

Maria Cecília Roxo Nobre Barreira &
Maria do Carmo Brant de Carvalho (orgs.)
Marta Tereza da Silva Arretche, Judy Baker,
Elba Siqueira de Sá Barretto, Denise Blanes,
Nabil Bonduki, Sônia Miriam Draibe,
Jean Hébrard, Regina Pahim Pinto &
Mariangela Belfiore Wanderley

Tendências e perspectivas na avaliação de políticas e programas sociais



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**



Maria Cecília 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)**

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

Tendências e perspectivas na avaliação de políticas e programas sociais /
Orgs. Maria Cecília 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 Cecília 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

CDD

361.61

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
<hr/> <i>Maria do Carmo Brant de Carvalho e Maria Cecília Roxo Nobre Barreira</i>	
Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas	13
<hr/> <i>Sônia Miriam Draibe</i>	
Uma contribuição para fazermos avaliações menos ingênuas	43
<hr/> <i>Marta Tereza da Silva Arretche</i>	
Avaliando o impacto de projetos em desenvolvimento voltados à pobreza	57
<hr/> <i>Judy Baker</i>	
Avaliação dos sistemas educativos e avaliação dos resultados escolares dos alunos	77
<hr/> <i>Jean Hébrard</i>	
Avaliação de programas educacionais: indagações metodológicas e disseminação de resultados	101
<hr/> <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
<hr/> <i>Maria do Carmo Brant de Carvalho e Denise Blanes</i>	
Monitorando a implantação de um programa social	163
<hr/> <i>Mariangela Belfiore Wanderley e Denise Blanes</i>	
Uma metodologia para avaliar programas de habitação	183
<hr/> <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.

○ 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 da avaliação de políticas e programas sociais.

○ 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 2ª 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 desafiadores 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

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.

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.

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.

Exploreemos 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

Tipos de Resultados (sentido amplo)			
Programas	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 sensu* 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,⁷ 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-

temos, 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.⁸

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 frequentemente, 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.

QUADRO 3 • QUADRO-RESUMO: DIMENSÕES E INDICADORES DE AVALIAÇÃO DE PROGRAMAS

Dimensões dos programas		Indicadores		
Indicadores de Desempenho				
Dimensões	Tipo	Exemplos de indicadores		
Metas	Físicas	Metas realizadas x metas previstas • 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				
Subprocessos ou sistemas da implementação	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		
	Processos de capacitação (de agentes e ou beneficiários)	• 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	Monitoramento	Regularidade; abrangência; agilidade na identificação de desvios e incorreções; capacidade/agilidade em recomendar correções (feedback)	
		Avaliação interna	Regularidade; abrangência; graus de participação e comprometimento dos atores e stakeholders; efetividade (em extrair lições, propor e recomendar melhorias, promover aprendizagem institucional)	
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)	• 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)	• 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*. LCSPR/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 D.C., 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. Metodologies 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.
- SKOCPOL, 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 Apoio-me, nesse item, sobretudo em Friedlander, Greenberg & Robins (1999); SIEMPRO (1999); Ezenenari, 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 desertaram 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 conceituação. Por isso, dispense-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, 1998; 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 Skocopol & Orloff, 1984; Skocopol & 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

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.

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.

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 excluíram 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 consequência, 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

A FORMULAÇÃO COMO UM CAMPO DE INCERTEZAS

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.

Judy Baker

Avaliando o
impacto de
projetos em
desenvolvimento
voltados à
pobreza*

JUDY BAKER é economista senior do Banco Mundial, responsável pelas regiões da América Latina e Caribe. Trabalha no Banco há dez anos, em assuntos relacionados à pobreza, distribuição de renda e programas sociais. Realizou estudos sobre a pobreza em diversos países dessa região, bem como o desenho, implementação e avaliação de programas sociais. Publicou, em junho de 2000, o livro *Evaluating the Impact of Development Projects on Poverty* (Avaliação de impacto de projetos em desenvolvimento voltados à pobreza).

Nos últimos anos, o Banco Mundial tem priorizado os projetos e políticas para a redução da pobreza, que se tornaram o foco principal de todo nosso trabalho. Apesar dos milhares de milhões que gastam os países financiadores, ainda não sabemos muito sobre o impacto de projetos específicos na redução da pobreza. Por exemplo, que impacto teve o projeto de capacitação profissional no emprego? Que impacto teve o projeto de alimentação escolar na assistência aos estudantes?

Fizemos um levantamento no Banco Mundial sobre as avaliações referentes a esses projetos, particularmente na região da América Latina e do Caribe, onde eu trabalho, e descobrimos que poucos deles foram avaliados. Após conversas com seus coordenadores, surgiram três razões que poderiam explicar a falta de maior número de avaliações:

- A falta de financiamento para os custos, que podem ser significativos;
- Em muitos países, o fato dos governos não quererem avaliações por razões políticas, especialmente nos casos em que interesses políticos determinam quais projetos e beneficiários devem ser selecionados;
- O fato de muitos coordenadores de projetos, no Banco e nos países onde trabalhamos, não saberem como fazer avaliações de impacto.

Para começar, ficou claro que o terceiro ponto seria o mais fácil de resolver, então decidimos preparar um guia de orientação sobre como avaliar o impacto de projetos, particularmente daqueles cuja meta, direta ou indireta, era a redução da pobreza. Podem ser incluídos nessa categoria, por exemplo, projetos em educação, saúde, microeconomia, infra-estrutura e outros temas.

Eis um resumo dos temas principais, que incluem:

- Conceitos e métodos quantitativos e qualitativos para avaliação de impacto;
- Etapas do projeto e sua execução;
- Lições que aprendemos e exemplos dos estudos de caso de avaliações válidas, além de alguma informação sobre custos.

Não discutirei os detalhes econométricos que estão incluídos no Capítulo 3 do livro de minha autoria, recentemente publicado.

CONCEITOS E MÉTODOS PARA AVALIAÇÃO DO IMPACTO

CONCEITOS

Uma avaliação completa inclui:

- Supervisão;
- Avaliação do processo;
- Avaliação do custo-benefício ou efetividade do custo;
- Avaliação do impacto.

Alguns projetos incluem todos estes componentes, mas a maioria somente inclui:

- Supervisão: verificar se o projeto foi executado como planejado;
- Avaliação do processo: revisar como o programa está operando e problemas na oferta dos serviços;
- Avaliação de custos e benefícios: calcular a relação do custo do programa em relação aos benefícios.

A avaliação de impacto é diferente. Esta determina se o programa teve os efeitos projetados em indivíduos, lares e instituições, assim como se tais efeitos podem ser atribuídos, de fato, à intervenção do projeto. Para isso é preciso controlar o elemento contrafactual ou a contrapartida: o que aconteceria se o projeto não existisse?

Para medir esse dado, utiliza-se um *grupo de controle ou comparação*, que se refere àqueles que não participaram do projeto ou não receberam os benefícios, mas que têm características similares aos participantes. O grupo que recebe os benefícios denomina-se *grupo de tratamento ou de intervenção*. Analisa-se, por exemplo, quando um recém-graduado de um programa de capacitação profissional consegue trabalho, se isso pode ser o resultado direto do programa ou se ele poderia conseguir o emprego de qualquer outro modo; se um empréstimo a um microempresário contribuiu de fato para o aumento de suas receitas ou se foram outros fatores do mercado que o ajudaram.

Este trabalho visa, especificamente, determinar o impacto dos projetos no que se refere aos pobres, mas os conceitos e métodos são os mesmos para a avaliação de impacto de qualquer outro tipo de pesquisa.

MÉTODOS

Para enfatizar o ponto mais importante, o coração de uma avaliação de impacto, há o estudo do contrafactual. Isso requer o uso de um grupo de controle ou comparação, formado por aqueles que não participam do programa.

As três formas principais são:

- Desenhos experimentais
- Desenhos quase-experimentais
- Métodos qualitativos

Desenhos Experimentais

São também chamados de *randomization*, ou amostra ao acaso, sendo considerados a melhor proposta. A intervenção, ou benefício, é aplicada

aleatoriamente entre os que são elegíveis; assim, o processo gera um grupo de controle. Este método é considerado o melhor, porque, em teoria, os grupos de controle criados ao acaso são a contrapartida perfeita, sem nenhuma distorção de seleção. No entanto, pode haver problemas de ética, como recusar benefícios de alimentação ou serviços médicos a alguns grupos que os solicitam.

Desenhos Quase-Experimentais

Pode-se usar este método quando não é possível constituir grupos de tratamento e grupos de controle por meio de um desenho experimental. Essas técnicas geram grupos de controle similares ao grupo de tratamento, mediante métodos econométricos como combinações, dupla diferença, comparações reflexivas etc. Em geral, os grupos são selecionados depois da intervenção, utilizando-se controles estatísticos. Tais métodos estão cada vez mais populares, pois podem ser baseados em pesquisas já existentes.

Entre os métodos estatísticos, *matching* ou métodos de sincronização são considerados uma boa alternativa para o projeto experimental. Os desenhos quase-experimentais têm problemas de distorção na seleção dos dados — o que requer um outro conjunto de detalhes econométricos —, que podem se referir a fatores não passíveis de controle, tais como aptidões, conexões familiares e outras.

Métodos Qualitativos

Esses métodos não focalizam o estudo comparativo, mas os processos de entendimento, comportamentos e condições, como são compreendidos por indivíduos e grupos que estão sendo investigados. Pode-se, por exemplo, ter um *insight* sobre como as famílias e as comunidades locais percebem um projeto, e como elas estão sendo afetadas pelas intervenções. Avaliações de impacto qualitativo usam métodos de avaliações rápidas ou avaliações participativas, que estão baseadas na participação dos beneficiários e gestores dos programas. Uma das principais desvantagens deste método é a falta da contrapartida e do tamanho das amostras, as quais, em geral, são muito menores.

Considerando-se os prós e os contras de cada método, a melhor maneira de se fazer uma avaliação de impacto é integrar técnicas quantitativas e qualitativas, incluindo a análise de custos e benefícios. Há muitas maneiras de incorporar tais técnicas durante cada fase de avaliação. Por exemplo, uma fase do projeto deve ser participativa, incluindo os beneficiários para se assegurar de que a avaliação contenha as perguntas corretas.

As qualidades das melhores práticas avaliativas incluem os seguintes pontos:

- Estimativa da contrapartida por projeto experimental ou projeto quase-experimental, usando-se métodos econométricos para gerar um grupo de controle;
- Controle das diferenças nos participantes, antes e depois do programa, estabelecendo seus efeitos. Há coleta de dados no início (*baseline*), como referência, e no seu seguimento (*follow-up*), com um prazo suficiente que permita observar os efeitos do programa;
- Grupos de comparação e de tratamento de tamanho suficiente, permitindo estabelecer inferências estatísticas significantes (*statistical significance*);
- Inclusão de análise do custo-benefício ou eficácia em função dos custos, para medir a eficiência do projeto;
- Inclusão de técnicas qualitativas, permitindo a triangulação dos resultados.

PASSOS-CHAVE NO DESENHO E NA IMPLEMENTAÇÃO DE AVALIAÇÕES DE IMPACTO

Existem muitas etapas na realização de um estudo de avaliação de impacto, o que pode custar muito em dinheiro e tempo. Tal fato enfatiza a necessidade de um estudo bem desenhado, de uma equipe bem qualificada e muito comprometida, com uma boa comunicação entre seus membros.

A planificação oportuna permitirá obter resultados no curso da implementação do projeto, nos tempos estabelecidos anteriormente, de modo

que os dados possam ser utilizados no aprimoramento de seu desenho, se necessário; assegurará, também, que os dados possam ser coletados durante os passos mais importantes do projeto.

A seqüência de cada um desses passos é fundamental, particularmente ao se assegurar a coleta dos dados necessários antes do início da implementação do projeto. Idealmente, a coleta dos dados deve ser feita no momento da identificação ou preparação do projeto, porque esses dados são mutáveis, não sendo possível identificá-los após o desencadeamento da ação projetada.

Ao tratar da identificação e preparação do projeto, mencionarei dez etapas.

1. DETERMINANDO A REALIZAÇÃO OU NÃO DE UMA AVALIAÇÃO

Realizar uma avaliação de impacto em determinados projetos talvez não seja apropriado, considerando-se sua complexidade e seu custo. Deve-se ter em conta qual método é o mais adequado, como a supervisão de resultados-chave ou o processo de avaliação. (Esses enfoques não devem ser considerados substitutos das avaliações de impacto, embora, com frequência, constituam-se como seus componentes decisivos.)

