemplo, o impacto esperado será a redução da incidência da poliomielite.

Finalmente, os efeitos referem-se a outros impactos do programa, esperados ou não, que afetam o meio social e institucional no qual se realizou. No caso da vacinação antipólio, o aumento do nível de informação da população em geral acerca da enfermidade pode ser um dos efeitos do programa; ou também o grau de satisfação dos pais com a organização do evento, ou, enfim, no plano institucional, a melhor capacitação dos agentes para executar ações de campanhas de vacinação.

Uma mesma pesquisa de avaliação pode avaliar, simultaneamente, os três tipos de resultados (em sentido amplo). Na maioria das vezes, infelizmente, limitam-se a identificar o primeiro grupo, o do desempenho do programa.

Podemos completar as definições anteriores com um rápido exame dos tipos de indicadores com que se pode avaliar resultados, impactos e efeitos. Aponto também — perdoem-me pelo didatismo — as perguntas que devem ser respondidas pelo avaliador, quando se dispõe a avaliar cada um daqueles planos de resultados.

RESULTADOS: COM QUE INDICADORES AVALIAR?

Cada tipo ou grupo de resultados (no sentido amplo) será melhor avaliado sempre que para tanto forem utilizados os indicadores adequados e os instrumentos corretos.

Indicadores de Resultados ou Desempenho

Desempenho ou resultados de um dado programa são aferidos por meio de indicadores que medem os graus, as quantidades e os níveis de qualidade com que as metas programáticas foram cumpridas. Ou seja, podem ser estimados ou medidos por indicadores que confrontam objetivos e metas, de um lado, e por realizações alcançadas pela atividade-fim, de outro.

A proporção das metas realizadas e de cumprimento do planejado ou os índices de cobertura da população-alvo são típicos indicadores de resultados. O nível de focalização do programa sobre os públicos-alvo pretendidos também constitui um bom indicador de resultados.

Perguntas básicas do avaliador — Esquemáticamente, as perguntas principais que o avaliador trata de responder, neste plano, podem ser assim exemplificadas: O programa realizou os objetivos pretendidos? Atingiu os grupos-alvo que pretendeu atingir? Cumpriu, e em que níveis de quantidade e de qualidade, as metas propostas?

Indicadores de Impacto e, a propósito, mais uma Classificação das Pesquisas de Avaliação

Os indicadores de impacto são muito mais difíceis de serem operados. Vejamos por quê.³

De acordo com a definição anterior, as avaliações de impacto tratam de medir as "mudanças" pretendidas ou, caso se queira, os resultados mais duradouros, verificados no grupo-objetivo. Pois bem, os indicadores de impactos devem ser capazes de medir os efeitos líquidos do programa — e somente do programa — sobre a população-alvo.

Voltemos aos exemplos. No caso do programa de saúde, deverão estimar a variação das taxas de incidência e a prevalência da poliomielite. No caso do programa do leite ou no da merenda escolar, deverão estimar as variações dos níveis de nutrição infantil ou do alunado.

Perguntas básicas do avaliador — No caso dos impactos, as perguntas do avaliador são principalmente as seguintes: O programa efetivamente produziu os resultados *strito senso* que pretendia alcançar? Afetou — e em que medida e sentido — a característica da realidade que queria transformar?

Há exigências metodológicas que tornam complexa a tarefa de avaliação de impacto. Uma delas é a definição da situação ou grupo de referência, com o qual se compararão os impactos do programa. Com o que comparar?

Uma possibilidade é confrontar o sujeito consigo mesmo, mediante uma comparação do tipo "antes — depois". Para tanto, é preciso dispor de indicadores da situação *ex ante*, ou seja, é necessário elaborar um diagnóstico da situação anterior ao início do programa, com a qual seus impactos serão comparados. O ponto de partida ou a *baseline*, aqui, é crucial, ou não haverá avaliação de impacto possível. Por exemplo, somente se poderá medir o impacto nutricional da merenda escolar caso se conheça o perfil nutricional do alunado antes e na ausência do programa. No caso da vacinação antipólio, o impacto de uma dada campanha somente poderá ser estimado se se dispuser de indicadores precisos de prevalência e incidência da poliomielite antes da vacinação.

Nem sempre as avaliações desse tipo são exequíveis. Quando os programas são contínuos, rotineiros e universais — como é o caso dos programas educacionais — é quase impossível determinar o ponto inicial. Além disso, avaliações diagnósticas são geralmente complexas e caras.

Quando não se processa nenhuma comparação, ou a comparação limita-se ao tipo "antes — depois" descrito, a avaliação realizada enquadra-se na família das avaliações não experimentais. Ora, a confiabilidade dos resultados depende dos métodos de seleção dos indivíduos a serem comparados e tende a decrescer na medida em que se reduz o caráter experimental da pesquisa. Por essa razão, mesmo as avaliações não experimentais não constituem as melhores alternativas de avaliação de impacto, seus resultados tendendo a se mostrar limitados.

Há, no entanto, alternativas cientificamente mais densas. É possível, por exemplo, estimar impactos de um dado programa por meio da comparação entre seus participantes e seus não-participantes. O segundo

grupo operará, então, como grupo de controle, em relação ao qual serão estimadas as diferenças da dimensão que se quer aferir.

Para que se cumpra o requisito básico das avaliações experimentais, os membros tanto do grupo de participantes (grupo focalizado) quanto do grupo de não-participantes (grupo de controle) devem ser aleatoriamente selecionados, ambos os grupos apresentando similaridades nos seus atributos, seja em relação aos fatores observáveis, seja aos não observáveis.

E se não for possível isolar, para comparação, um grupo de não-participantes?

Se o nosso campo fosse o das ciências exatas, poderíamos propor a alternativa clássica: aplicar o programa para um grupo e, simultaneamente, deixar de aplicá-lo para o grupo de controle. Ora, isso nem sempre é possível, nem eticamente recomendável. Nos nossos exemplos, envolveria a suspensão da merenda escolar para um grupo, ou a vacinação para outro! Em tais circunstâncias, a estratégia mais freqüente é a trabalhar com tipos ou desenhos quase-experimentais de avaliação, nos quais a comparação se faz com um grupo que opera como contrafactual ao programa, a sua seleção podendo se dar por distintas modalidades.⁴ Supõe-se, ainda assim, que a aleatoriedade da seleção, em cada um dos grupos, tenha sido observada.⁵

Não basta identificar bem o grupo de controle. A aferição dos efeitos líquidos de um programa implica, naturalmente, isolar e controlar todas as variáveis — da mesma natureza ou de natureza distinta — que possam interferir nos resultados tanto dos participantes quanto dos não-participantes. O suposto, aqui, é o de que a única diferença entre os dois grupos seja a o impacto do programa que está sob avaliação. Ora, o controle das variáveis intervenientes não é nada trivial. Vamos a alguns exemplos de vieses que podem alterar ou contaminar os resultados.

A determinação do tempo transcorrido entre o final do programa e o momento da avaliação é, quase sempre, um fator crucial, já que se pode supor que quanto maior for o período entre aqueles dois momentos, maiores serão as probabilidades de ocorrências de eventos que venham também afetar os resultados.

Outros vieses, se não controlados, podem afetar a comparação com o grupo de controle, mesmo quando esse tenha sido selecionado por critérios adequados. Por exemplo, no caso de um programa de leite para crianças desnutridas, pode ocorrer que o grupo de controle esteja sendo afetado, ou por um outro programa similar, ou mesmo por variações não controladas na dieta alimentar doméstica. Em certos programas — tomemos como exemplo um programa de capacitação para jovens — a própria decisão de dele participar pode introduzir viés nos resultados. Ou seja, na comparação entre os dois grupos, certas diferenças supostas como impactos — por exemplo, estar empregado ou ter melhores rendimentos — podem decorrer menos do treinamento obtido, antes do fator motivação, em geral associado a outras características dos sujeitos.

Difícil ou não, quando se quer produzir uma boa avaliação de impactos, o controle dessas variáveis é decisivo, de preferência por meio de métodos estatísticos que possam isolar e medir o impacto relativo de cada um dos fatores que influenciam os resultados — por exemplo, a análise multivariada, com técnicas de modelagem ou regressão logística.⁶

Indicadores de Efeitos

Como dissemos, os efeitos de um dado programa referem-se a determinadas alterações — esperadas ou não, outras que não as especificamente pretendidas — provocadas pelo programa e que, mais do que o próprio grupo-alvo focalizado, afetam as instituições ou organizações participantes, os agentes que implementam o programa ou o meio social em que este se realiza. Muitas vezes referidos como impactos indiretos, tendem a abranger os efeitos mais duradouros do programa sobre:

- os agentes implementadores, a comunidade local e grupos particulares de interessados na sua execução (*efeitos sociais*);
- as instituições governamentais e não-governamentais associadas à sua implementação (*efeitos institucionais*).

Perguntas básicas também se impõem aqui. A implementação do programa logrou produzir, nas instituições envolvidas, alguma aprendizagem

institucional? Produziu melhora de suas capacidades institucionais (state capabilities)? Estimulou ou criou condições para a continuidade, ampliação ou multiplicação da experiência (sustentabilidade)? As metodologias utilizadas puderam ou podem ser aproveitadas/transferidas para outras experiências? Foram constatadas melhoras nas capacidades profissionais dos agentes da implementação, como resultado das atividades desenvolvidas por eles próprios no interior do programa?

Passemos agora às avaliações de processo. Talvez estejam aqui minhas contribuições mais úteis.

AS AVALIAÇÕES DE PROCESSOS: EM BUSCA DE MODELOS ANALÍTICOS

Permitam-me iniciar essas notas com uma imagem sobre as avaliações de processos.

As políticas ou os programas têm vida. Nascem, crescem, transformam-se, reformam-se. Eventualmente estagnam, às vezes morrem. Percorrem, então, um ciclo vital, um processo de desenvolvimento, de maturação e, alguns deles, de envelhecimento ou decrepitude. É este ciclo (ou alguns de seus momentos) que constitui o objeto das avaliações de processos.

As políticas e os programas também têm, em contrapartida, carne e osso, melhor, têm corpo e alma. São decididas e elaboradas por pessoas, são dirigidas às pessoas ou ao seu *habitat*, são gerenciadas e implementadas por pessoas e, quando isso ocorre, são avaliadas também por pessoas. Ora, as pessoas ou os grupos de pessoas que animam as políticas, fazem-no segundo seus valores, seus interesses, suas opções, suas perspectivas, que não são consensuais, nem muito menos unânimes, como sabemos. Ao contrário, o campo onde florescem as políticas e programas pode ser pensado como um campo de força, de embates, de conflitos, que se sucedem e se "resolvem" ao longo do tempo.

Se for assim, as avaliações de processos serão ainda mais completas ao

tenderem a se apoiar também em conceitos e modelos de análise capazes de captar o sentido e a lógica de programas movidos por interesses, conflitos, eventuais negociações. Mesmo esquematicamente, vale a pena identificar os principais entre esses conceitos e dimensões de análise.

AS ESTRATÉGIAS DA IMPLEMENTAÇÃO

Os conteúdos políticos das políticas constituem tema clássico dos estudos de políticas públicas, base para a tão freqüentemente citada diferenciação *policy — politics — polity*. Não é bem desses conceitos gerais que me interessa tratar aqui, até para evitar repetições desnecessárias. Quero restringir-me especificamente às avaliações de processo e a algumas dimensões do que poderia chamar a *economia política das políticas públicas*.

Comumente, a dimensão da política das políticas públicas remete aos momentos de formulação e, sobretudo, às decisões iniciais, mas não somente àquelas referidas aos conteúdos materiais dos programas ou ao seu desenho. Seu processo de implementação também repousa em orientações e preferências, envolvendo cálculos estratégicos, escolhas e decisões por parte dos agentes que o conduzem e o implementam ou dele se beneficiam.

Em outras palavras, é possível identificar, em cada política ou programas, sua estratégia de implementação, constituída primordialmente por decisões acerca de características ou dimensões de processo, tais como a dimensão temporal; os atores estratégicos a serem mobilizados, nos diferentes estágios, para apoiar a implementação do programa; os subprocessos e estágios pelos quais se desenvolverá a implementação etc. Quando completa, a avaliação de processo deve envolver também a avaliação das estratégias que orientaram a implementação, aferindo em que medida tiveram ou não tiveram êxitos, garantiram ou dificultaram o sucesso do programa.

Sem pretender abarcar todas as dimensões que uma análise desse tipo pode abranger, refiro-me resumidamente às que julgo principais e que, a meu ver, devem estar na ordem de preocupações do avaliador da implementação, mesmo quando ele não vá trabalhar exaustivamente com a questão estratégica.

A Dimensão Temporal

Explicitamente ou não, a estratégia de implementação envolve uma escolha acerca da temporalidade das ações, escolha que, em princípio, corresponde a certos cálculos políticos quanto aos apoios e resistências que enfrentará o programa. Em geral, as alternativas aqui polarizam-se entre estratégias incrementais e estratégias de choque. Nas primeiras, a implementação se desdobra por um dado período, suficientemente longo para que, através dos vários estágios, logre-se construir, desde o interior mesmo do programa, as alianças ou coalizões de apoio capazes de vencer as resistências. Já as estratégias de choque, nas quais o grosso da implementação se faz de uma só vez e num momento único do tempo, supõem que ingredientes, tais como a surpresa, e medidas, tais como o insulamento burocrático dos que têm poder decisório, sejam capazes de vencer as resistências esperadas e inibir o surgimento de novas.

Atores Estratégicos e Matrizes de Conflito e Cooperação

Ao reconstituir as estratégias da implementação, o avaliador enfrenta questões do seguinte tipo: Quais são os atores (institucionais e ou individuais) que sustentam o programa? Entre todos, foram ou não alguns deles especialmente mobilizados para apoiar o programa, podendo mais à frente operar como núcleos de coalizões amplas de sustentação? Outros, também cruciais, teriam sido esquecidos ou desprezados, tal encaminhamento tendo significado um erro estratégico dos que têm poder decisório? A preocupação vale também para os atores que não aderem ou que resistem ao programa, é claro.

A identificação das várias matrizes político-institucionais de cooperação ou conflito, que permeiam e dinamizam o processo de implementação, é decisiva para a localização dos pontos de acordo e dos consensos mínimos, sendo indispensável às negociações e, sobretudo, à sustentabilidade dos programas.

Parcerias e Redes de Apoio

Se, nas considerações anteriores, a referência foi sobretudo aos atores in-

ternos, aqui quero referir-me aos diferentes grupos de interesse, aos parceiros e às redes de apoio que, mais amplamente, relacionam-se com o programa avaliado. São eles muitos e bastante diferenciados: sem nenhuma elegância, poder-se-ia considerá-los "produtores" e "consumidores" da política em questão. Podemos localizá-los desde os beneficiários mais diretos até atores, institucionais ou não, relacionados, por exemplo, com seu financiamento, passando mesmo pelos concorrentes, ou seja, pelos atores vinculados a programas similares e competitivos.⁴

Nas avaliações de processo, além da identificação desse campo diferenciado de interesses, é importante reconhecer que a própria avaliação nele se enraíza; sua credibilidade será tanto maior quanto mais se apóie em instrumentos e mecanismos de precisa identificação desses interesses e de suas distinções. Não será por isso que as avaliações participativas vêm ganhando importância?

A análise política das políticas públicas não é matéria fácil, nem sempre possível de ser realizada e, muito freqüentemente, desagrada a muitos. Ainda assim, a consideração das questões mais estratégicas da implementação é útil para a correta identificação dos fatores de processo que operam ou como facilitadores ou como obstáculos à consecução do programa.

ESTRUTURA E PROCESSO DA IMPLEMENTAÇÃO

Quero dedicar-me à questão da implementação. Entretanto, para que nos entendamos bem acerca de conceitos, faço de início a distinção, no ciclo de vida dos programas, de duas grandes etapas: a formulação e a implementação.

- **Formulação:** o processo de formulação de uma dada política ou programa é bastante complexo e tende a ser longo. Teoricamente, compreende a formação da agenda pública (a introdução do tema ou mais diretamente das demandas na agenda social e, posteriormente, pública); a produção e o confronto de alternativas por parte dos di-

ferentes grupos de atores; os processos de filtragens e de apropriação/domesticação da policy por parte dos agentes, segundo o legado ou as tradições e culturas organizacionais; a formulação e decisão; as definições de estratégias de implementação.⁹

- **Implementação:** inclui tanto as atividades-meio, que viabilizam o desenvolvimento do programa, quanto a atividade-fim, ou a execução propriamente dita, antes que se torne rotineiro.

Veremos logo mais quais são esse subprocessos nos quais se desdobra o processo mais geral da implementação.

O que se avalia ao se avaliar a implementação?

Uma resposta simples poderia ser a seguinte:

As avaliações de processo, de natureza qualitativa, buscam identificar os fatores facilitadores e os obstáculos que operam ao longo da implementação e que condicionam, positiva ou negativamente, o cumprimento das metas e objetivos. Tais fatores podem ser entendidos como condições institucionais e sociais dos resultados.