O dado mais importante para a decisão de fazer ou não uma avaliação talvez seja a necessidade de existir um forte apoio político e financeiro, que pode ser difícil, mas é requisito prévio para que se tenha continuidade. Os gestores e financiadores do programa devem estar seguros de que a avaliação é necessária para levantar dados que serão pertinentes para a tomada de decisões relativas à melhoria, expansão ou redução do programa avaliado. Deve-se também ter certeza da legitimidade do desenho da avaliação e, por conseguinte, dos resultados, particularmente quando esses não são tão positivos como seria esperado.

Financiar uma avaliação de impacto continua sendo um tema difícil para os administradores e financiadores dos programas. Em geral, é mais fácil mobilizar fundos se o projeto é inovador, pode ser reproduzido, supõe grandes aportes de recursos e compreende intervenções bem definidas.

2. DEFININDO COM CLAREZA OS OBJETIVOS DA AVALIAÇÃO

Nesta área, objetivos claros são essenciais para identificar as informações necessárias, definir os indicadores de resultados e efeitos, assim como para criar uma estratégia de avaliação sólida, proporcionando respostas às perguntas propostas. É importante haver consenso sobre os objetivos antes do começo da avaliação. O uso de um enfoque do marco lógico (*log*) é uma boa maneira de identificar as metas do projeto e as necessidades de informação, em cujas bases se criará a avaliação.

3. EXAMINANDO A DISPONIBILIDADE DE DADOS

Há muitos tipos de dados para estudos de avaliação de impacto. Estes podem incluir uma gama que varia de entrevistas transversais, ou de painel, a entrevistas abertas qualitativas. É ideal que esta informação exista em âmbito individual, a fim de assegurar que se possa avaliar o verdadeiro efeito.

Em muitos casos, a avaliação de efeitos aproveitará certo tipo de dados existentes ou se apoiará em alguma entrevista atual, o que pode poupar custos consideráveis. Alguns exemplos dos tipos de dados que podem ser usados incluem: entrevistas de renda e gastos nos lares; estudos dos níveis de vida (LSMS); entrevistas do mercado de trabalho; registros sobre assistência, repetência e rendimento escolar; dados referentes à saúde pública, sobre mortalidade infantil, incidência de diferentes enfermidades e outros.

Obviamente, seria mais barato usar dados existentes, pois há necessidade de uma certa criatividade quando se pensa a respeito disto. Muitas pesquisas também podem estar nas etapas de planejamento ou em curso. Se planejarmos uma entrevista que meça os indicadores requeridos, a avaliação poderá tomar uma amostra da população-alvo durante o curso da investigação geral.

4. DESENHANDO A AVALIAÇÃO

Estando os objetivos claros e os dados disponíveis, é possível começar a fase de proposta do estudo de avaliação de efeitos. A escolha de metodologias dependerá da questão formulada, da distribuição cronológica, das restrições orçamentárias e da capacidade de aplicação da avaliação. Devem-se analisar os prós e os contras dos distintos tipos, discutindo-se antes para determinar que metodologias são mais adequadas e que técnicas quantitativas e qualitativas podem se integrar para se complementarem entre si.

Menciono alguns pontos relevantes:

- *Clareza e especificidade*: são importantes na identificação das perguntas da avaliação.
- *Distribuição cronológica e pressupostos*: o mais importante é a possibilidade de se começar o desenho da avaliação, antes que o projeto seja implementado e seus resultados sejam necessários. É igualmente útil identificar, de início, quais pontos durante o ciclo do projeto necessitam da informação da avaliação. Ter os resultados corretos de forma oportuna pode ser decisivo para a definição da política a ser adotada, como, por exemplo, avaliar durante uma revisão do projeto, perto de um período de eleições ou na ocasião da tomada de decisões com respeito à continuidade do projeto.

Alguns métodos precisam de mais tempo para serem implementados do que outros.

A capacidade de implementação é muito importante e alguns dos temas da execução da avaliação são bastante difíceis, particularmente em alguns países onde há pouca experiência na aplicação de avaliações de políticas e programas.

5. FORMANDO A EQUIPE DE AVALIAÇÃO

A composição da equipe de avaliação é muito importante, pois requer uma gama de capacitações diferenciadas, como, por exemplo, analistas

quantitativos, qualitativos e com experiência com o tipo de projeto que está sob investigação.

A qualidade e utilidade final da avaliação de efeitos podem ser aprimoradas, coordenando-se, desde o início, os membros da equipe e os tomadores de decisão. Para tanto, é importante identificar os membros da equipe o quanto antes, decidindo suas funções e responsabilidades e estabelecendo mecanismos de comunicação nos momentos-chave da avaliação.

6. ELABORANDO OS DADOS

No caso de avaliações que geram suas próprias informações, são requeridos passos decisivos para o desenho dos instrumentos de coleta de dados, amostra, trabalho de campo, manejo e acesso de dados.

Não analisarei o processo, passo a passo, para a realização de uma pesquisa, mas mencionarei simplesmente os passos principais, que incluem:

- Desenhar e selecionar amostras;
- Elaborar instrumentos de coleta de dados;
- Reunir e capacitar pessoal para trabalho de campo;
- Realizar teste-piloto;
- Coletar dados;
- Acessar e gerenciar as informações.

7. COLETANDO DADOS ATUAIS

Inicia-se a fase de *execução do projeto*, que inclui a coleta de dados atuais.

8. ANALISANDO OS DADOS

Como nas outras etapas do processo, a análise dos dados de avaliação, sejam eles quantitativos, sejam qualitativos, requer a colaboração entre os analistas, os geradores de dados e as autoridades responsáveis, a fim de esclarecerem as perguntas e garantirem resultados oportunos e

de qualidade. Seguramente, surgirão problemas com a depuração e a interpretação dos dados durante a análise, e serão necessários os aportes de diversos membros da equipe. É também importante lembrar que o andamento da análise sempre demora mais do que o previsto.

9. REDIGINDO OS RESULTADOS E ANALISANDO-OS COM AS AUTORIDADES RESPONSÁVEIS E OUTRAS PARTES INTERESSADAS

Este ponto pode ser óbvio, mas é muito importante que os produtos e as informações sejam fáceis de ler e tenham relevância política. O administrador da avaliação deve planejar a geração de diversos produtos para políticos, administradores do programa, doadores e ao público em geral, jornalistas e acadêmicos. Se possível, é bom tornar a informação acessível a todos.

Esses produtos terão maior aplicação na política se incluírem recomendações claras e práticas provenientes da análise dos efeitos. As autoridades que decidem devem estar dispostas a considerar os resultados finais.

10. INCORPORANDO OS RESULTADOS NO DESENHO DO PROJETO

Se a avaliação está bem planejada e os objetivos são claros, isto deve estar bem implícito. Os problemas de implementação podem ser difíceis, mas esse é outro tema.

EXEMPLOS DE AVALIAÇÕES DE EFEITO DE “PROCEDIMENTOS ÓTIMOS”

Os quinze estudos de caso, incluídos em meu livro, foram selecionados de uma variedade de avaliações realizadas pelo Banco Mundial, por outras agências de doadores, instituições de pesquisa e empresas de consultoria privada. Foram escolhidos como exemplo de “prática ótima” e por seu rigor metodológico, que tentava refletir diversos exem-

plos de diferentes setores e regiões. Ainda que cada avaliação de impacto tenha seus pontos fortes e fracos, as lições aprendidas dessas experiências deveriam ajudar ao administrador de projetos ou ao analista político, que buscavam desenhar e implementar um futuro trabalho. Muitos exemplos vieram da América Latina, já que nessa região são feitas mais avaliações do que em outras.

Entre esses casos, existem avaliações de projetos na área da educação, saúde, programas de trabalho, microcrédito, alimentação e nutrição, infra-estrutura e agricultura.

Dos exemplos, tiramos alguns temas gerais.

PLANIFICAÇÃO ANTECIPADA E CUIDADOSA DO DESENHO DA AVALIAÇÃO

A preparação cuidadosa durante as etapas iniciais da identificação do projeto assegurará que se reúnam as informações corretas e que se possam empregar os resultados para ajustes dos componentes do projeto durante o processo. Com o planejamento antecipado e cuidadoso, é possível incorporar todos os elementos que contribuem para uma avaliação de efeitos rigorosa, como uma pesquisa de referência com um grupo de controle ao acaso e dados qualitativos sobre os processos que poderiam influir nos efeitos.

Um exemplo que é particularmente bom é o Projeto de Nutrição e Desenvolvimento Pré-Escolar de Uganda. Esta avaliação foi planejada com muitos detalhes, estando seus termos de referência inclusos no livro já citado.

AVALIAÇÕES SEM DADOS DE REFERÊNCIA

Na prática, muitas avaliações não têm dados adequados, havendo que se trabalhar com os dados existentes e, por essa razão, sem informações referenciais. Isto significa que os grupos de controle devem ser construídos. Alguns casos exemplares são os projetos de avaliação de Alimentos para Educação, de Bangladesh, Probecat, do México, Programas de Trabalho Ativo da República Tcheca e Projeto Trabalhar, da Argentina.

Na ausência de uma referência, devem ser criados controles usando-se os métodos de sincronização ou *matching*, mas isto pode ser bastante complicado. Tome-se como exemplo o caso dos métodos de sincronização (*propensity score matching*), que usou o Projeto Trabalhar da Argentina para criar um grupo de controle, com dados transversais sobre os participantes e os não-participantes do programa.

O Projeto Trabalhar II, da Argentina, é um programa de emprego público. Focou-se em prover emprego para pessoas de baixa renda, em pequenos subprojetos de infra-estrutura social e econômica, selecionados por grupos da comunidade. A avaliação de impacto foi desenhada para ponderar se a renda dos participantes era mais alta do que seria se tal programa não existisse. Os métodos mais comuns usados para calcular a renda dos lares, sem a intervenção do programa, não eram factíveis, no caso do Projeto Trabalhar. Nenhum estudo ao acaso (*randomization*) fora realizado para construir um grupo de controle com finalidade de comparação da renda dos beneficiários do projeto, nem se contava com pesquisas de referência, o que eliminava a possibilidade de realizar uma avaliação do antes e do depois.

A avaliação do Projeto Trabalhar usou, então, em seu lugar, os dados existentes, criando um grupo de controle mediante a reunião de participantes e não-participantes do programa nacional, em um conjunto de variáveis socioeconômicas, como sociedade civil, gênero, habitação, percepções subjetivas de bem-estar e filiação a partidos políticos e associações de vizinhos, usando uma técnica chamada qualificação de propensão.

O estudo demonstra o uso abundante dos dados existentes da pesquisa nacional em lares — a Entrevista de Desenvolvimento Social (EDS) —, para gerar o grupo de comparação, combinado com uma pesquisa mais reduzida dos participantes do Projeto Trabalhar, conduzida especificamente para os fins da avaliação. Essa pesquisa mais reduzida foi desenhada cuidadosamente, de modo que se usou o mesmo questionário que a EDS e as mesmas equipes de entrevistadores, tendo sido realizada aproximadamente ao mesmo tempo para se poder efetivar com sucesso o exercício de comparação. Esta técnica foi possível porque a pesquisa nacional de lares estava sendo realizada, então os avaliadores podiam

aproveitar esta oportunidade para tirar amostras dos participantes do Projeto Trabalhar. O fato das mesmas equipes de entrevistadores terem sido utilizadas para as entrevistas nacionais e as do projeto permitiu obter vantagens na coleta de dados.

LIMITAÇÕES AO ELABORAR CONTROLES ADEQUADOS

Os estudos aleatórios (*randomization*) ou os controles experimentais são às vezes possíveis, mas não factíveis no que se refere a políticas nacionais e locais. Neste caso, é possível efetuar os estudos aleatórios aproveitando-se de qualquer plano para testar o projeto em certas áreas restritas, que podem ser selecionadas, inicialmente, de forma aleatória, com possíveis áreas futuras planejadas como controles. Com o tempo, podem-se incluir, aleatoriamente, comunidades adicionais ao projeto.

Três exemplos ilustram como manejar uma situação na qual o estudo aleatório não era exequível relativamente à política ou a outro aspecto.

- No Vietnã, será avaliado um projeto de transporte rural com informação limitada e sem estudo aleatório.
- O Fundo Hondurenho de Investimento Social proporciona um exemplo de como criar um grupo de controle em um projeto dirigido pela demanda, usando-se uma comparação *ex post*, tendo por base um corte transversal único de dados. Avaliar projetos dirigidos pela demanda pode ser especialmente difícil, dado que não se sabe, quais projetos ou quais comunidades estarão deles participando de forma antecipada.
- A avaliação do Fundo de Investimento Social da Bolívia, na região do Chaco, proporciona um bom exemplo de como incorporar o estudo aleatório em projetos dirigidos pela demanda, de forma que permita a focalização.

MÉTODOS DE COMBINAÇÃO

Na maioria das avaliações, é requerida mais de uma técnica para se obterem resultados sólidos que aborem as várias perguntas da avaliação. Cada pergunta poderia necessitar diferentes técnicas, incluídas

dentro de um desenho de projeto. Três casos ilustram como se combinaram diversas técnicas em avaliação: o Fundo Social, da Bolívia, a Avaliação Trabalhar, da Argentina, e a Reforma Escolar, da Nicarágua.

REFORMA DA AUTONOMIA ESCOLAR DA NICARÁGUA

Em 1993, o governo da Nicarágua tomou medidas para implementar uma importante decisão de descentralização no setor da educação, outorgando autonomia orçamentária e de administração a escolas selecionadas, primárias e secundárias. A meta das reformas é melhorar a aprendizagem dos alunos: à medida que a administração de escolas se torna mais democrática e participativa, dando maior liberdade e controle de seu orçamento às administrações locais, há melhoria na qualidade do ensino, impulsionando o rendimento escolar dos alunos.

O efeito dessa reforma foi avaliado, usando-se uma combinação de técnicas quantitativas e qualitativas que examinam o resultado e o processo de descentralização.

O propósito do *componente qualitativo* é esclarecer se estão realizando ou não, nas escolas, as reformas previstas de administração e financiamento, pesquisando-se os aspectos positivos ou negativos. O *componente quantitativo* reforça esses resultados, respondendo à pergunta: As mudanças na administração e financiamento das escolas realmente produzem melhores resultados na aprendizagem dos alunos?

Os *resultados qualitativos* indicam que o êxito das reformas depende, em grande parte, do cenário no qual está inserida a escola (quer dizer, o nível de pobreza da comunidade), ao passo que os *resultados quantitativos* mostram que o aumento da tomada de decisões das escolas está associado, de forma significativa, ao aumento do rendimento dos alunos.

Foram utilizadas metodologias e fontes de dados diferentes, mas complementares, para combinar com os diferentes enfoques. Por um lado, a *avaliação quantitativa* seguiu um desenho quase-experimental, no qual se compararam as qualificações das provas de uma amostra de alunos de escolas autônomas (grupo de tratamento) com os resultados de uma

amostra de escolas públicas não-autônomas e de escolas privadas (grupo de comparação). Os dados para este componente da avaliação foram compilados de um painel de duas entrevistas em escolas-lares testadas, além das qualificações de provas de rendimento dos alunos. Por outro lado, o desenho da *avaliação qualitativa* consistia de uma série de entrevistas a informantes-chave e de conversações com grupos-objetivo, com diferentes pais e funcionários da escola, advindos de uma subamostra das escolas autônomas e tradicionais, incluídas no estudo quantitativo.

○ uso de técnicas de pesquisa qualitativa e quantitativa gerou uma valiosa combinação de resultados úteis e pertinentes para a política.

○ *trabalho quantitativo* proporcionou uma ampla visão geral e estatisticamente válida das condições e resultados das escolas; o *trabalho qualitativo* melhorou tais conclusões, informando por que alguns resultados esperados do programa de reforma tiveram êxito, enquanto outros fracassaram, ajudando, portanto, a orientar os ajustes da política. Além disso, como o trabalho qualitativo é mais intuitivo, mais acessível e, por conseguinte, mais interessante para o pessoal do Ministério, ele gerou uma facilitação na credibilidade do processo de avaliação na esfera ministerial.

CUSTOS E FINANCIAMENTO

Como vêem, existem muitos custos potenciais em um estudo de impacto, que incluem o tempo dos pesquisadores e a coleta de dados.

Somente algumas avaliações de nossa amostra tinham dados sobre os custos, mas temos dez avaliações feitas com a participação do Banco Mundial com estimativas dessa natureza. Esses custos não incluem o valor a ser pago para os funcionários dos países conveniados, cujo montante pode ser significativo. Em certos países, o governo pode ter algumas pessoas trabalhando em um projeto, por um ou dois anos.

○ custo médio estimado para a avaliação de impacto foi de 433 mil dólares. Isto reflete uma flutuação entre 263 mil dólares — como a avaliação de um programa de capacitação de habilidades vocacionais

para jovens desempregados de Trinidad e Tobago — e 878 mil dólares — como a avaliação do Fundo de Investimento Social da Bolívia.

O gasto com avaliações de impacto para a manutenção dos projetos reflete, em média, 0,6% de seu custo total (que às vezes inclui financiamento de vários doadores), ou 1,3% do custo do empréstimo do Banco Internacional de Reconstrução e Desenvolvimento (BIRD) ou do crédito da Associação Internacional de Desenvolvimento (AID).

Os componentes mais custosos para a manutenção das avaliações indicadas foram a coleta de dados e os assessores locais e internacionais. Em muitos dos casos, os custos incluíram viagens do pessoal local para se reunir com funcionários do Banco Mundial, e de pesquisadores para sessões de estratégia e capacitação, porque um objetivo-chave era a criação de capacidade em pessoal do país-cliente.

O financiamento de uma avaliação de impacto pode vir de um projeto com outros recursos governamentais, de uma subvenção para pesquisa ou de um doador externo.

A informação para uma amostra de avaliações do Banco Mundial indica que, apesar de muitos países assumirem a maior parte dos custos de avaliação, a execução exitosa de uma avaliação de impacto requer consideráveis recursos externos, além dos previstos em um empréstimo ou crédito para um projeto.

Esses recursos vieram das seguintes fontes:

- empréstimo ou financiamento de crédito do Banco Mundial para a coleta e processamento de dados;
- verbas do governo, destinadas aos salários pagos ao pessoal local nomeado para o esforço da avaliação;
- subvenções para pesquisa do Banco Mundial e subvenções de doadores bilaterais, que financiaram a assistência técnica de assessores com conhecimentos especializados, requeridos especificamente para a avaliação; e

- pressuposto de gastos gerais do Banco Mundial, proporcionado pelo tempo do pessoal para orientar a avaliação de efeitos e, com frequência, para participar ativamente do trabalho analítico.

ASSUNTOS DE ECONOMIA POLÍTICA

Como mencionei anteriormente, o apoio político é muito importante. A decisão de efetuar uma avaliação depende muito de um apoio político sólido.

Muitos governos não vêem o valor de efetuar projetos de avaliação e, portanto, não desejam investir recursos nisto. Além do mais, podem ter receio ao permitir uma avaliação independente, que venha revelar resultados contrários à política governamental, em particular em regimes autoritários ou fechados. Os governos mais abertos, no entanto, consideram as avaliações e a divulgação dos resultados como parte importante do processo democrático.

As avaliações também são sensíveis à mudança política. Três dos oito estudos de efeitos indicados pelo Banco, mencionados no livro, foram cancelados por problemas de economia política. As alterações dos regimes ou nos postos-chave dentro de uma administração governamental e as mudanças de estratégia política podem afetar não só o esforço de avaliação, mas fundamentalmente a execução do programa que se está avaliando.

Um exemplo é a Venezuela, que redesenhou uma avaliação de um programa de saúde e nutrição materna e infantil três vezes, com três contrapartidas diferentes de clientes, posto que o governo transferiu a responsabilidade da avaliação de um organismo a outro. Cada mudança esteve acompanhada de uma renegociação de contrato com a empresa do setor privado que se tinha prontificado a realizar a coleta de dados e a maior parte da análise para a avaliação. Quando a legitimidade da terceira contrapartida do governo começou a ser questionada, a empresa anulou o contrato e a avaliação foi abandonada.

Tais incidentes aconteceram durante um período de mudança política, caracterizado por várias reorganizações de gabinete, que terminaram com o fracasso do governo eleito, o qual servia de contrapartida do projeto, de modo que a avaliação dificilmente era a única que se via afetada pelas repercussões da instabilidade política.

NOTA

* Palestra proferida em espanhol e traduzida para o português por Ana Rojas Acosta e Maria Zélia Martins, com revisão final de Maria Cecília Roxo Nobre Barreira. Este trabalho está baseado no livro do mesmo título e pode ser encontrado no site do Banco Mundial (<http://www.worldbank.org/poverty/library/impact.htm>) — BAKER, Judy L. *Evaluating the Impact of Development Projects on Poverty: a handbook for practitioners*. The World Bank, Washington, D. C., USA, 2000.

Jean Hébrard

Avaliação dos
sistemas
educativos e
avaliação dos
resultados
escolares dos
alunos*

JEAN HÉBRARD é inspetor-geral da educação nacional do Ministério da Educação Nacional da França. Desenvolve programas educativos em diversos países, mais particularmente na América Latina. Coordenou, durante muitos anos, a cooperação educativa bilateral entre o Brasil e a França. É professor associado ao Centro de Pesquisas sobre o Brasil Contemporâneo da École des Hautes Études en Sciences Sociales de Paris.

Os diferentes dispositivos de avaliação dos sistemas educativos avançados podem ser remetidos a três grandes modelos. O primeiro, o mais tradicional, verifica a adequação do trabalho dos atores à regulamentação e aos objetivos definidos pelo organismo de tutela (trata-se, em geral, de um sistema público). O segundo, bastante difundido atualmente, tenta medir a pertinência das escolhas feitas por um operador que intervém, de maneira excepcional, em um dado momento da história do sistema educativo, na perspectiva de seu desenvolvimento. O terceiro, que só se desenvolveu durante a explosão escolar da metade do século XX (aumento da duração e do nível de escolaridade do ensino de massa), interessa-se, sobretudo, pelos resultados obtidos pelas diferentes unidades de ensino, que conservam um forte grau de liberdade na organização de seu trabalho.

Os primeiros correspondem aos sistemas estatais centralizados, como o da França, desde o fim do século XIX até os anos 80; os segundos, aos países ou regiões que sofrem um atraso em seus investimentos educativos e utilizam os recursos dos grandes organismos internacionais para superá-lo; os últimos, aos sistemas contratados nos quais os financiamentos são distribuídos não somente em função dos efetivos escola-

rizados, mas dos projetos de cada unidade e dos resultados obtidos no fim do contrato. Na maioria dos países, as universidades (públicas ou as que recebem subsídios públicos) foram as primeiras a se engajar nesse tipo de relação com suas tutelas. Há cerca de uma década, os ensinos de 1º e 2º graus passam a fazer o mesmo com seus organismos de tutela, não sem novos problemas. Estes dispositivos tendem a substituir tanto os de primeiro tipo (verificação da adequação às regras) quanto os do segundo (*expertise* de escolhas operacionais).

Para tentar compreender como se dá essa evolução, vamos nos deter aqui na avaliação dos sistemas de ensino básico (escolas primárias) que se dedicam à alfabetização. Em grande parte dos países, tais sistemas parecem ser os menos aptos a entrar em uma outra lógica de avaliação que não aquela da conformidade à regulamentação pública (é caso dos sistemas públicos centralizados) ou da pertinência das escolhas operacionais (caso dos pedidos de financiamentos internacionais).

Dois exemplos, bastante distantes um do outro, podem esclarecer a problemática e apoiar a discussão em bases mais concretas. O primeiro é o de um projeto de auxílio à melhoria do funcionamento de 25 escolas públicas em quatro estados do Brasil (Projeto Escola que Vale), envolvendo uma fundação (Fundação Vale do Rio Doce) e uma organização não-governamental (Cedac). Tal projeto visa experimentar e desenvolver um modelo de intervenção “clínica” em políticas públicas de escolarização (Estado, município), cujo interesse é evidente para os sistemas educativos pouco centralizados, como os de muitos países sul ou norte-americanos. O dispositivo de avaliação que foi implementado tem, *a priori*, uma dupla função: permitir à equipe responsável pelo projeto um olhar exterior sobre as diferentes etapas de sua ação e prestar contas aos sócios capitalistas do bom emprego do financiamento obtido.

O segundo exemplo é o da implementação, por um Estado (a França), de uma lei escolar (lei de orientação, de 1989), que implica um novo uso das avaliações regulamentares (controle de conformidade) e de avaliações dos resultados dos alunos (redução de fracassos), no âmbito de uma avaliação permanente de políticas públicas. Neste caso, a avaliação leva em conta a *expertise* de funcionários públicos (inspetores) e o

trabalho dos diretores ministeriais (direção de avaliação). Ela envolve todos os níveis de uma administração — nível central e nível descentralizado — e diferentes atores do sistema, como professores, diretores, conselheiros pedagógicos, inspetores.

AVALIAÇÃO DE UM PROJETO DE AUXÍLIO À MELHORIA DO FUNCIONAMENTO DE ALGUMAS ESCOLAS PÚBLICAS EM QUATRO ESTADOS DO BRASIL.