A questão básica que norteia investigações dessa natureza é a de detectar os condicionamentos, no plano dos processos, dos êxitos ou dos fracassos dos programas. É também a de saber se outras alternativas de processos garantiriam melhores resultados, ou, inversamente, se os mesmos resultados poderiam ser alcançados com alternativas menos caras ou mais rápidas de processos ou sistemas.

Como avaliar processos de implementação?

Registro, na seqüência, um esboço de metodologia de análise de processo, que venho elaborando ao longo da minha experiência em avaliações de programas sociais. Embora não acabada, tem se mostrado útil e frutífera, podendo certamente interessar aos colegas pesquisadores de políticas públicas.

CICLO DE VIDA DOS PROGRAMAS: UMA METODOLOGIA DE AVALIAÇÃO DE PROCESSOS

Já distinguimos, anteriormente, as etapas de formulação e de implementação, ou seja, supomos que, teoricamente, quando se inicia o movimento da implementação, tenham já sido superadas as etapas de formulação (e decisão) e definidas as estratégias da própria implementação.

Para fins de pesquisa, é conveniente captar o que chamo a *anatomia do processo geral de implementação*, identificando seus principais subprocessos ou sistemas:

QUADRO 2 • SISTEMAS OU SUBPROCESSOS DA IMPLEMENTAÇÃO

Sistemas ou subprocessos da implementação:

- Sistema gerencial e decisório
- Processos de divulgação e informação
- Processos de seleção (de agentes implementadores e ou de beneficiários)
- Processos de capacitação (de agentes e ou beneficiários)
- Sistemas logísticos e operacionais (atividade-fim)
 - financiamento e gasto
 - provisão de recursos materiais
- Processos de monitoramento e avaliação internos

Vejamos, sumariamente, as preocupações de análise de cada sistema ou subprocesso.

Sistema Gerencial e Decisório

Seja qual for a inscrição institucional de um dado programa, seu desenvolvimento apóia-se numa dada estrutura organizacional — cujas características obviamente interessam ao avaliador — e está submetido a um sistema gerencial e decisório específico, que "conduz" ou dirige a implementação.

A estrutura hierárquica desse sistema, seus graus de centralização ou descentralização, de autonomia ou dependência das partes, de um lado, e as características da gestão do tempo e a capacidade de implementar decisões de que dispõem os gerentes, de outro lado todas essas são

características a serem avaliadas por meio de indicadores quantitativos e qualitativos adequados.

É igualmente importante, em muitos dos casos, compreender a natureza e os atributos dos gerentes, ou melhor, da autoridade que conduz o processo. Pertence ela aos próprios quadros do programa ou da política, ou lhes é externa? Ou seja, é constituída por funcionários do programa ou, contrariamente, constitui-se meramente pelo *staff*, de presença *ad hoc*? Goza de liderança e de legitimidade?

Processos de Divulgação e Informação

Seja qual for a dimensão do programa, é de se supor que, para implementar-se, apóie-se em mínima divulgação e circulação de informações, pelo menos entre aqueles mais diretamente afetados, os que o vão implementar ou os que dele vão se beneficiar. A suposição, portanto, é a de que parte do seu sucesso dependa da adequação dos meios de divulgação junto aos seus públicos internos e externos, bem como de informação qualificada (isto é, clara e suficiente) entre agentes implementadores e beneficiários.

A pergunta que, a respeito, se faz o avaliador é sobretudo a seguinte: as informações básicas sobre o programa — seus objetivos, seus modos de operação, seus componentes, seus prazos etc. — chegaram aos agentes executores e à população interessada nas quantidades, qualidades e com a antecipação necessária à realização das atividades?

Podem daí ser extraídos os indicadores que “medirão” os níveis de qualidade desse processo: clareza, abrangência e suficiência das informações; agilidade do fluxo etc.

Sistemas de Seleção

Os programas envolvem, em geral, algum tipo de seleção, seja de agentes que o implementarão, seja do público-alvo a que se dirige. Aqui também podem ser apontadas as questões básicas a serem enfrentadas pelo avaliador. Que sistemas e critérios de seleção foram utilizados no recrutamento dos agentes da implementação e dos grupos beneficiados?

A divulgação dos processos seletivos foi suficientemente ampla, atingindo a todos os potencialmente interessados? Competência e exame de mérito foram, em alguma medida, usados como critérios nos sistemas de seleção de instituições e agentes estratégicos do programa? Houve adequação entre os tipos de processos seletivos e os objetivos do programa?

Sistemas de Capacitação

Em qualquer programa, é imprescindível verificar a capacidade dos agentes para cumprir as tarefas que lhe cabem na implementação. Os sistemas nos quais se capacitaram podem ser externos ou internos ao programa. Quando internos e específicos, esta é indispensável. Além disso, pode ocorrer que também a capacitação dos beneficiários seja condição de sucesso do programa, logo, sua avaliação também se torna decisiva.

O que interessa saber, a respeito da capacitação, quando se quer avaliar a implementação?

Vejamos algumas perguntas do avaliador: Os prazos, os sistemas e os conteúdos da capacitação dos agentes implementadores foram, em qualidade e quantidade, adequados e pertinentes às atividades que deverão realizar? Os diferentes conteúdos, referidos às distintas atividades a serem enfrentadas pelos implementadores (gerenciais, didáticas, de supervisão ou monitoramento, de desenvolvimento social etc.), foram adequada e equilibradamente contemplados na pauta da capacitação? Terminada a capacitação, os implementadores se sentiram seguros para empreender as tarefas esperadas?

Sistemas Internos de Monitoramento e Avaliação

É possível identificar uma espécie de hierarquia de procedimentos avaliativos relacionados aos programas e políticas públicas. Vejamos.

Mesmo que não se faça de modo explícito, a implementação dos programas tende a ser monitorada ou supervisionada pelos seus gerentes, sobretudo nos seus estágios iniciais. Já avaliações internas tendem a ser mais raras, restringindo-se, em geral, a examinar os registros administrativos ou a coletar opiniões dos agentes implementadores. Mais raras

ainda são as auditorias externas e, finalmente, as avaliações externas, elaboradas com alguma isenção e pautadas por procedimentos científicos. É verdade que essas duas últimas formas vêm gradativamente se disseminando, com claros benefícios de transparência para as políticas.

Quais são os problemas de avaliação aqui presentes? Algumas perguntas ajudam a entendê-los. Seja qual for a forma avaliativa utilizada, trata-se de saber: Este procedimento é regular? Seus conteúdos e métodos contribuíram para a melhora da implementação e para a maior efetividade do programa? Ocorreram correções de processos e procedimentos, como resultado das evidências coletadas no monitoramento? Seus resultados foram adequadamente sistematizados e socializados entre instituições e técnicos participantes?

Sistemas Logísticos e Operacionais (Atividade-Fim)

Até agora, tratamos das atividades-meio, mas o modo de implementar e de operacionalizar a atividade-fim, antes que o programa se torne rotineiro, é também de fundamental importância para que se avalie a implementação.

Nessa dimensão, os parâmetros de suficiência — de recursos e de tempo — em geral são decisivos. Independentemente de como são financiados os programas, trata-se de saber se os recursos financeiros são suficientes para os objetivos propostos ou se podem ser maximizados, em benefício dos próprios objetivos e metas. E, tão importante quanto, precisa-se saber se chegam onde devem chegar, e se nos prazos adequados.

Já no que tange aos recursos materiais, vale lembrar que, em muitos programas, a base material específica na qual se apóiam — equipamentos coletivos, equipamentos de comunicação, serviços de transportes etc. — é crucial para o seu êxito. Medidas de suficiência e de qualidade parecem ser as mais adequadas.

Examinamos até aqui os conceitos de resultados e processos, assim como as questões de avaliação que geralmente envolvem. É possível avaliar simultaneamente esses dois planos dos programas? Vejamos então os conceitos que têm tal pretensão.

INDICADORES AGREGADOS: AVALIAÇÕES DE EFICÁCIA, EFICIÊNCIA E EFETIVIDADE DOS PROGRAMAS

Processos e resultados podem, devem e freqüentemente são apreendidos em conjunto nos estudos de avaliação, por meio de conceitos que tratam exatamente de relacioná-los. É o caso dos conceitos de eficácia, eficiência e efetividade, abusivamente utilizados na pesquisa de avaliação, embora seu entendimento longe esteja de ser consensual.

Vou propor aqui algumas definições, mesmo correndo o risco de "remar contra a corrente", como me alertou um aluno há pouco tempo.

EFICÁCIA

No plano mais geral, a eficácia de um dado programa diz respeito à relação entre características e "qualidades" dos processos e sistemas de sua implementação, de um lado, e os resultados a que chega, de outro.

Assim, será mais eficaz — atingirá seus resultados em menor tempo, menor custo e com mais alta qualidade — aquele programa que se apoiar em processos e sistemas adequados de implementação. Ou seja, quando se examinam processos e sistemas da implementação, consideram-se eficazes aqueles nos quais predominam os fatores facilitadores de resultados em quantidades e qualidades máximas e em períodos mais curtos de execução. Ao contrário, perde em graus de eficácia o programa, cujos processos e sistemas de implementação (um ou mais) apresentam eles próprios ineficácias, tendendo a operar como obstáculos ao atingimento dos resultados.¹⁰

A avaliação de eficácia é, por natureza, uma avaliação de processo. A pergunta básica que estará respondendo o avaliador, ao medir a eficácia, pode ser assim formulada: Quais são ou foram, na implementação, os fatores de ordem material e, sobretudo, institucional que operaram como condicionantes positivos ou negativos do desempenho dos programas?

EFICIÊNCIA E EFETIVIDADE ¹¹

A eficiência diz respeito às qualidades de um programa, examinadas sob os parâmetros técnicos, de tempo e de custos.

Em uma definição relativamente restrita, pode ser entendida a relação entre produtos (bens e serviços) e custos dos recursos (insumos e atividades), confrontada com uma norma ou parâmetro (Cepal, 1998a, p.18). Medidas de custo/benefício seriam, nesta acepção, os melhores indicadores da eficiência de um dado programa.

Adicionado o parâmetro tempo — maior produção/menores custos/menor tempo —, temos aí um particular ângulo da eficiência, o da produtividade. Quando se consideram as alternativas tecnológicas do processo, pode-se chegar a uma outra medição de efetividade (também denominada eficiência técnica).

As perguntas típicas para a avaliação de eficiência são do seguinte tipo:

- As atividades ou processos se organizam, tecnicamente, da maneira mais apropriada?
- Esta condição opera com os menores custos para a produção máxima dos resultados esperados?
- Esta condição opera no menor tempo possível?

Já o conceito de efetividade refere-se à relação entre objetivos e metas, de um lado, e impactos e efeitos, de outro. Ou seja, a efetividade de um programa se mede pelas quantidades e níveis de qualidade com que realiza seus objetivos específicos, isto é, as alterações que se pretende provocar na realidade sobre a qual o programa incide. Ora, como vimos antes, estas alterações são de dois tipos: impactos e efeitos. Logo, a efetividade significa as quantidades e/ou os níveis de qualidade com que o programa atinge os impactos esperados e promove os efeitos, previsíveis ou não.

Os efeitos, como já assinalamos, podem ser agrupados em duas dimensões: social e institucional. Se tal diferenciação for útil, como penso ser,

devemos distinguir a avaliação da efetividade social e da efetividade institucional, uma vez que são também muito distintos os indicadores que se referem a uma e outra dimensão.

De fato, por efetividade social, especificamente, entende-se a capacidade do programa em afetar o capital social do meio em que se realiza, em especial a rede de articulações e parcerias específicas que facilitam sua execução. O conceito remete também aos níveis de adesão e satisfação dos agentes implementadores e da população-alvo.

Entre os principais indicadores a serem mobilizados para a avaliação da efetividade social de um dado programa, sugiro:

- os indicadores de capital social, que estimam os graus de confiança, as redes sociais, os graus de apoios, enfim, as energias associativas e organizacionais, presentes no meio social e que se relacionam com o programa;
- os níveis de satisfação dos agentes implementadores, dos usuários e de grupos de interesse.

O conceito de efetividade institucional, por sua vez, remete às afetações provocadas pelo programa sobre as organizações e instituições responsáveis ou envolvidas na implementação. Entre os conceitos e dimensões que podem captar tais efeitos, registro principalmente aqueles mais caros à análise neo-institucionalista, tais como os de capacidade institucional, aprendizagem institucional ou os referentes a hábitos ou comportamentos culturais das organizações,¹² todos eles remetendo, afinal, às condições institucionais de sustentação e multiplicação dos programas.

Como espero ter mostrado, as medidas de eficácia, eficiência e efetividade tratam de relacionar, cada uma a seu modo e nos respectivos planos, as avaliações de processo e as de resultados.

O quadro seguinte expõe, de forma resumida, esses planos e conceitos de avaliação de resultados e processos, exemplificando, com alguns indicadores, as possibilidades de avaliação a que nos estamos referindo.

Dimensões dos programas	Indicadores	
	Indicadores de Desempenho	
Dimensões	Tipo	Exemplos de indicadores
Metas	Físicas	<p>Metas realizadas x metas previstas</p> <ul style="list-style-type: none"> • Produção: proporções do cumprimento das metas; regularidade da prestação dos serviços • Público-alvo: cobertura; focalização
Metas	Financeiras	Metas realizadas x metas previstas
	Indicadores de Eficiência	
Recursos/Prazos	Custo	Relação custo/benefício ou custo/efetividade
	Indicadores de Eficácia	
	Sistema gerencial e decisório	Competência dos gerentes; capacidade de implementar decisões; graus centralização/descentralização
	Processos de divulgação e informação	Diversificação dos canais; suficiência e qualidade das mensagens; % público atingido; adequação de prazos; agilidade do fluxo
	Processos de seleção (de agentes implementadores e ou de beneficiários)	Publicização; competitividade do processo; qualidades dos sistemas de aferição de mérito; adequação do grupo selecionado aos objetivos do programa
Subprocessos ou sistemas da implementação	Processos de capacitação (de agentes e ou beneficiários)	<ul style="list-style-type: none"> • Competência dos monitores • Duração e qualidade dos cursos: conteúdos; didáticas; avaliações dos beneficiados
	Sistemas logísticos e operacionais (atividade-fim) financiamento e gasto; provisão de recursos materiais	Suficiência dos recursos; prazos e fluxos; qualidade da infra-estrutura e material de apoio
	Processos de monitoramento e avaliação internos	<p>Regularidade; abrangência; agilidade na identificação de desvios e incorreções; capacidade/agilidade em recomendar correções (feedback)</p> <p>Regularidade; abrangência; graus de participação e comprometimento dos atores e stakeholders; efetividade (em extrair lições, propor e recomendar melhorias, promover aprendizagem institucional)</p>
		<p>Monitoramento</p> <p>Avaliação interna</p>
	Indicadores de Efetividade	
Impactos	Variações/mudanças na realidade de intervenção	Taxas de crescimento do fenômeno / intensidade da mudança
Efeitos	Efetividade Social (Capital social; opinião e satisfação dos atores)	<ul style="list-style-type: none"> • Envolvimento de organizações da sociedade civil • Avaliação dos atores sobre programas • Graus de satisfação dos beneficiários; graus de adesão / satisfação dos agentes e beneficiários • Indução/reforço a associações/redes/parcerias
	Efetividade Institucional (Capacidades institucionais; sustentabilidade e reprodutividade)	<ul style="list-style-type: none"> • Graus de indução a mudanças / inovações / autonomia • Níveis de aprendizado institucional • Rotinização / institucionalização dos processos • Capacidade de transferência de metodologia

OBSERVAÇÕES FINAIS

Há duas fortes convicções, neste esboço metodológico, que espero ter revelado ao leitor.

A primeira é a de que a mera avaliação de desempenho é muito limitada, não faz justiça e mal avalia os programas, embora constitua passo indispensável de qualquer avaliação. Isoladamente, esse plano restrito dos resultados, na melhor das hipóteses, poderá dizer muito sobre a realização do programa, mas seguramente pouco ou nada sobre sua efetividade ou a qualidade de seus processos. Em outras palavras, apenas registrar percentuais de cumprimentos de metas físicas ou financeiras não significa avaliar um programa. Uma efetiva avaliação ocorre exatamente quando se põem em relação tais resultados e os processos pelos quais foram produzidos (avaliação de processo e medidas de eficácia e eficiência); tais resultados e as mudanças que o programa provoca na realidade sobre a qual incide (avaliação de impactos e efeitos; medidas de efetividade).

A outra convicção é a de que, tanto quanto os resultados, também os processos devem e podem, na maioria das vezes, ser submetidos a avaliações. Somente assim — mediante suas próprias avaliações — é que se pode superar o nível meramente descritivo com que, em geral, os processos são tratados nas pesquisas de avaliação de programas e políticas públicas.

Quero com isso dizer que, além de possível, é recomendável que variáveis qualitativas, tais como aquelas que habitualmente descrevem os atributos de processos, recebam tratamento quantitativo. É esta uma das alternativas para se proceder à efetiva avaliação do nível de eficácia (ou de qualidade) de dado subprocesso ou sistema da implementação.