O Programa Escola Que Vale, voltado à formação continuada de professores do ensino público fundamental, tem por finalidade dotar seus sujeitos de novas competências capazes de gerar práticas educativas mais densas de valores, conteúdos e processos de aprendizagem.

Os professores são sujeitos históricos que vivem e aprendem muita coisa em seu cotidiano, que apreciam a beleza, que usufruem sua cultura e que, diante de uma boa oportunidade, podem fazer uma “alquimia” de tudo isso e se apresentar também com produtores de cultura.

A estratégia proposta para a capacitação do professor centra-se no desenvolvimento de Projetos de Trabalho, estruturados de forma a contemplar a ênfase na leitura/escrita/comunicação oral, a resolução de problemas, a prática da cidadania, o uso de recursos tecnológicos, vendo a comunidade como sujeito e locus de conhecimento.

O programa introduz, igualmente, a oferta de insumos que buscam assegurar condições para a capacitação dos docentes, assim como responder de forma convergente às demandas dos demais sujeitos implicados: a escola (incluindo aqui melhorias no equipamento); os professores da rede (incluindo oficinas de sensibilização e capacitação); os alunos, com acesso a tecnologias e situações pedagógicas para que possam aprender melhor e de forma significativa; os pais e a comunidade, para que possam assumir seus papéis de sujeitos de aprendizagem.

O Programa assenta-se em parcerias com as prefeituras, secretarias municipais de educação, equipes de escolares, pais e comunidade, visando a um processo de mobilização e irradiação capaz de produzir impactos efetivos na melhoria da aprendizagem dos alunos, na apropriação de valores democráticos e nos vínculos entre escolas e comunidade local. (Pesquisa avaliativa do Programa Escola Que Vale — Relatório Parcial. Maria do Carmo Brant de Carvalho (Coord.). São Paulo, agosto de 2000.)

○ Projeto Escola que Vale foi elaborado pelo Centro de Estudos e Documentação para Ação Comunitária — Cedac, criado, há alguns anos, por Beatriz Cardoso. Dentre suas diferentes ações, a equipe do Cedac já havia tido a oportunidade de desenvolver uma ação de formação, particularmente original, de professores rurais no Estado de São Paulo, baseada em uma rede de troca de correspondências que permitia ultrapassar o isolamento dos professores de escolas muito distantes umas das outras. Nesta ação, foram elaborados instrumentos didáticos de grande qualidade e uma filosofia de formação centrada mais na troca de experiências concretas do que em um metadiscorso pedagógico.

○ Projeto Escola que Vale¹ é mais ambicioso, uma vez que visa definir um modelo de intervenção no sistema educativo público, leve o bastante para não parecer intempestivo às autoridades de tutela e suficientemente eficaz para que sua ação ultrapasse rapidamente as poucas escolas envolvidas (25 no total). Trata-se de um projeto que tenta se automultiplicar. Baseia-se, portanto, menos em modelos clássicos de ensino e de formação — encontrados nas faculdades de educação ou em institutos de formação pedagógica — do que na criação de competências de autoformação dos próprios atores (“construir um modelo de intervenção cooperativo num sistema público de ensino, que ofereça condições de trabalho e permita a apropriação do processo por parte de todos os implicados²). Esta característica conduziu os responsáveis a imaginar um dispositivo no qual a formação pessoal não apenas é inseparável da formação profissional; mais que isso, a orienta.

Lançado em 1999, o Projeto Escola que Vale³ desenvolveu, até o momento, um cronograma de trabalho no qual todas as etapas encadeiam-se com precisão:

- elaboração do projeto e negociação com as instituições financiadoras;
- enquete em municípios passíveis de se interessar pelo projeto;
- solicitação de uma pesquisa sobre 45 escolas sondadas por um organismo exterior (realização de monografias sobre cada uma delas, permitindo explicitar as dificuldades específicas de cada uma e as possibilidades dinâmicas dos atores), visando facilitar a escolha;
- escolha dos Estados (quatro), dos municípios (seis) e das escolas (25) envolvidas;⁴
- implementação do dispositivo de acompanhamento das escolas participantes por um coordenador local e um coordenador regional; estabelecimento das modalidades de trabalho do conjunto assim constituído (escolas, municípios, coordenadores, Cedac);
- primeira ação de formação: ateliês de prática artística pessoal (oficinas de arte), dirigidos aos professores;
- elaboração de instrumentos destinados a promover, nas classes, “eventos didáticos” ao longo do ano 2000: agenda, proposição de “projetos” didáticos;
- no início de 2000, lançamento dos “projetos”, ativação do acompanhamento local pelos coordenadores, supervisão do acompanhamento pelo Cedac;
- no fim do primeiro semestre, avaliação da etapa realizada por uma equipe externa; emendas ao projeto e lançamento de projetos para o 2º semestre.

Como a avaliação intervém nesse esquema? Ela ocupa um lugar determinado no projeto inicial e, de modo geral, visa oferecer à equipe do Cedac um olhar externo em cada uma das etapas de seu trabalho. **Permite também a esta ONG prestar contas à fundação e à empresa que financiam e incentivam a operação.** Eis aqui um esquema bastante simples encontrado em todos os programas desse gênero, particularmente quando são financiados por grandes organismos internacionais. No entanto, observando mais de perto, a articulação entre avaliação e

desenvolvimento do projeto revela-se mais complexa e merece maior atenção.

Pode-se considerar que este projeto funciona com três lógicas de avaliação bem distintas. A *primeira* se refere a um dispositivo de *expertise* externa, que intervém, *a posteriori*, no momento da concepção do projeto. Ela visa, portanto, os atores situados na origem do projeto (essencialmente o Cedac, mas também os organismos de apoio) e se baseia em suas produções escritas — o projeto propriamente dito, instrumentos de análise, resultados das análises, instrumentos de formação ou instrumentos didáticos etc. O pedido dos operadores que solicitaram esta *expertise* formula-se como controle e conselho: controle da coerência das diferentes propostas (de sua pertinência em relação aos dados disponíveis etc., e conselho para a melhoria da concepção.

Para o especialista, trata-se primeiro de “compreender” o projeto, isto é, de apreciar sua legibilidade. Para tanto, ele dispõe de sua experiência, bem como de sua própria cultura, e trabalha por assimilação: no projeto, tudo o que é “reconhecido” — isto é, referente a uma de suas categorias de análise — constitui sua base de apreciação. O projeto é imediatamente inscrito em uma tipologia de projetos educativos possíveis (a tipologia de que o especialista dispõe) e avaliada na esfera desta categoria. Tal categoria se define, em geral, por um ideal-tipo (ou por um conjunto de características e normas).

Em sua singularidade, o projeto é então examinado com base nas diferenças com o modelo ao qual é assimilado: diferenças aceitáveis, diferenças possíveis, diferenças que correm o risco de criar conflitos irreduzíveis ou de colocar em risco a coerência do projeto. É evidente que, para o especialista, tais diferenças constituem uma série de experiências novas, que vêm enriquecer seu repertório de modelos. O julgamento da *expertise*, neste nível, é pouco capacitado para verificar a adequação do projeto às situações concretas dos sistemas educativos (ou de parcelas desses sistemas) sobre os quais se apóia.

É neste momento que intervém o *segundo dispositivo de avaliação* implementado pelo Cedac. Sua originalidade reside no fato de que ele

“externaliza” todo o dispositivo de enquete sobre a realidade estudada e solicita, então, a uma equipe de pesquisa,⁵ a análise da situação do sistema educativo em questão, assim como das características dos atores e das populações sobre as quais o projeto vai incidir. Esta equipe terá por missão apreciar a adequação do projeto a esta realidade, bem como os resultados obtidos. Os relatórios produzidos nessa ocasião permitiram, até o momento:

- lançar um olhar avaliativo global sobre o projeto do ponto de vista de uma ética das ações sociais;
- facilitar as escolhas dos locais de experimentação (tipos de dificuldades, capacidade de mutação etc.);
- seguir a evolução das ações previstas e regular os investimentos dos atores;
- apreciar o impacto do projeto durante sua realização, em cada um dos eixos definidos como prioritários (alunos, professores, escolas, famílias, comunidade, parceiros da rede municipal).

A metodologia da equipe encarregada deste trabalho se desenvolve em três direções complementares:

- análise dos dados estatísticos disponíveis (macroindicadores) em nível federal, nacional e municipal; assim, as escolas participantes são situadas em um contexto mais amplo, permitindo apreender suas características demográficas, econômicas e sociais, bem como o grau de desenvolvimento do sistema educativo no qual elas se inserem (índices de escolarização, de repetência, de evasão escolar; nível de formação dos professores, despesas dos municípios etc.).
- observação clínica dos diferentes atores do dispositivo local (crianças, famílias, professores, coordenadores, equipes municipais); eles são consultados acerca do grau de concordância sobre o programa, os efeitos do programa que já se fazem perceber, as dificuldades que sentem etc.
- análise das respostas dadas pelos professores nos questionários em relação às atitudes das crianças e à apreciação dos efeitos do programa nelas observados.

Como podemos notar, esta segunda lógica de avaliação concerne essencialmente às representações dos professores, tanto no que se refere ao seu próprio trabalho, quanto ao trabalho dos alunos. Ela é centrada nas dinâmicas de mudanças e nas atitudes. Tenta apreciar, por meio de discursos cruzados de diferentes atores, todos os signos que permitem afirmar que o programa produz efeitos significativos na percepção da realidade, nas hierarquias de valores e na mobilização do sujeito (ou do grupo).

O terceiro dispositivo de avaliação pode ser considerado um processo *online*, que, em todos os níveis, exige dos atores momentos de regulação cruzada que parecem ser decisivos. É o caso, por exemplo, das notas regularmente enviadas pelo coordenador regional ao Cedac. Quando ele visita uma localidade, organiza uma série de reuniões — com o coordenador local, professores, supervisores, diretores, o Secretário Municipal de Educação e seus colaboradores —, cada uma delas é gravada (áudio) e torna-se objeto de um relatório que é enviado ao Cedac. Cada participante faz uma auto-avaliação e uma avaliação espontânea da ação em andamento, referente ao nível imediatamente inferior: os professores avaliam seu trabalho e o de seus alunos; os supervisores avaliam seu trabalho de supervisão e o trabalho dos professores etc. Tendo em vista esse material abundante, cabe ao coordenador regional triar aquilo que julga pertinente, tanto para controlar sua ação quanto para informar o nível superior (no caso, a equipe central do Cedac) e receber, deste último, uma avaliação de sua ação.

A leitura desses documentos⁶ mostra como os atores exercem efetivamente seu papel no dispositivo da avaliação *online*. De modo geral, eles chamam a atenção sobre as distorções que aparecem entre os grandes eixos do projeto geral e as possibilidades de sua implementação:

- incompreensão dos objetivos;
- reflexos de resistência dos professores;
- impossibilidade de desenvolvimento de certos aspectos do projeto por inadequação de instrumentos;
- subobjetivos desnecessários, etc.

No caso específico aqui examinado, a coordenadora regional do Maranhão

chama a atenção da equipe do Cedac sobre o que parece ser um dos problemas da implementação: no Projeto Escola que Vale, faltaria (no nível manifesto) uma série de proposições precisas e explícitas sobre o acesso aos três instrumentos de base das ações pedagógicas previstas — a palavra, a escrita, a imagem. Isto se traduz, na transcrição das falas dos professores durante as reuniões de trabalho, pela seleção de proposições referentes às dificuldades dos alunos neste aspecto, apresentando-as como contraditórias em relação ao projeto.

Assim, parece que uma prática fortemente centrada na formação contínua de professores e no acompanhamento da preparação didática de seu trabalho (supondo ademais que esta formação e este acompanhamento serão suficientes para promover uma dinâmica de mudança das práticas profissionais) vem se chocar com as representações das supostas capacidades dos alunos, construídas por todos os atores envolvidos.

As respostas sugeridas pela coordenadora regional merecem também ser examinadas na medida em que tentam sair desta dificuldade, por vias didáticas e formativas, e não por via avaliativa. Para ela, com efeito, as atitudes dos professores são interpretadas prioritariamente como reticências a suprimir mediante melhoria da formação (particularmente a formação pessoal).