Vou tomar apenas um exemplo, para não estender-me em demasia. Suponha-se que, em um determinado programa de alfabetização de adultos, o processo seletivo principal seja o da seleção dos professores alfabetizadores. Quais são, em geral, os atributos de um processo desta natureza? E com que critérios avaliá-los? Sem dúvida, estarão presentes, nestas

definições, valores e escolhas do avaliador, mas, em geral, pode-se esperar de um bom processo seletivo que ele esteja apoiado em ampla divulgação (de modo que todos os potenciais candidatos estejam informados), que seja claro acerca de critérios e procedimentos de escolha (ou dos seus critérios de justiça), que apure mérito e que, portanto, apóie-se em processos competitivos, em instrumentos adequados etc.

Como discriminar, entretanto, as variações de qualidade entre as várias ocorrências de seleção de professores? Pode-se esperar que, em alguns casos, o processo tenha registrado alta qualidade, uma vez que tenha cumprido todas aquelas condições estabelecidas pelo avaliador. Mas pode ocorrer que não tenha cumprido todas, apenas algumas, operando então a um nível médio. Ou, finalmente, pode suceder que tenham sido tantas as "lacunas" que, no máximo, pode-se avaliar o processo seletivo como insuficiente ou ruim. Componho, no quadro seguinte, este exemplo:

QUADRO 4 • AVALIAÇÃO DE PROCESSOS: ATRIBUTOS E DIFERENCIAÇÃO DE NÍVEIS DE QUALIDADE DE PROCESSOS SELETIVOS

Atributos	Níveis de qualidade do processo de seleção de professores alfabetizadores		
	Alta (cumpre todas as condições abaixo)	Média (cumpre todas e apenas as condições abaixo)	Insuficiente (cumpre todas e apenas as condições abaixo)
Transparência/publicização	<ul style="list-style-type: none"> divulgação: todos os públicos canais: edital + 3 ou mais canais clareza: expõe objetivos e regras indica prazo de inscrição 	<ul style="list-style-type: none"> divulgação: apenas para professores da rede canais: edital + 2 canais expõe objetivos, mas não as regras não indica prazo de inscrição 	<ul style="list-style-type: none"> nenhuma divulgação apenas edital afixado não expõe objetivos nem regras não indica prazo de inscrição
Competitividade e critérios de justiça	<ul style="list-style-type: none"> apuração de mérito: melhor colocação em provas experiência: exame de currículo 	<ul style="list-style-type: none"> posição na carreira experiência: exame de currículo 	<ul style="list-style-type: none"> indicação / recomendação da autoridade
Instrumentos	<ul style="list-style-type: none"> prova objetiva prova didática entrevista com roteiro 	<ul style="list-style-type: none"> prova objetiva entrevista 	<ul style="list-style-type: none"> entrevista

Sistemas de pontuações ou ponderações podem ser utilizados para quantificar os atributos de processo e estimar, ao fim, o nível de qualidade alcançado. O que aqui se fez para o subprocesso de seleção pode e deve ser feito também para todos os outros subprocessos e sistemas da implementação.

Fecho este texto insistindo na tese de que, em avaliações de implementação, não basta tão somente descrever processos e sistemas. Mesmo quando se trata de variáveis qualitativas, é preciso — e é possível — ir além, caso se queira efetivamente avaliar o processo de implementação de um programa e relacioná-lo com os resultados, buscando explicar, então, também por fatores internos à própria política, os seus êxitos ou fracassos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAKER, J. *Evaluating the Poverty Impacts of Projects: A Handbook for Practitioners*. LCSRP/PRMPO. Washington: The World Bank, 1999.
- BARNOW, B., CAIN, G., GOLDBERGER, A. S. Issues in the analysis of selectivity bias. *Evaluation Studies*, v.5, n.1, p.42-59, May 1980.
- BID. *Evaluation: a Management Tool for Improving Project Performance*, Washington DC, Draft, 1995.
- CALMON, Kátia M. N. Avaliação de programas e a dinâmica da aprendizagem organizacional. *Planejamento e Políticas Públicas*, n.19, p.3-70, jun. 1999.
- CEPAL *Gestión de Programas Sociales en América Latina*. Série Políticas Sociales, v.1, n.25, Cepal, Santiago, 1988a.
- CEPAL *Metodología para el Análisis de la Gestión de Programas Sociales*. Série Políticas Sociales, v.2, n.25, Cepal, Santiago, 1988b.
- COLEMAN, James S. *Foundations of Social Theory*. Cambridge, MA: Harvard University, 1990.
- DRAIBE, Sônia M. Desenho do Projeto. In: *Projeto. Análise Qualitativa dos Programas Inovadores do Comunidade Solidária*. Campinas: NEPP/UNICAMP, 1998.
- DRAIBE, Sônia M. Análise qualitativa integrada. Relatório final n.2. In: *Projeto. Análise qualitativa dos programas inovadores do Comunidade Solidária*. Campinas: NEPP/UNICAMP, 1999.
- ERSKINE, Angus, MARGARET MAY, M., ALCOCK, Peter (Edit.) *The Students Companion to Social Policy*. Blackwell Publishers, 1998.
- EZEMENARI, K., RUDQVIST, A., SUBBARAO, K. *Impact Evaluation: a note on concepts and methods*. Revised Draft 5/99, The World Bank, Washington, 1999.
- FRIEDLANDER, D., GREENBERG, D., ROBINS, Ph. *Methodologies for Determinating the Effectiveness of Training Process*. In: BARNOW, Burt, KING, Christopher T. (Ed.) *Improving the ODDs*. Washington: The Urban Institute Press, 1999.
- HATRY, HARRY. *Performance Measurement*. Washington: The Urban Institute Press, 1999.
- IDB *Evaluation: A Management Tool for Improving Project Performance*. Inter-American Development Bank. Washington: Draft, 1995.
- PALFREY, C. et al. *Policy Evaluation in the Public Sector*. Inglaterra: Averbury, Ashgate Publishing, 1992.
- PIERSON, P. *Dismanteling the Welfare State? Reagan, Thatcher and Politics of Retrenchment*. Cambridge: Cambridge University Press, 1994.
- PIERSON, P. *The New Politics of the Welfare State*. *World Politics*, v.48, n.2, p.143-79, 1996.
- PROPPER, Carol. *Efficiency, Equity and Choice*. In: ALCOCK, P., ANGUS. *To Social Policy*. London: Blackwell, 1998.
- SIEMPRO — SISTEMA DE INFORMACIÓN, MONITOREO Y EVALUACIÓN DE PROGRAMAS SOCIALES. *Gestión Integral de Programas Sociales*. Orientaciones e Resultados. Buenos Ayres: BIRD/ UNESCO/ Secr. Desarrollo Social, 1999.
- SKOCPOL, T., IKENBERRY, J. *The Political Formation of the American Welfare State*. *Comparative Social Research*, n.6, p.87-148, 1983.
- SCKOCPOL, T., ORLOFF, Ann S. *Why not equal protection? Explaining the politics of Public Social Spending in Britain, 1900-191 and the United States, 1880-1920*. *American Sociological Review*, v.49, p.726-50, 1984.

NOTAS

1 Preparado para apoiar a exposição do Seminário Avaliação de Políticas e Programas Sociais, este texto resume partes de trabalho mais amplo, que trata de metodologia de avaliação de políticas públicas, em fase de elaboração.

2 Por tais características, as avaliações *ex ante* são também chamadas, em inglês, de "formative" evaluation, enquanto as do tipo *ex post* são denominadas "summative" evaluation (BID, 1995, p.4-6), expressões cuja tradução não faz nenhum sentido em português, não se recomendando em nenhum caso o seu uso.

3 Apóio-me, nesse item, sobretudo em Friedlander, Greenberg & Robins (1999); SIEMPRO (1999); Ezemenari, Rudqvist & Subbarao (1999).

4 Entre as alternativas mais conhecidas de definição do grupo de controle, estão as seguintes: candidatos inscritos, mas não selecionados para o programa; participantes que desistiram do programa; integrantes da mesma área geográfica e ou do mesmo grupo-alvo que, entretanto, não se candidataram ao programa; indivíduos similares, mas de fora da área geográfica coberta pelo programa; finalmente, não-participantes selecionados em bases nacionais ou regionais de dados (censos etc.).

5 Recentemente, em pesquisa que avaliou o impacto de um programa de capacitação de jovens, na ausência de fontes censuais de informação, com base nas quais pudessem ser sorteadas amostras, utilizei a técnica da amostra-sombra para compor o grupo de controle, que replicava todas as características da amostra de egressos, salvo o ter feito aquele exato curso de capacitação. Embora freqüente, esta não é a melhor alternativa; no caso citado, infelizmente, era a única possível.

6 Na pesquisa *Avaliação da descentralização dos programas do FNDE e da Merenda Escolar*, que coordenei no Núcleo de Estudos de Políticas Públicas — NEPP, Unicamp, obtive sucesso com a aplicação de regressão logística para determinar as chances (os níveis de risco) de determinados tipos de escola implementarem bem os programas, isolando variáveis tais como o tamanho da escola, o tipo municipal ou estadual e o nível de escolaridade do diretor. Ver Draibe, 1998. Na avaliação de impacto de programas norte-americanos de capacitação, Barnow, Cain & Goldberger (1980) tiveram êxito em isolar o efeito participação por meio de análise de regressão que incluiu, como variáveis de controle, todos os critérios usados na seleção dos candidatos ao programa.

7 Ao longo da implementação, tendem a variar os tipos predominantes de interação entre atores — interações cooperativas, conflitantes ou de *advocacy* — seja no interior da coalizão de apoio, seja na relação com outros grupos de atores.

8 Há vários esforços de classificação dos diferentes grupos interessados ou *stakeholders*, como prefere o jargão da moda. Os manuais do Banco Interamericano de Desenvolvimento distinguem, por exemplo, entre os *stakeholders*, os primários, os secundários e os patrocinadores (IDB, 1995).

9 É ampla e conhecida a bibliografia de teoria organizacional e análise de políticas públicas em que se apóia esta conceitualização. Por isso, dispensei-me de referi-la, com exceção de alguns títulos da corrente neo-institucionalista, mais à frente indicados.

10 Esta definição difere totalmente das que consideram apenas a relação metas-tempo, como a seguinte: "... [entende-se a eficácia] como o grau em que se alcançam as metas de produção de um projeto em um dado período de tempo, independentemente de seus custos" (CEPAL, 1998a, p.18).

11 Os conceitos aqui usados não se identificam plenamente com as definições da tradição da teoria econômica neoclássica, em especial do *welfare economics*. Nesta, a eficiência refere-se à relação entre insumos e produtos, mais amplamente entre custos e benefícios. Traduzida na linguagem da curva de Pareto, é eficiente a produção que, realizada num dado nível, se faz ao menor custo e com máximo benefício líquido. Já a efetividade, também chamada eficiência técnica, refere-se ao máximo da produção ou realização possível (dos objetivos), segundo a melhor alternativa tecnológica. O conceito de eficiência, então, é mais amplo e envolve mesmo o de efetividade. Se a relação custo-efetividade remete à realização dos objetivos específicos com menores custos, a eficiência significa a extensão em que a razão custo-benefício é maximizada (cf. Propper, 1996; Palfrey et al., 1992).

12 Na análise de processo, são também de grande utilidade os conceitos de legado e principalmente o de *path dependency*. Ver Skocpol & Orloff, 1984; Skocpol & Ikenberry, 1983; Pierson, 1994, 1996. Para um bom estudo da bibliografia sobre aprendizagem organizacional, ver Calmon, 1999.

Marta Tereza da Silva Arretche

Uma contribuição
para fazermos
avaliações menos
ingênuas

Os manuais de avaliação de políticas públicas nos ensinaram que a avaliação da eficácia, da eficiência ou da efetividade de programas públicos deve sistematicamente levar em consideração os objetivos e a estratégia de implementação definidas por seus próprios formuladores,¹ pois seria inteiramente fora de propósito que o avaliador tomasse em consideração objetivos e/ou metodologias externos àqueles estabelecidos pelos próprios programas. Avaliar segundo critérios alheios aos estabelecidos pelos formuladores implica necessariamente uma avaliação negativa, pois não é plausível esperar que um programa realize o que não estava em seu próprio horizonte de implementação.

Supor, no entanto, que um programa público possa ser implementado inteiramente de acordo com o desenho e os meios previstos por seus formuladores também implicará uma conclusão negativa acerca de seu desempenho, porque é praticamente impossível que isto ocorra. Em outras palavras, na gestão de programas públicos, é grande a distância entre os objetivos e o desenho de programas, tal como concebidos por seus formuladores originais, e a tradução dessas concepções em intervenções públicas, tal como elas atingem a gama diversa de seus beneficiários e provedores.

MARTA TEREZA DA SILVA ARRETCHE é doutora em Ciências Sociais, na área de Políticas Públicas pela Universidade de Campinas — Unicamp. Professora de Ciência Política na Universidade do Estado de São Paulo — Campus Araraquara, é autora de vários trabalhos publicados.

Essa distância não diz respeito a fatores de ordem moral ou ética, que possam ser atribuídos a interesses escusos de formuladores e implementadores. Na verdade, tal distância é uma contingência da implementação, que pode ser, em grande parte, explicada pelas decisões tomadas por uma cadeia de implementadores, no contexto econômico, político e institucional em que operam (Rossi & Freeman, 1993; Mitnick & Backoff, 1984; Mladenka, 1984).²

Para superar uma concepção ingênua da avaliação de políticas públicas, que conduziria necessariamente o avaliador a concluir pelo fracasso do programa sob análise, é prudente, sábio e necessário, então, admitir que a implementação *modifica* as políticas públicas.

POR QUE A IMPLEMENTAÇÃO MODIFICA AS POLÍTICAS?

Qualquer programa tem duas dimensões, que podem ser analiticamente desagregadas: objetivos e uma metodologia ou estratégia pela qual se pretende que estes objetivos sejam atingidos. A definição da estratégia é uma escolha entre outras alternativas possíveis, o que significa que um mesmo objetivo pode ser atingido por diversas modalidades de ação.

Não são todas as instâncias da máquina estatal que têm autoridade para criar programas; ao contrário, a possibilidade de formulá-los supõe uma dada inserção institucional em cargos que confirmam algum nível de centralização da autoridade. Os objetivos e estratégias de um programa expressam, portanto, as decisões e as preferências de uma autoridade central (em qualquer nível em que esta autoridade esteja inserida). Mesmo programas de muito reduzida complexidade e escala supõem um agente que tomou decisões e executores encarregados de implementá-las. Alternativamente, programas cujo desenho final supôs um complexo processo decisório serão necessariamente implementados por agentes que não participaram do processo de formulação.

A implementação, por sua vez, corresponde a uma outra fase da "vida" de um programa, na qual são desenvolvidas as atividades pelas quais se pretende que os objetivos, tidos como desejáveis, sejam alcançados. Embora possa ocorrer coincidência entre a figura dos formuladores e a figura dos implementadores, é muito raro que isto ocorra. Na prática, qualquer política pública é de fato feita pelos agentes encarregados da implementação. Esta constatação não é nova, pois Lipsky (1980), já argumentava que

(...) as decisões dos burocratas que desempenham atividades-fim, as rotinas que eles estabelecem, os expedientes que eles criam para contornar a incerteza e as pressões, eis o que vêm a ser de fato as políticas públicas. Eu argumento que a política pública não é bem entendida se observada pelo ângulo da legislatura ou do alto nível da administração, porque em diversos aspectos importantes ela é realmente feita nos escritórios lotados e nos encontros diários da burocracia. (...) trabalhadores, clientes e os cidadãos em geral "experimentam" a burocracia que desempenha atividades-fim por meio dos conflitos que estes encontram para tornar o desempenho de suas funções mais consistente com suas próprias preferências e compromissos. (Lipsky, 1980, p.xii)

Um programa é, então, o resultado de uma combinação complexa de decisões de diversos agentes. É claro que, nessa cadeia de interações, a concepção original, tal como apresentada na formulação, é, sem dúvida, muito importante, porque as decisões tomadas durante esta fase já excluam diversas alternativas possíveis. Mas esta é apenas uma das dimensões da vida de um programa. Na realidade, a implementação efetiva, tal como se traduz para os diversos beneficiários, é sempre realizada com base nas referências que os implementadores de fato adotam para desempenhar suas funções.³

A despeito dos esforços de regulamentação da atividade dos implementadores, estes têm razoável margem de autonomia para determinar a natureza, a quantidade e a qualidade dos bens e serviços a serem oferecidos. Neste sentido, eles têm, com efeito, a prerrogativa de *fazer a política*. É esta autonomia que, por sua vez, lhes permite atuar segundo seus próprios referenciais.

A IMPLEMENTAÇÃO COMO UM CAMPO DE INCERTEZAS

Em princípio, as vontades, os interesses, as lealdades e as concepções ideológicas dos diversos agentes envolvidos em um programa público dificilmente serão inteiramente coincidentes. Portanto, quanto mais complexo for um programa, maior será a variedade de interesses e concepções envolvidos em sua execução e, por conseqüência, mais fortes serão as tendências à não-convergência.