AVALIAÇÃO DE UMA POLÍTICA PÚBLICA NA FRANÇA: PAPEL DO DISPOSITIVO DE AVALIAÇÃO NACIONAL DOS ALUNOS DA ESCOLA PRIMÁRIA NA IMPLEMENTAÇÃO DA LEI DE ORIENTAÇÃO DE 1989

O processo francês que desejo analisar agora é conseqüência de um ato legislativo: o voto parlamentar, em 1989, de uma “lei de orientação”, a qual define objetivos ambiciosos para o conjunto da escolaridade pré-obrigatória e obrigatória (entre dois e dezesseis anos). É solicitado aos professores tentar de tudo para que nenhum aluno saia do sistema educativo sem uma qualificação, no mínimo equivalente ao nível V da escala de qualificações profissionais (Certificado de Aptidão Profissional — CAP, ou Brevê de Estudos Profissionais — BEP).⁷

No momento em que a lei foi votada, mais de 20% dos alunos de cada faixa etária não atingiam o nível V e saíam do sistema educativo sem qualificação. No contexto do mercado de trabalho, extraordinariamente reduzido nos anos 1980-1990, esta situação conduzia, de modo inexorável, ao desemprego de longa duração e a situações escolares extremamente precárias.⁸ No que tange à pré-escola (crianças de dois a seis anos — mais de 95% das crianças escolarizadas desde os três anos e 100% com quatro anos) e à escola elementar (escolarização obrigatória de seis a onze anos), a lei prevê três dispositivos:

- predominância da organização da escolaridade em três ciclos de três anos cada⁹ (em vez de ciclos anuais): só é possível repetir uma vez ao longo do ciclo e o programa deve ser percorrido segundo um ritmo decidido pela equipe de professores do ciclo, os quais dispõem de uma hora semanal de reunião para harmonizar seu trabalho;
- obrigação de cada escola redigir, a cada três anos, um projeto educativo¹⁰ (“projeto de escola”), que é homologado pela administração e que permite, em relação aos objetivos definidos em nível nacional (programas dos conteúdos e competências a atingir ao fim do ciclo), adaptar a ação da escola às especificidades dos alunos que a frequentam e aos recursos humanos e materiais de que dispõe;
- implementação de uma avaliação nacional (em língua materna e em matemática) de todos os alunos em fim de ciclo II e em fim de ciclo III;¹¹ definição de um nível de competência mínimo para aceder ao ciclo superior que, quando não atingido, não aciona uma repetência, mas uma medida de responsabilidade diferenciada (projeto pedagógico de ajuda personalizada).

É interessante notar como o dispositivo tradicional de avaliação de conformidade à regulamentação nacional pode se adaptar a uma lei que, de um lado, exige a abertura de um espaço maior à iniciativa local, mas, de outro, introduz um sistema de avaliação dos alunos, o qual permite situar precisamente a eficácia de cada escola ou de cada unidade administrativa encarregada das escolas; isto é, tenta completar o controle de conformidade por um controle sobre o rendimento do sistema.

Na França, o controle mais próximo das escolas primárias (pré-elemen-

tares e elementares) é realizado por um “inspetor da Educação Nacional”,¹² responsável por uma circunscrição com cerca de 300 a 350 professores (trinta a quarenta escolas urbanas, sessenta a oitenta escolas rurais). Ele é cercado por uma equipe de dois ou três professores, que não estão mais em sala de aula e que possuem o diploma de professor-formador,¹³ além de uma secretária. Ele assegura a gestão administrativa das escolas relacionadas à sua tutela (na França, os municípios), a formação contínua dos professores e a recreação pedagógica das escolas; enfim, ele inspeciona cada professor ao menos uma vez a cada três anos, e, nesta ocasião, confere-lhe uma nota que contribui para sua evolução na carreira. Ele próprio está diretamente subordinado a um inspetor de academia, que dirige os serviços da Educação Nacional no âmbito do departamento,¹⁴ distribui os recursos humanos alocados pelo Estado e impulsiona a política nacional tal qual é definida pelo ministro.

As três medidas, previstas pela lei de orientação de 1989 e imediatamente aplicáveis, na verdade modificaram as práticas docentes de maneira muito diferente. Parece que, como é freqüente no caso dos sistemas centralizados, as medidas que demoram mais para entrar em vigor são aquelas que implicam mais os atores (os professores em suas classes, as equipes de professores nas escolas) e seus atos profissionais. Aquelas que atingem os atores intermediários (professores-formadores, inspetores, administração descentralizada) têm necessidade de um certo tempo de gestação, mas se efetivam com mais rapidez, transformando a prática desses profissionais. E, enfim, aquelas que implicam diretamente a administração central são imediatamente colocadas em prática e influenciam com rapidez as práticas dos professores em sala de aula, mas não acionam tão depressa as atitudes esperadas, particularmente quando exigem certa iniciativa. Em geral, elas são reinterpretadas pelos professores e adaptadas tanto às condições de trabalho quanto às representações que estes fazem do seu ofício.

Mais de dez anos depois da lei entrar em vigor, pode-se afirmar que a organização da grade escolar em ciclos de três anos e a pedagogia diferenciada que ela implica são encontradas apenas excepcionalmente quando se visita as escolas francesas. Apesar de todos os esforços de

formação e dos reiterados incentivos dos poderes públicos, os professores parecem não conseguir sair de uma pedagogia frontal¹⁵ e conceber velocidades variáveis de percurso do ciclo, para cada um de seus alunos. Na verdade, eles não dispõem ainda de instrumentos didáticos que facilitariam seu trabalho ou, quando existem, eles não sabem (ou não querem) deles se utilizar.

Dois saídas parecem, contudo, se delinear: um certo número de editores de livros didáticos começa, timidamente, a propor livros abrangendo a totalidade do ciclo e não somente um de seus anos; para responder às reiteradas solicitações de seus inspetores, numerosas equipes de professores de ciclo fabricam, disciplina a disciplina, referenciais de competências simplificados de fim de ciclo e boletins de avaliação passíveis de serem transmitidos às famílias. Para os professores, é a ocasião de harmonizar seu trabalho e de melhor dividir os esforços dos alunos durante os três anos do ciclo.

É interessante observar que os inspetores não levam em conta estas reticências em suas avaliações sobre os professores, quando de sua inspeção, e que, desde 1989, os sucessivos ministros nunca solicitaram uma pesquisa sobre o assunto.

A segunda medida — projeto de escola — foi colocada em prática mais rapidamente. Entretanto, isso foi feito a custo de modificações significativas em seus objetivos, que só puderam ser corrigidas recentemente. Com efeito, em um primeiro momento, as escolas confundiram “projeto de ensino” e “projeto de equipamento”. Elas se lançaram na realização (com financiamento obtido graças aos créditos de incentivo) de bibliotecas, de reformas no pátio, em salas de informática, na decoração das escolas etc. Nenhum desses projetos foi verdadeiramente pensado como modificador da prática dos professores em sala de aula. Na maioria das vezes, eles levam em consideração certo consumismo escolar (“é preciso ter uma biblioteca!”) ou ainda a concorrência entre escolas (“mesmo em um bairro difícil, podem-se ter computadores!”).

Do mesmo modo, as escolas confundiram, com frequência, “projeto de ensino” com “projeto festivo”, imaginando que programar grandes festas ou espetáculos excepcionais permitiria agrupar todas as classes em

torno de um objetivo comum. É verdade que eles foram estimulados a caminhar nesta direção pela ação do Ministério da Cultura que, nos anos 80, disponibilizou créditos importantes para que os jovens artistas, trabalhando nas escolas, pudessem tanto beneficiá-las, por meio de suas competências, como encontrar complementos salariais em um período difícil para as profissões artísticas. O mesmo ocorreu com o Ministério da Juventude e dos Esportes em relação aos jovens atletas.

Foi em razão de uma avaliação bastante clássica da política governamental que esta situação pôde ser modificada. Na verdade, um dos atos mais importantes do primeiro governo de esquerda, realizado pelo presidente François Mitterand, em 1981, foi a criação de “zonas de educação prioritária”(ZEP)¹⁶ nos bairros considerados difíceis. Essas ZEPs foram pensadas tomando-se por base o modelo de educação compensatória desenvolvido pela Grã-Bretanha alguns anos antes. As escolas dessas zonas se beneficiavam de um adicional importante de recursos humanos e financeiros. Ora, a implementação de avaliações nacionais dos resultados escolares em 1989 e sua reedição anual pareciam mostrar que o diferencial entre escolas em ZEPs e escolas fora das ZEPs continuava tristemente idêntico ano após ano, apesar dos importantes esforços realizados tanto pelo Estado como pelos próprios professores. Em alguns casos, ele até mesmo se agravava à medida que as condições sociais desses bairros, particularmente sensíveis à crise econômica dos anos posteriores à Guerra do Golfo, se degradavam.

Em 1996, ocasião do retorno ao poder de um novo governo socialista, que se sentia herdeiro da política dos anos 80, decidiu-se confiar uma missão de avaliação à inspeção geral da Educação Nacional.¹⁷ Tratava-se de compreender o que poderia levar uma zona de educação prioritária (ZEP) a fracassar ou ser bem-sucedida em suas políticas de compensação.

Para estabelecer seu diagnóstico, a equipe encarregada da avaliação¹⁸ utilizou indicadores recém-elaborados pela Direção de Avaliação e da Prospectiva (DEP), os quais permitem levar em conta, de maneiras diferentes, a composição social de uma escola para julgar os resultados dos alunos.¹⁹ Estabelecem-se os indicadores com base em uma amostra representativa da população francesa da qual se conhece a distribuição

socioeconômica (fazendo parte da amostra, segmentos de cada categoria socioprofissional).

Quando se quer apreciar o nível relativo (não mais absoluto) atingido por uma escola ou município, nesses mesmos indicadores, eles são corrigidos em função da composição social específica dessa população, de maneira a neutralizar as variáveis sociológicas. Isto permite calcular, “na constância de todas as variáveis”, o “valor educativo agregado” efetivamente pelos atores dessa zona, eliminando-se os fenômenos que seriam devidos tão-somente à realidade socioprofissional das famílias envolvidas. Esse valor agregado pode, é claro, ser positivo ou negativo.

Ao final dessa avaliação, observou-se que as ZEPs mais eficazes se caracterizam por:

- serem pequenas (um ou dois colégios, cada um com seiscentos a setecentos alunos, e as escolas é que lhes encaminham alunos, ou seja, têm de três a quatro mil alunos no máximo);
- seus alunos serem escolarizados precocemente na pré-escola (todos, aos três anos; uma parte importante, aos dois anos);
- seus projetos de escola serem prioritariamente voltados para as aprendizagens, mesmo que não deixem de lado as ações sociais ou culturais;
- fazerem uso regular das avaliações nacionais para controlar seu trabalho.

Tendo o ministro decidido tornar públicos os resultados, o debate que se seguiu (por ocasião de um fórum nacional, em que todas as equipes das ZEPs estavam representadas) permitiu levantar algumas fortes oposições. Em particular, observou-se que certos professores haviam preferido evitar qualquer tipo de conflito, ocupando o tempo dos alunos, o máximo possível, com recreações de forte valor cultural e social, na jornada diária e ao longo do ano. Eles chegaram, com frequência, a acalmar as populações difíceis pelas quais eram responsáveis sem, no entanto, obter melhores resultados escolares.

Em direção oposta, outras equipes tinham centrado seus esforços na aprendizagem, obtendo resultados escolares bastante estimulantes (um real “valor agregado” e, algumas vezes, valores médios superiores às mé-

dias nacionais), sem criar, no entanto, situações sociais mais conflituosas. No segundo caso, ocorria também que a escolarização precoce das crianças na pré-escola (a partir de dois anos) era mais centrada nas atividades de aprendizagem que nos jogos e festas. Evidentemente, a realidade de diferentes zonas prioritárias se dividia proporcionalmente entre esses dois modelos.

Tais constatações e os debates que se seguiram permitiram repensar, globalmente, a política dos projetos de escola. Em particular, eles abalaram profundamente as convicções dos inspetores das circunscrições que, até então, haviam aceitado o que, hoje, parece ser uma deriva “social” do dispositivo. Assistiu-se, então, em alguns anos, à redefinição completa da noção de projeto e à sua articulação com as avaliações nacionais dos resultados dos alunos.

Daí em diante, quando uma escola renova seu projeto, ela começa por fazer um retrato da população com a qual trabalha. Ela avalia os indicadores socioeconômicos: divisão de categorias socioprofissionais das famílias, número de famílias que se beneficiam de recursos sociais, ajuda para o pagamento da alimentação escolar, número de famílias monoparentais etc. Em um segundo momento, a escola tenta avaliar a eficácia até então demonstrada, graças aos indicadores de sucesso escolar: índices de repetência, resultados nas avaliações nacionais, locais, número de crianças que precisam de um acompanhamento especial oferecido pelas redes de ajuda²⁰ etc. Em um terceiro momento, faz-se o balanço dos recursos humanos e materiais disponíveis para os próximos três anos do projeto, assim como dos recursos suplementares de que a escola poderia se beneficiar.

Considerando essas três dimensões, a escola tenta otimizar sua ação de maneira a aumentar sua eficácia, respondendo, o mais diretamente possível, às necessidades de seus alunos. Assim, os elementos que até então ocupavam o essencial do projeto — instalação de equipamentos, organização de eventos etc. —, longe de desaparecer, tornam-se instrumentos a serviço do projeto e continuam a dar sentido à ação da escola no seu entorno. Os inspetores que impulsionaram esta política vêm

reforçado seu papel de formadores, na medida em que intervêm, de maneira bastante ativa, na elaboração do projeto, objeto de uma série de idas e vindas entre a escola e a inspeção.

Os esforços das escolas são compartilhados, ao longo da fase de elaboração, por diversos encontros entre diretores que comparam suas propostas. Enfim, perante as famílias — que evidentemente participam da elaboração do projeto por intermédio de seus representantes eleitos no “Conselho de Escola” —, os professores de uma mesma escola possuem uma meta coerente e passível de ser defendida e argumentada. Desta forma, o projeto de escola tem um papel importante na regulação do funcionamento do sistema educativo primário francês, tanto no que tange às próprias equipes de professores (uma vez que estas assumem plenamente sua construção e seu acompanhamento) quanto no que se refere aos atores intermediários (inspetores), que têm nele um instrumento de diálogo e de estímulo em relação às escolas menos propositivas.

Foi a própria avaliação nacional dos resultados dos alunos que se revelou, no entanto, a mais fácil de ser implementada e, ao mesmo tempo, a mais rica em conseqüências (positivas e negativas). Seu funcionamento relativamente centralizado pode explicar seu sucesso, mas também alguns de seus defeitos.

Todos os anos, os inspetores de circunscrição são incentivados a apresentar projetos, em colaboração com professores que selecionam, visando à elaboração de baterias de provas, cujos critérios são definidos por uma comissão nacional (uma por nível e por disciplina). Esta comissão seleciona, entre todos os projetos, as provas julgadas mais pertinentes e constrói um dossiê de avaliação (cerca de trinta provas por disciplina e por nível), que é encaminhado a todas as classes envolvidas no início das aulas.

No fim da aplicação das provas, o professor encarregado corrige os trabalhos dos alunos e transpõe os resultados para tabelas (manuais ou informatizadas), que lhe permitem calcular as tendências centrais de sua classe em uma série de indicadores²¹ e, no caso dos alunos, definir sua distância, positiva ou negativa, em relação à média. Calcula-se também, para cada aluno, sua posição em uma escala de leitura composta de

quatro valores: não possui competências de base; possui competências de base; possui competências notáveis; possui competências excepcionais.²² Os resultados são enviados ao inspetor da circunscrição, que estabelece médias para cada escola e para o conjunto da circunscrição. Reagrupados no âmbito do departamento e, depois, da academia, esses resultados permitem definir médias para cada um dos estágios do sistema educativo.

Paralelamente, cerca de três mil alunos são sorteados para compor uma amostra representativa das escolas francesas; suas provas são agrupadas na Direção de Avaliação, analisadas e informatizadas de modo a dispor dos mesmos escores em escala nacional. Este duplo caminho permite remediar a ausência de escalonamento das provas e obter, em nível local — mas não em nível nacional — comparações longitudinais. Com efeito, como as provas da avaliação mudam todo ano, não é possível comparar diretamente os resultados obtidos por amostras nacionais sucessivas. Além disso, é possível seguir, ano após ano, a evolução de uma dada classe (isto é, um professor), de um ciclo (isto é, os professores de um ciclo), de uma escola (ou seja, os professores desta escola), de uma cidade, de uma circunscrição, de um departamento ou de uma academia em relação à amostra nacional.

A princípio, essa avaliação havia sido pensada na perspectiva do novo funcionamento em ciclos. Ela deveria permitir aos professores melhor conhecer os novos alunos que chegavam em suas classes e melhor analisar suas dificuldades. Assim, pensava-se que era possível promover um ensino mais individualizado, que permitisse, a cada aluno que não tivesse conquistado todas as competências do ciclo anterior, beneficiar-se de um ensino específico antes de se lançar em novas aprendizagens.

Na verdade, esse dispositivo foi desviado de seus objetivos iniciais. Primeiramente, pelos inspetores de circunscrição, que viram aí uma forma de acrescentar a seus instrumentos de avaliação tradicional uma ferramenta que lhes permitisse avaliar mais as equipes do que os indivíduos e, sobretudo, ultrapassar o controle sempre contestável (e contestado) dos recursos didáticos utilizados. Com as avaliações nacionais, podem dizer aos professores de uma escola que eles deveriam modificar seu

modo de trabalhar, não porque não utilizem tal ou tal método pedagógico, mas porque seus resultados são insuficientes. De qualquer forma, os inspetores remetem aos professores a responsabilidade pela escolha de instrumentos didáticos, contentando-se em verificar se as aprendizagens são efetivamente construídas.

Tal fato representa uma mudança de atitude muito importante em um país onde os inspetores sempre acharam que tinham o direito de escolher um método a outro e, algumas vezes, de impor os instrumentos que eles próprios publicavam. Esta mudança de atitude tem uma consequência direta sobre as ações de avaliação das pessoas tradicionalmente encarregadas de tais tarefas. Observaram-se, progressivamente, inspetores de escolas finalizarem as inspeções individuais requeridas pela regulamentação em vigor. Este trabalho foi inúmeras vezes acompanhado por ações de formação na própria escola, substituindo os estágios mais tradicionais nos quais os professores de determinadas escolas eram reunidos em torno do tema de formação. Nesses “estágios de escola”, as pessoas habituaram-se, gradualmente, a examinar os resultados da escola nas avaliações nacionais, a refletir em conjunto sobre as formas de melhorá-las e a pedir formação complementar sobre essas ações específicas. Com frequência, era também o momento de redigir adendos ao próprio projeto de escola.

É interessante observar, todavia, que, nesse primeiro momento, as respostas aos problemas escolares das crianças foram sempre concebidas como respostas globais; questiona-se mais a estratégia pedagógica da escola e a prevenção dos fracassos do que as respostas apropriadas aos fracassos constatados. Continuava muito difícil agir no nível individual, sobre cada criança com dificuldades, quando uma aprendizagem tinha sido falha. De qualquer forma, na impossibilidade de avançar em direção ao ensino diferenciado (em relação aos ciclos) e de articulá-lo às avaliações nacionais, tanto as equipes de professores quanto as de inspeção se voltaram para uma articulação entre as avaliações nacionais e o projeto de escola.

Para que a situação avançasse um pouco mais, foi necessária uma intervenção do próprio ministro. Ela ocorreu sob a forma de uma circular publicada em 28 de novembro de 1998 (quase dez anos após a promul-

gação da lei),²³ solicitando que toda criança, cujas graves dificuldades foram detectadas pelas avaliações nacionais de fim de ciclo II (ou seja, depois da entrada na escrita), fosse foco de um projeto pedagógico de ajuda personalizada, controlado pelo inspetor da circunscrição, até que ela tivesse conseguido, pelo menos, o nível de base do fim do ciclo III.

Três anos após a entrada em vigor desta circular, é claro que a maioria das escolas continua envolvida neste problema. Quando os inspetores foram particularmente atentos em promover ações, parece que a elaboração de novas avaliações (anuais, às vezes trimestrais) foi mais frequentemente privilegiada do que a elaboração de estratégias de respostas às dificuldades.

DISCUSSÃO

O exame desses dois casos particularmente contrastantes é instrutivo por mais de uma razão. Um e outro colocam, de maneira evidente, a questão do estatuto das avaliações nos sistemas educativos. Eles permitem compreender como a preocupação legítima de avaliar as políticas públicas pode ser facilmente desviada de suas finalidades.

No campo educativo, é evidente que a avaliação pode ser compreendida de múltiplas formas e que a oposição tradicional entre avaliação formativa e avaliação somativa não é suficiente para resolver a complexidade da questão.

Na medicina, a problemática central se tornou aquela que opõe o dever de colocar em prática a terapia consensual (dever em relação ao qual se pode pedir satisfação ao médico) ao dever de resultados (em relação ao qual a responsabilidade do médico é, por enquanto, raramente considerada).

Em educação, parece que os dois modelos se opõem. Por um lado, há aquele especialmente relativo ao 2º grau, em que o dever de resultados é raramente solicitado, uma vez que se considera que cabe ao aluno contribuir com sua parte de investimento no processo que o conduzirá ao sucesso. O resultado disso é que o professor, como o médico, só é responsável pela escolha da melhor estratégia didática. Por outro lado, no

que se refere ao 1º grau, cada vez mais se espera que a escola assuma totalmente a exigência de eficácia. Na verdade, pensa-se que cada indivíduo dispõe de uma capacidade de alfabetização quase “natural”, assimilando-se tranqüilamente *literacy* e linguagem.²⁴ Todo ser humano poderia, assim, ter acesso à leitura e à escrita, como tem acesso à linguagem, desde que, tanto num caso como no outro, o meio da criança seja favorável ao desenvolvimento de suas competências.

Como mostram os grandes debates públicos sobre alfabetização ou sobre iletrismo, isto leva a novas exigências de avaliação. Não se espera mais da escola que ela desenvolva os melhores métodos (consenso didático), mas, simplesmente, que seja bem-sucedida. Para a escola primária, este sucesso se mede, cada vez mais, conforme um critério único: a alfabetização dos alunos. No entanto, esta alfabetização também não é isenta de problemas específicos.

Na verdade, o nível de alfabetização exigido depende, evidentemente, das políticas escolares desenvolvidas. Nos sistemas educativos avançados, por exemplo, a recente passagem de um ensino de massas curto (centrado no primeiro grau) a um ensino de massas longo (centrado no encadeamento da escola primária/escola média) modificou profundamente a questão da alfabetização. Para garantir que todos os alunos (e não somente aqueles de famílias de elevado capital cultural) cheguem ao ensino secundário, fundado na transmissão maciçamente escrita dos conhecimentos, é necessário que, desde a escola primária, sejam desenvolvidas competências de leitura e de escrita autônomas, que jamais seriam exigidas no ensino primário curto. Portanto, a escola deve prestar contas de sua capacidade de dotar todos os alunos que dela saem de competências de leitura, que lhes permitam tirar proveito do ensino secundário.

Vemos, assim, como as políticas públicas de educação não podem deixar de lado dispositivos de avaliação da eficácia das escolas (e, portanto, dos professores que nela trabalham). Em todo lugar, esses dispositivos tendem a substituir aqueles que, até então, objetivavam ou estabelecer a conformidade às regras em vigor ou julgar a pertinência dos meios implementados. Isto resulta em uma quebra dos consensos e em certo

afastamento dos especialistas de avaliação no que se refere ao controle das didáticas. Evidentemente, isto reforça a autonomia das escolas e a dos professores.

Esta orientação geral cria, contudo, novas dificuldades que os diferentes sistemas educativos ainda não conseguiram resolver por completo. Uma delas é a contradição, cada vez mais freqüente, entre o uso interno das avaliações (para pilotar o sistema, para melhorar os resultados individuais etc.) e o uso que a imprensa faz delas sobre a opinião pública. A preocupação legítima dos jornalistas em denunciar as insuficiências das políticas públicas, bem como sua pouca diligência em mostrar seu sucesso, conduz as instituições responsáveis por essas políticas a desconfiar das avaliações e a tentar limitar seu impacto. Este é um ponto que deve ser debatido entre os profissionais da informação. O mesmo ocorre com os operadores (ONGs, por exemplo), em relação aos organismos que financiam suas ações, que esperam, com freqüência, um retorno imediato do investimento.

Atualmente, os sistemas educativos não podem mais dispensar a avaliação. Ela funciona em todos os níveis do dispositivo escolar, permitindo, por si só, uma regulação permanente das ações. É preciso também colocá-la a serviço da consolidação da real eficácia do investimento educativo. Isto supõe que a avaliação intervenha menos sobre os meios do que sobre os efeitos, além de que, definitivamente, não contribua para restringir a iniciativa profissional dos professores (fechando-os em modelos pré-determinados), mas para restituir-lhes a responsabilidade por suas iniciativas, controlando, em última instância, sua eficácia educativa.

NOTAS

1 A descrição deste projeto por uma pessoa estrangeira e exterior ao Cedac, mas participante de sua avaliação desde o início, corre o risco de apresentar vieses importantes. Trata-se, aqui, então, do projeto como é percebido por um avaliador externo, e não tal como ele foi concebido e implementado.

2 *Programa Escola que Vale*, sem data, p.2. (Mimeogr.)