Imaginemos a implementação de um programa federal, de escala nacional, cujas regras de operação suponham a cooperação dos três níveis de governo, em um país federativo e multipartidário, como o Brasil, em que prefeitos e governadores têm autonomia política e podem estar ligados a partidos distintos. Políticas públicas compartilhadas por governos ligados a partidos que competem entre si tendem a produzir comportamento não-cooperativos, pois na base das relações de implementação haveria uma incongruência básica de objetivos, derivada da competição eleitoral. Este cenário — perfeitamente factível, dada a multiplicidade de programas descentralizados existentes — não pode ser encarado como um problema. É simplesmente um dado da realidade, que impõe um razoável grau de incerteza quanto à convergência de ações dos implementadores e, por extensão, da perfeita adequação entre formulação original e implementação efetiva.

Suponhamos, ainda, um programa que envolva uma extensa rede de organizações não-governamentais, cujos gestores aceitem participar do

programa por razões inteiramente pragmáticas, mas não partilhem dos mesmos objetivos e concepções dos formuladores e financiadores. É razoável supor que tais agências desempenhem suas funções de acordo com suas próprias referências e até mesmo de acordo com seus próprios objetivos. Este comportamento dos implementadores — perfeitamente possível, dada a variedade de programas públicos assentados sobre a noção de parcerias público x privado — é um cenário bastante provável e apenas corresponde à realidade da implementação de programas que contam com agentes cujos interesses e concepções são diferentes.

Assim, mais que uma fase (técnica e obediente) que sucede à formulação (política) de programas, a implementação pode ser encarada como um jogo em que uma autoridade central procura induzir agentes (implementadores) a colocarem em prática objetivos e estratégias que lhe são alheios (Mitnick & Backoff, 1984). Nessas circunstâncias, para obter adesão e "obediência" aos objetivos e ao desenho do programa, a autoridade central deve ser capaz de desenvolver uma bem-sucedida estratégia de incentivos. É por isto que

(...) qualquer modelo adequado de implementação deve ter relações de incentivo, implícitas ou explícitas, como um componente-chave. (Mitnick & Backoff, 1984, p. 68)

Como conseqüência, problemas na estrutura de incentivos podem explicar grande parte das dificuldades da autoridade central para obter sucesso na implementação de seus programas, mesmo que tenham sido cumpridos todos os demais requisitos necessários. Desse modo, a implementação é, de fato, uma cadeia de relações entre formuladores e implementadores, e entre implementadores situados em diferentes posições na máquina governamental. Isto implica que a maior proximidade entre as intenções do formulador e a ação dos implementadores dependerá do sucesso do primeiro em obter a adesão dos agentes implementadores aos objetivos e à metodologia de operação de um programa.⁴

Como vimos, o grau de sucesso de um programa depende diretamente do grau de sucesso na obtenção da ação cooperativa de outros atores — governamentais e/ou não-governamentais, a depender do desenho do programa —, cuja "obediência" não é um dado automático. Na verdade, a taxa de sucesso depende da cooperação não apenas dos agentes institucionalmente vinculados aos demais níveis de governo, mas de centenas de potenciais beneficiários e provedores, cujo comportamento precisa ser coordenado (Stein, 1984). E isto não é um problema a ser constatado pelo avaliador, mas um dado da realidade a ser incorporado à análise.

Cientes deste fato, os agentes formuladores e implementadores experientes e com poder decisório tendem a escolher estratégias de implementação pautadas muito mais por sua potencial aceitação do que por sua esperada eficiência ou efetividade. Portanto, a reação esperada dos agentes implementadores tende a ser um elemento decisivo na definição do conteúdo das políticas públicas.

Além disto, políticas não são formuladas em condições de irrestrita liberdade. Dado que a formulação de um programa — com seus objetivos e desenho — é em si mesma um processo de negociações e barganhas, seu desenho final não será necessariamente o mais adequado, mas, sim, aquele em torno do qual foi possível obter algum grau de acordo ao longo do processo decisório.

Tampouco os formuladores concebem estratégias de implementação dispondo de completa e perfeita informação acerca da realidade sobre a qual se pretende interferir. Na verdade, com quase absoluta frequência, a informação é imperfeita, incompleta e até mesmo baseada em expectativas quanto ao comportamento futuro das variáveis da realidade.

Por estas razões, desenhos de programas não são peças internamente coesas e ajustadas. Nestas circunstâncias, é possível que algumas das especificações do programa possam operar, no plano local, de modo contrário aos objetivos do programa. Isto significa que a metodologia

proposta produz, no plano local, implicações contrárias às originalmente previstas, dado o fato de que o desenho do programa é formulado em condições de razoável incerteza, com base nos efeitos esperados de uma dada estratégia de operação. Mais que isto: a diversidade de contextos de implementação pode fazer que uma mesma regulamentação produza impactos inteiramente diversos em diferentes unidades da federação.⁵

Em suma, ainda que uma adequada estrutura de incentivos seja vital para a implementação de programas — e ela é tanto mais necessária quanto maior a complexidade dos interesses envolvidos — fatores inerentes à formulação impactam o próprio desenho da estratégia de implementação e, por extensão, o conteúdo das políticas. Mas, uma vez adotada uma determinada estratégia, seus efeitos (reais) somente serão conhecidos ao ser implementada a política, dados os limites de informação e autonomia decisória de que dispõem as agências formuladoras.

INCORPORANDO A IMPLEMENTAÇÃO À AVALIAÇÃO

Na prática real da administração pública, a implementação de políticas ocorre em um ambiente caracterizado por contínua mutação, mutação esta que é inteiramente *alheia à vontade dos implementadores*.

Em primeiro lugar, os recursos, as prioridades e a influência relativa dos agentes encarregados da implementação freqüentemente mudam (Rossi & Freeman, 1993, p.28). Metas quantitativas são fixadas para cada período, de acordo com uma previsão dos recursos (futuramente) disponíveis; e, com absoluta freqüência, tendências não previstas no comportamento da economia influem no montante de recursos disponíveis, atingindo a possibilidade de realização das metas previstas. Objetivos conflitantes entre si também tendem a receber ênfases diferentes ao longo do ciclo de vida de um programa, expressando alterações na geometria de poder no interior da máquina pública. Ou, ainda, é possível que um fato novo, que pode ser desde um escândalo de natureza política até uma alteração (inesperada) em uma variável econômica extremamente relevante, implique modifica-

ções nas formas de desembolso de recursos em prazos bastante curtos de tempo. Mais ainda: é perfeitamente possível que, por diversas razões não previstas no momento da formulação, a importância relativa das agências encarregadas de funções diferentes seja alterada ao longo de um curto período de tempo.

Em segundo lugar, os interesses e a influência dos *stakeholders* de um dado programa podem mudar entre o momento da formulação e o da implementação, alterando de modo significativo (para mais ou para menos) sua disposição para colaborar no momento em que o programa é colocado em prática (Rossi & Freeman, 1993, p.28).

Uma adequada metodologia de avaliação não deve, portanto, concentrar-se em concluir pelo sucesso ou fracasso de um programa, pois, como espero haver demonstrado, independentemente da "vontade política", da ética ou do interesse dos formuladores e implementadores, a distância entre formulação e implementação é uma contingência da ação pública. Com efeito, uma adequada metodologia de avaliação deve investigar, em primeiro lugar, os diversos pontos de estrangulamento, alheios à vontade dos implementadores, que implicaram que as metas e os objetivos inicialmente previstos não pudessem ser alcançados.

Além disso, uma adequada metodologia de avaliação deveria concentrar-se no exame das razões pelas quais a distância entre os objetivos e a metodologia de um programa — tal como previstos por seus formuladores —, bem como sua implementação efetiva, ocorre por decisão dos próprios agentes implementadores.

Como enunciado anteriormente, todo programa tem objetivos e um desenho de implementação, que se expressam em uma legislação geral e em uma regulamentação específica. E a implementação de um programa supõe necessariamente a *relação* de uma autoridade central com agentes implementadores. Nesses termos, as *referências* (Mitnick & Backoff, 1984), isto é, os princípios que efetivamente pautam a ação dos agentes implementadores, deveriam ser a "legislação" e as "normas determinadas pela agência formuladora".

Uma primeira pergunta a ser feita deveria ser se os agentes implementadores *de fato conhecem o programa*. Caso, por razões da capacidade de informação da agência formuladora, estes não venham a conhecer efetivamente os objetivos e as regras de operação do programa, eles tenderão a eleger uma outra referência para a implementação — seus próprios objetivos ou os objetivos dos *stakeholders*.⁶

Uma segunda indagação possível diz respeito à *aceitação dos objetivos e regras do programa*. Pode ocorrer que, embora conhecendo tais objetivos, os implementadores discordem das prioridades estabelecidas e elejam suas próprias prioridades (burocráticas, de sua clientela, ou ainda de grupos de interesses) como sua referência de atuação.