3 Seguimos, aqui, o programa de trabalho tal qual foi definido nos documentos *Programa Escola que Vale*, sem data, 9p. (Mimeogr.); *Etapas realizadas 1999 e Documento de Logística 2000*, São Paulo: Cedac, janeiro 2000, 76p.

4 Marabá e Parauapebas (Pará), São Luís e Açailândia (Maranhão), João Neiva (Espírito Santo), Caldas Altas (Minas Gerais).

5 Equipe dirigida por Maria do Carmo Brant de Carvalho (Programa de Pós-Graduação em Serviço Social da PUC-SP e Cenpec).

6 Privilegiamos aqui os documentos produzidos pela coordenadora responsável pelo Estado do Maranhão, cujo trabalho pudemos acompanhar *in loco*.

7 Lembremos que o nível IV, imediatamente superior, é aquele do *baccalauréat*, isto é, o exame que sanciona o fim dos estudos secundários e dá acesso à universidade.

- 8 Seria preciso inventar, em 1990, um salário mínimo de inserção — RMI — para dar algum recurso a todos aqueles cuja qualificação era insuficiente, sendo excluído do mercado de trabalho.
- 9 O ciclo I (ciclo das primeiras aprendizagens) é constituído pelos três primeiros anos da escola maternal (crianças de 2 a 4 anos), o ciclo II (ciclo das aprendizagens fundamentais) é constituído pelo último ano da escola maternal e os dois primeiros anos da escola elementar (crianças de 5 a 7 anos); o ciclo III (ciclo das aprendizagens continuadas) é constituído pelos três últimos anos da escola elementar (crianças de 8 a 10 anos). Depois disso, os alunos passam para o sistema secundário (quatro anos no colégio, seguidos de três anos no liceu). Somente a escola elementar e o colégio fazem parte do ensino obrigatório. Ver: Ministère de l'Éducation Nationale, Direction des Écoles. *Les Cycles à l'école primaire*. Paris: CNDP/Savoir Livre, 1990.
- Manteremos os termos "colégio" e "liceu", quando citados pelo autor, pois delimitam etapas da escolaridade na educação francesa. Grosso modo, elas são correspondentes da 5ª à 8ª série e ao Ensino Médio, respectivamente. (N. T.)
- 10 Ministère de l'Éducation Nationale, Direction des Écoles. *Le projet d'école*. Paris: CNDP/Savoir Livre, 1990.
- 11 A avaliação se dá durante a primeira semana de escolarização no ciclo superior, para evitar sobrecarregar os professores que foram responsáveis pelo ciclo que acabou de finalizar. Os resultados das avaliações e sua análise são publicados a cada ano na revista *Éducation et Formation*.
- 12 O inspetor de Educação Nacional é recrutado mediante análise de documentação, entre os professores do 1º ou 2º Grau (portanto, tendo realizado pelo menos três anos de estudos universitários), que tiveram uma experiência profissional como docentes (ao menos cinco anos, mas, na realidade, em média dez anos). Eles recebem, ainda, uma formação complementar de um ano.
- 13 CAFMF (Certificado de Aptidão às Funções de Professor-Formador).
- 14 Na França, o recorte administrativo comporta a comuna (administrada por um presidente da Câmara Municipal e um Conselho Municipal, eleitos), o cantão (sem administração própria, mas representado por um conselheiro-geral eleito), o departamento (administrado por um conselheiro-geral e por seu presidente, que são eleitos e controlados por um governante — *préfet* — nomeado pela administração central).
- 15 Em francês, o termo "pedagogia frontal" se refere, especificamente, a um modo de comunicação entre alunos e professores, no qual o professor se coloca "frente ao aluno", e não "com ele". Em português, o termo poderia se aproximar de "pedagogia tradicional", que, no entanto, é mais abrangente. (N. T.)
- 16 Na França, são as zonas periurbanas, com grandes edifícios a módicos alugueis, chamadas *banlieues*, onde se concentram as famílias sem emprego ou de baixa renda, que são, geralmente, de origem africana ou norte-africana.
- 17 Na França, cada ministro dispõe de uma inspeção geral, composta de altos funcionários e encarregada de missões de avaliação das políticas públicas de sua administração. Os relatórios feitos a pedido do ministro lhe são entregues, sendo ele quem decide torná-los públicos ou não. Esta avaliação é diferente daquela feita pelo Tribunal de Contas, que funciona autonomamente, cujos relatórios são sempre públicos (seu controle concerne apenas ao bom emprego do orçamento público do Estado e aos fundos alocados em estabelecimentos ou organismos públicos).
- 18 MOISAN, Catherine, SIMON, Jacky. *Les ZEP em France: Rapport au ministre de l'Éducation Nationale*. Paris: Documentation Française, 1996.
- 19 Este indicador corrigido foi primeiramente desenvolvido para avaliar os resultados obtidos pelos liceus no *baccalauréat*, a fim de se opor às listas dos laureados, regularmente publicada pela imprensa, significativos apenas da composição social da população escolarizada. Outro indicador permitia calcular a taxa real de alunos cujos resultados positivos, ao longo de toda a sua escolaridade, podiam ser efetivamente atribuídos ao estabelecimento escolar frequentado (a fim de evitar que a reputação de um liceu se estabelecesse sobre sua capacidade de triar os alunos desde as séries iniciais, conservando apenas os melhores).
- 20 As redes de ajuda especializada para os alunos em dificuldade (Rased) são dispositivos que envolvem um psicólogo escolar e dois professores especializados, responsáveis pelas dificuldades de aprendizagem; eles intervêm a pedido dos professores. Existe, em média, uma rede de ajuda para cada 2 mil a 3 mil alunos.
- 21 Por exemplo, em língua materna, para o nível de fim de ciclo II, há um escore global para língua materna, um escore para leitura, produção de textos, conhecimentos dos instrumentos da língua.
- 22 A posição é estabelecida se o aluno for bem-sucedido em, pelo menos, 75% dos itens definidos (*a priori*) como características do nível em questão.
- 23 Circular reforçada por um novo texto editado em novembro de 1999.
- 24 Sobre este debate, ver CHARTIER, Anne-Marie, HÉBRARD, Jean, *Discours sur la lecture, 1880-2000*. Paris: Fayard, BPI-Centre Georges Pompidou, 2000.

* Texto original em francês, traduzido para o português por Beatriz Penteado.

Elba Siqueira de Sá Barretto
Regina Pahim Pinto

Avaliação de
programas
educacionais:
indagações
metodológicas
e disseminação
de resultados*

ELBA SIQUEIRA DE SÁ BARRETTO é licenciada em Pedagogia e doutora em Sociologia pela Universidade de São Paulo. Professora da Faculdade de Educação da USP, atua também como pesquisadora da Fundação Carlos Chagas e como editora responsável da revista *Cadernos de Pesquisa*.

REGINA PAHIM PINTO é mestre em Sociologia e doutora em Antropologia pela Universidade de São Paulo. É pesquisadora senior do Departamento de Pesquisas Educacionais da Fundação Carlos Chagas e editora-executiva da revista *Cadernos de Pesquisa*.

Os programas educacionais, que têm como alvo o ensino básico e que estão a cargo de administrações públicas, não costumam ser avaliados; a regra entre nós tem sido a de fazer suceder uma série de medidas, reformas, inovações, sem que se disponha de elementos suficientes para uma apreciação mais isenta e fundamentada de seus resultados e repercussões na população.

A despeito da tradição incipiente na área, a avaliação intencional de algumas das ações governamentais no campo da educação tem se intensificado, sobretudo a partir do processo de transição democrática nos anos 80, quando as próprias administrações passaram a se mostrar mais sensíveis e permeáveis a esse tipo de apreciação.

Entre as avaliações, podem-se encontrar, em geral, aquelas realizadas a pedido do próprio poder público ou sob a forma de pesquisa acadêmica, feitas por iniciativa dos investigadores. As primeiras caracterizam-se por uma relação mais formal entre clientes institucionais e avaliador, tendo em vista explicitamente a busca de soluções e a tomada de decisões. As pesquisas de caráter acadêmico são realizadas independentemente dos agentes ou dos órgãos promotores da política; buscam, com prioridade, o estabelecimento de conclusões de caráter mais amplo,

podendo levar à identificação de fenômenos recorrentes e princípios mais generalizáveis (Fletcher, 1995; Vianna, 1997).

No livro *Estado da arte: avaliação na educação básica*, em que procedemos a um levantamento dos artigos encontrados em dez das principais publicações científicas da área no período de 1990 a 1998, dentre os 217 textos arrolados, um grupo deles foi categorizado como “avaliação de políticas educacionais” (Barretto & Pinto, 2000). É sobre alguns desses estudos que desenvolveremos considerações neste trabalho.

Em primeiro lugar, cumpre assinalar que há uma série de dificuldades conceituais para delimitar o campo de análise das políticas públicas, bem como o de sua avaliação propriamente dita, visto que os referentes teóricos que lhes servem de base ainda não estão fortemente consolidados. Os estudos que se reportam a análises de políticas reguladas, formuladas e, preponderante ou exclusivamente, executadas pelo Estado procuram relacionar uma série de resultados a determinadas configurações políticas, e, pelo fato de possuírem necessariamente uma conotação valorativa, pode-se-lhes imputar, *grosso modo*, um caráter avaliativo.

Não obstante, além desses trabalhos, há estudos cuja pretensão é estabelecer uma relação mais precisa entre os resultados alcançados pelas políticas e determinados desenhos institucionais (Arretche, 1999). Estes são realizados com a intenção e a sistematicidade próprias do campo da avaliação, visando verificar em que medida as metas propostas por uma política estão sendo atingidas, de que maneira e a que custo.

A avaliação de políticas sociais, nesse sentido mais restrito (e dentre elas as que dizem respeito à educação), tem, na literatura especializada, frequentemente se restringido ou mesmo se confundido com a avaliação de programas, que constituem um elenco de ações ordenadas com propósitos definidos e mais claramente delimitados dentro de uma determinada política. Nesse caso, além de verificar se os objetivos dos programas estão sendo alcançados, as avaliações se prestam a examinar os processos pelos quais eles são implantados e os efeitos que lhes podem ser atribuídos.

As avaliações de programas educacionais, entre nós, têm como referência, de um lado, a longa tradição de avaliação na própria área e, de outro, as contribuições mais recentes da ciência política. Na primeira tradição, a

avaliação educacional no Brasil, caudatária do pensamento anglo-saxônico e, em particular, dos estudos norte-americanos, acompanhou a transição do conceito de avaliação como medida — voltada para os resultados da aprendizagem do aluno e a cargo de psicólogos e psicopedagogos — para uma perspectiva mais ampla. Esta passou a focalizar como objeto de estudo os programas educacionais e o seu contexto, e a ocupar-se de projetos de larga escala, tendo, como avaliadores, profissionais do campo das ciências sociais.

Tal mudança de ênfase deu-se, no exterior, em decorrência de circunstâncias históricas, por vezes muito bem definidas. Nos Estados Unidos, por exemplo, a avaliação de programas, ou pesquisa de avaliação, foi incrementada no contexto de “guerra à pobreza”, como resultado de dispositivo legal de 1965, que obrigava a avaliação de todos os projetos financiados pelo poder público voltados para o ensino básico. O objetivo era fornecer um retorno aos gestores e prestar contas à população acerca dos recursos aí envolvidos, o que levou a grandes esforços nessa direção. A prática generalizada da avaliação mostrou, no entanto, que muitas das abordagens não eram adequadas (Vianna, 1995).

Nessa trajetória, vários autores, bastante conhecidos no Brasil, darão nova dimensão metodológica à avaliação, que se torna um campo profissional definido após 1970. Os referenciais adotados à época apontam para a necessidade de se levar em conta o grande número de variáveis presentes nos programas educacionais e de se atuar no campo em que os atores sociais naturalmente se movem, buscando-se superar *designs* de “controle” amplamente utilizados na pesquisa experimental ou quase-experimental de base psicológica.

Essas transformações transitam pela preocupação da avaliação voltada para a tomada de decisões, como em Stufflebeam; pela avaliação responsiva de Stake, a qual se orienta mais para as atividades do programa do que para as intenções por ele declaradas; pelo flexionamento feito por Scriven, que passa da defesa de uma avaliação centrada em objetivos à necessidade de que não sejam separados os fins dos efeitos secundários e sociais de um programa. Os enfoques caminham também pela avaliação iluminativa de Parlett e Hamilton, proposta na Inglaterra; pela observação participante e pela pesquisa-ação.