Finalmente, uma terceira questão diz respeito às *condições institucionais para implementação de programas*, isto é, à situação na qual os implementadores, embora conhecendo e concordando com os objetivos do programa, não possam, por diversas razões — que podem variar desde impedimentos associados à capacidade fiscal e administrativa das agências implementadoras até compromissos e lealdades de ordem política —, realizar os objetivos previstos.⁷

Problemas não previstos no momento da formulação podem levar os implementadores a promover adaptações ao desenho original, tendo em vista garantir que os objetivos de um dado programa sejam realizados (Rossi & Freeman, 1993). Trata-se de uma medida de "adaptação" para viabilizar a realização dos objetivos do programa, que implica, por sua vez, que o mesmo não seja implementado de acordo com o desenho previsto.⁸

A TÍTULO DE CONCLUSÃO

A implementação de políticas supõe (pelo menos) dois tipos de agentes situados em pontos distintos: agentes encarregados da formulação dos termos de operação de um programa — os quais têm autoridade para definir

os objetivos e o desenho de um programa — e agentes encarregados de executá-lo, traduzindo suas concepções em medidas concretas de intervenção. Na prática, são estes últimos que fazem a política.

A incongruência de objetivos, interesses, lealdades e visões de mundo entre agências formuladoras e a gama diversa de implementadores implica que dificilmente um programa atinja plenamente seus objetivos e que seja implementado inteiramente de acordo com seu desenho. Esse fenômeno tende a ocorrer mesmo em políticas públicas cuja implementação não suponha a ação integrada de diversos níveis de governo, porque essa incongruência é derivada da complexidade mesma de um processo de implementação de políticas. Por esta razão, todo modelo de implementação deve contemplar a formulação de uma adequada estrutura de incentivos.

No Brasil, os programas de ação pública são crescentemente realizados mediante integração de vários níveis de governo e de agentes não-governamentais. Atualmente, raros são os programas que não são implementados no contexto de um sistema de relações intergovernamentais e de parcerias as mais diversas com a sociedade civil. Ter em mente as consequências institucionais desses fatores é de fundamental importância para aumentar a probabilidade de sucesso de programas de larga escala.

Para o avaliador, é igualmente fundamental, contudo, ter em mente que a implementação modifica o desenho original das políticas, pois esta ocorre em um ambiente caracterizado por contínua mutação. Mais que isto, devem levar em conta que os implementadores é que fazem a política, e a fazem segundo suas próprias referências. Nessas circunstâncias, cabe investigar a autonomia decisória dos implementadores, suas condições de trabalho e suas disposições em relação à política sob avaliação.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALONSO, Analucia Faggion, GUIMARÃES, Érico. Avaliação do Programa de Renda Mínima de Ribeirão Preto. Araraquara, 1998. (Trabalho final de estágio de conclusão de curso de Administração Pública). (Mimeogr.)
- ARRETCHE, Marta T. S. Tendências nos estudos sobre avaliação. In: RICO, Elizabeth Melo. Avaliação de Políticas Sociais: uma questão em debate. São Paulo: Cortez, 1998.
- ARRETCHE, Marta T. S. Políticas Sociais no Brasil: Descentralização em um Estado Federativo. *Revista Brasileira de Ciências Sociais*, v.14, n.40, p.111-41, 1999.
- EDWARDS III, George C. *Public Policy Implementation*. Greenwich: Jai Press, 1984.
- FIGUEIREDO, A., FIGUEIREDO, M. Avaliação política e avaliação de políticas. *Análise e Conjuntura*, v.1, n.3, set., -dez. 1986
- LAZIN, Frederick A., ARONI, Samuel. Games Federal Officials Play: The Case of Housing Assistance Plans. In: EDWARDS III, George C. *Public Policy Implementation*. Greenwich: Jai Press, 1984. p.157-68.
- LIPSKY, Michael. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. Nova York: Russel Sage Foundation, 1980.
- MITNICK, Barry M., BACKOFF, Robert W. The Incentive Relation in Implementation. In: EDWARDS III, George C. *Public Policy Implementation*. Greenwich: Jai Press, 1984. p.59-122.
- MLADENKA, Kenneth R. Implementing Urban Services. In: EDWARDS III, George C. *Public Policy Implementation*. Greenwich: Jai Press, 1984. p.171-92.
- PETERSON, Paul, ROM, Mark C. *Welfare Magnets: A New Case for a National Standard*. Washington: Brookings Institution, 1990.
- PETERSON, Paul., *The Price of Federalism*. New York: The Twentieth Century Fund., 1995.
- RIKER, William H. *The Development of American Federalism*. Norwell: Kluwer Academic Publishers, 1987.
- ROSSI, Peter H., FREEMAN, Howard E. *Evaluation: A Systematic Approach*. Newbury Park: Sage Publications, 1993.
- SCHAPS, Ronald, RIKER, William. Disharmony in Federal Government. In: RIKER, William. *The Development of American Federalism*. Norwell: Kluwer Academic Publishers, 1987.
- SKOCPOL, Theda. Bringing the State Back In: Strategies of Analysis in Current Research. In: EVANS, Peter B., RUESCHEMEYER, Dietrich, SKOCPOL, Theda. *Bringing the State Back In*. Cambridge: Cambridge University, 1985.
- STEIN, Robert M. Policy Implementation in the Federal Aid System: The Case of Grant Policy. In: EDWARDS III, George C. *Public Policy Implementation*. Greenwich: Jai Press, p.125-55, 1984.
- TENDLER, Judith. *Bom governo nos trópicos*. Rio de Janeiro: Revan, Brasília: Enap, 1998.
- VALADEZ, Joseph, BAMBERGER, Michael. *Monitoring and Evaluating Social Programs in Developing Countries: A Handbook for Policymakers, Managers and Researchers*. Washington: World Bank, 1994.

NOTAS

- 1 Eu mesma defendi este argumento em trabalhos anteriores. Ver Arretche, 1998.
- 2 A implementação de um programa supõe sua inserção em "uma rede ou sistema complexo, no qual mesmo as ações de um agente [implementador] obediente podem produzir resultados imprevistos e talvez indesejados" (Mitnick & Backoff, 1984, p.59).
- 3 A burocracia encarregada da implementação pode, por exemplo, ter (mesmo que subjetivamente) preferência por uma parcela dos beneficiários potenciais que não corresponda à população-alvo do programa. Ou, ainda, pode ter compromissos com provedores que não tenham sido privilegiados pela estratégia de implementação prevista em seu desenho.
- 4 Tandler (1998) mostra que a alta dedicação dos "servidores públicos" foi uma das condições essenciais para a taxa de sucesso de quatro programas públicos implementados no estado do Ceará. A autora atribui este comportamento dos agentes implementadores a um conjunto de incentivos governamentais incorporados à estratégia de implementação.
- 5 Imaginemos o impacto fiscal da implementação do Fundo de Manutenção e Desenvolvimento do Ensino e Valorização do Magistério (Fundef). A diversidade de situações nacionais no que diz respeito à capacidade fiscal dos governos locais e à situação das matrículas nas redes municipal e estadual — o contexto de implementação de cada governo local — provocará simultaneamente efeitos benéficos e efeitos devastadores em diferentes governos locais. Tais efeitos, por sua vez, tenderão a produzir diferentes reações dos implementadores. Estes não poderiam ser, no entanto, inteiramente previstos pelos formuladores, dadas justamente a diversidade e a complexidade de situações no território nacional.
- 6 Imaginemos um programa de transferência de recursos às unidades escolares de todo o país para que estas ofereçam merendas escolares a seus alunos. O número de diretores escolares e de encarregados da gestão da merenda escolar é tão grande que os custos da difusão da informação podem ser muito elevados. Esforços para minimizar tais custos podem implicar um baixo ou insuficiente nível de conhecimento dos implementadores, criando as condições institucionais para que estes desempenhem suas funções de acordo com seus próprios referenciais.
- 7 Ao avaliar a implementação do Programa de Renda Mínima em Ribeirão Preto, Alonso & Guimarães (1998) concluíram que as assistentes sociais encarregadas da seleção dos beneficiários conheciam as regras legais de seleção, concordavam com as normas estabelecidas, mas selecionavam de fato segundo critérios que lhes pareciam subjetivamente mais sensatos.
- 8 Suponhamos um programa de renda negativa de âmbito municipal, que distribua mensalmente cheques de valor inversamente proporcional à renda familiar mensal dos beneficiários, os quais são selecionados segundo um teste-de-meios que exija extensa documentação. É muito possível que os agentes encarregados da seleção dos beneficiários flexibilizem os requisitos de comprovação de "necessidade", de modo a garantir que os mais necessitados venham a ser beneficiados.