Assim, de modo progressivo, os esquemas excessivamente formais dos modelos positivistas de avaliação de programas passam a ser animados pela consideração dos diferentes sujeitos sociais envolvidos nas ações que se buscam avaliar, com a incorporação das perspectivas fenomenológica e antropológica. Nesse percurso, ganham ênfase as análises de corte qualitativo e se observa a transformação da postura neutra e distante do avaliador para a disposição de aceitar a negociação de sentidos entre avaliador e avaliado. Algumas abordagens, em última instância, manifestam-se ainda a favor da apropriação total do sentido e do controle do saber pelos grupos que, no dizer de Depresbiteris (1989), expressam uma aprendizagem coletiva na tomada de decisões e no comprometimento com a ação grupal, como no caso da pesquisa participante.

A preferência pelos estudos qualitativos entre os pesquisadores brasileiros do campo da educação termina, contudo, por diluir a atenção aos aspectos de cobertura, financiamento e desenho institucional dos programas, cruciais na sua avaliação.

A segunda vertente de influência, a das ciências políticas, contribui para recuperar a dimensão macro da avaliação das políticas e programas educacionais entre nós, trazendo aportes significativos ao setor nas duas últimas décadas. A pesquisa de avaliação derivada desse campo recorre a outros autores, e embora ambas tradições lancem mão do instrumental de análise comum às ciências humanas, e à pesquisa social, em particular, pode-se dizer que o diálogo entre elas mal se iniciou.

A avaliação das diferentes políticas sociais pelo ângulo das ciências políticas é relativamente recente, estando ainda em processo de constituição de um campo estruturado de estudos, conforme assinalam Figueiredo & Figueiredo (1986). Na avaliação de programas feita por pesquisadores da área, é ainda freqüente o viés neutralista e comportamentalista que também marcou o campo educacional, os estudos usualmente se limitando a aferir em que medida as metas de determinado programa estão ou não sendo alcançadas, sem entrar no mérito propriamente político do que está sendo avaliado.

A despeito das limitações apontadas, a reflexão no âmbito das ciências políticas refere a discussão da avaliação a critérios de equidade e a

princípios de justiça política e social, lançando novos olhares sobre os processos de formulação da agenda governamental, da implementação de programas e da apreciação de seus impactos na população.

Antes de fazer comentários mais específicos sobre o material selecionado no *Estado da arte*, cumpre-nos prestar alguns esclarecimentos sobre a opção pela consulta aos periódicos e suas implicações. As revistas acadêmicas de educação contêm parte bastante representativa do material difundido sobre a variedade da produção científica na área. Por serem um dos veículos de circulação mais ágil entre pesquisadores, educadores e estudiosos do assunto, constituem instrumento privilegiado para o diálogo e o intercâmbio de idéias entre um público em que estão concentrados os formadores de opinião, os quais contribuem para formular e alimentar o ideário educacional no país. Essas publicações acabam se tornando um importante termômetro do peso da produção teórica sobre os temas concernentes ao campo. Ao nos reportarmos a elas convém frisar, portanto, que não estamos falando do *locus* de produção da avaliação de políticas ou programas educacionais, mas sim da perspectiva de sua disseminação.

A propósito, o primeiro destaque a ser feito quanto aos textos selecionados diz respeito ao espaço exíguo que a avaliação de políticas educacionais ocupa no conjunto dos artigos sobre avaliação nos periódicos: apenas 20% deles compõem a categoria.

AVALIAÇÃO DA AGENDA E DAS PROPOSTAS DE GOVERNO

Um dos aspectos das políticas públicas ainda muito pouco explorado nas pesquisas é o que se refere à avaliação da agenda e das propostas de governo. No caso da agenda, costumam contar, na sua formulação, as mobilizações sociais de diferente natureza, as negociações político-partidárias, as demandas de organismos internacionais, as negociações internas dos segmentos que pleiteiam o poder. À agenda, segue-se a formulação de políticas: são delineados os planos, programas e projetos que lhe darão materialidade, estabelecidas as prioridades e prevista a alocação de recursos.

Dentre os textos selecionados, a rigor não foram encontradas propriamente pesquisas de avaliação da agenda ou das propostas de governo, no sentido que permite uma apreensão mais profunda do processo pelo qual algumas questões postas no âmbito da sociedade se convertem em temas a serem tratados pelas políticas públicas, ou na acepção que possibilita identificar os segmentos que colocam certas questões e aqueles cujas demandas são incorporadas. Essa abordagem possibilitaria detectar conflitos, graus relativos de autonomia e de poder de atores sociais, dentro ou fora do Estado.

Quanto aos artigos em tela, trata-se antes de ensaios que tangenciam o objeto de estudo, provenientes, muitas vezes, de uma espécie de avaliação de contexto.

Vários dos artigos enquadrados nesse grupo representam falas do discurso oficial sobre as políticas do setor. Seus autores são, via de regra, dirigentes ou ex-dirigentes de Secretarias de Estado e do Ministério da Educação, que apresentam suas agendas políticas e discutem planos, programas ou projetos de governo. Com base na análise de informações sistematizadas, que contribuem para caracterizar e dimensionar em larga escala os problemas educacionais dos sistemas ou redes de ensino, os textos buscam justificar as prioridades estabelecidas nas respectivas agendas e nos programas, chegando, por vezes, a discutir estratégias para a sua execução ou para o seu redirecionamento. Trata-se, geralmente, de informações provenientes de censos; de indicadores de fluxo e eficiência; do exame de tendências detectadas na oferta do ensino em diferentes níveis e modalidades e no comportamento da demanda; da discussão relativa ao estilo de gestão do sistema público. Termos como *compromisso* com determinados segmentos — população usuária da escola ou professores, por exemplo —, *qualidade do ensino*, *equidade*, *autonomia* e *descentralização* povoam esses discursos, que recorrem a avaliações diagnósticas do sistema educacional, buscando a legitimação política das ações governamentais, especialmente por meio da racionalidade técnico-científica.

Argumentos que expressam a busca de legitimação de natureza mais essencialmente política das principais estratégias e ações adotadas na

gestão desses aparatos públicos são mais escassos. Uns poucos textos, escritos por especialistas não diretamente ligados às administrações públicas, buscam ainda, em tom opinativo, revisar a pauta governamental, seja procurando fazer um rápido balanço dos principais problemas enfrentados pela política educacional brasileira — analisando as medidas e ações do Estado quanto à adoção de regulamentações, de políticas redistributivas e do sistema de avaliação –, seja tentando examinar a avaliação da educação no Mercosul, à luz das relações entre os países membros e os organismos multilaterais, a evocar uma nova ética de cooperação internacional.

AVALIAÇÃO DE PROGRAMAS EDUCACIONAIS

O grupo mais numeroso de artigos arrolados na categoria focaliza os programas educacionais e inclui pouco menos da metade dos 42 títulos. Os artigos contêm análises que podem constituir-se em avaliações de programas como um todo ou de aspectos específicos. Alguns desses textos na verdade não fazem, a rigor, uma avaliação *stricto sensu*, quer porque se trate de ensaios que apreciam certos programas adotando, como no caso do exame da agenda, um estilo eminentemente opinativo, quer porque apenas os mencionem para subsidiar reflexões suscitadas pelo seu estudo. Nesses casos, as informações fornecidas ao leitor não são suficientes para que se possa ter uma idéia mais completa dos propósitos do programa e do seu delineamento. Há, por fim, os que se reportam a questões de caráter teórico-metodológico referentes à avaliação.

Nos anos mais recentes, em particular na última década, as avaliações sob demanda do poder público têm aumentado em decorrência de exigências de organismos multilaterais, cuja importância vem sendo crescente na definição e financiamento das políticas do setor. Não obstante, foram poucos os artigos que puderam ser enquadrados como avaliações propriamente ditas de programas, o que leva a supor que essa modalidade de estudo não chega facilmente ao conhecimento do grande público leitor da área da educação.

Investigações dessa natureza são realizadas sobretudo por agências de planejamento e/ou instituições de pesquisa, visto que costumam constituir estudos de grande porte e são onerosas. Elas exigem, freqüentemente, a composição de numerosas equipes de trabalho e razoável infra-estrutura que dê suporte à sua operacionalização; não raro, envolvem procedimentos altamente especializados de análise. Esses estudos muitas vezes deixam de ser divulgados em veículos de ampla circulação, ficando restritos a relatórios extensos, apenas entregues aos órgãos que os requisitaram.

As avaliações de programas de cunho acadêmico também não são abundantes na bibliografia consultada.

De modo geral, os artigos que se reportam à avaliação de programas nos periódicos têm a intenção de estender certas reflexões suscitadas pelas análises a um público maior, dada a natureza do veículo escolhido. Por isso, nem sempre seus autores explicitam a metodologia empregada, os procedimentos utilizados ou os modelos teóricos que subsidiaram a análise, o que, em muitos casos, dificulta sobremaneira o esclarecimento, para o leitor, do desenho utilizado na avaliação. Em outros casos, ainda, a seleção de aspectos parciais dos programas para análise impede mesmo a compreensão da sua lógica própria.

O exame detido dos textos arrolados como avaliações de programas suscita reflexões que serão registradas a seguir.

DOIS ESTUDOS EMBLEMÁTICOS SOBRE O CICLO BÁSICO

No contexto das avaliações que se reportam ao programa como um todo, ainda que possam privilegiar determinadas regiões ou localidades, dois textos focalizam o ciclo básico de alfabetização.¹ Sobre os ciclos de alfabetização implantados nas redes escolares de vários estados a partir do primeiro lustro dos anos 80 — os quais possibilitam ao aluno cursar os dois anos iniciais do ensino fundamental sem solução de continuidade — já foram realizados muitos estudos, particularmente nas universidades dos estados onde as medidas foram intro-

duzidas. Não obstante, a maior parte das pesquisas, realizadas como exigência para as dissertações de mestrado e teses de doutoramento, costuma ter circulação limitada e divulgação muito restrita. Por coincidência, os dois textos que se ocupam do tema nos periódicos selecionados apresentam abordagens que são bastante representativas das condições institucionais em que elas são produzidas.

O primeiro desses estudos incide sobre ações envolvidas no ciclo básico em dois estados, Minas Gerais e São Paulo, tratando dos aspectos políticos ligados à sua formulação e implementação, bem como da sua efetividade como medida que se propõe a melhorar o fluxo dos alunos na escola básica. Partindo de um delineamento experimental de pesquisa, aplica o modelo de fluxo desenvolvido por Fletcher e Ribeiro, utilizando dados dos censos educacionais do Ministério da Educação e das PNADs (Pesquisa Nacional de Amostra por Domicílio), referentes às matrículas de alunos do ciclo básico dos referidos estados e de alunos das primeiras e segundas séries da rede escolar do Estado do Rio Grande do Sul, tomado como grupo de controle. Ao acompanhar a trajetória escolar de coortes de alunos entre 1981 e 1989, constata que a introdução do ciclo básico não teria alterado essencialmente o fluxo escolar nos dois estados, passando a discutir as condições culturais e os condicionantes escolares que estariam contribuindo para o insucesso das medidas (Neubauer da Silva & Davis, 1993).

Embora não se trate de investigação feita sob demanda do poder público, o fato dela ter sido realizada em instituição especializada em pesquisa facilitou o acesso das pesquisadoras não só a processamentos especiais das estatísticas educacionais, como a uma metodologia sofisticada de análise de dados censitários. Isso permitiu a realização de um estudo de larga escala, que não teria condições de ser feito na maior parte das pesquisas realizadas no âmbito dos cursos de pós-graduação em educação no país,² em especial por falta de acesso à metodologia empregada. Apenas nos últimos anos começam a consolidar-se, em algumas universidades brasileiras, núcleos de estudos sobre avaliação, os quais, trabalhando numa perspectiva interdisciplinar que alia sobretudo a formação de educadores à de matemáticos e estatísticos, teriam condições de dar alento, entre outras, a investigações dessa natureza.

Já o segundo estudo publicado sobre o ciclo básico retrata bem o tipo de abordagem mais freqüente que a universidade costuma utilizar para análise de medidas de governo que levam a alterações no funcionamento das redes escolares. O artigo tece considerações sobre o programa, com base em pesquisa realizada em quatro escolas da rede estadual paulista, com o objetivo de avaliar o êxito da medida e o nível de adesão a ela e, nesse sentido, apontar aspectos a serem apoiados ou reformulados. Recorre a estudos de caso que procuram investigar em que medida o clima e a organização das escolas favorecem a efetivação das mudanças introduzidas, a aprendizagem básica da leitura e escrita pelos alunos e a alteração dos índices de retenção. Procura investigar a postura dos professores quanto à adesão ao programa, com base nas mudanças da prática pedagógica e na sua opinião sobre as inovações, e verificar o desempenho dos alunos em relação ao currículo, procurando detectar a presença ou não dos fatores que garantiriam o sucesso da implantação das medidas propostas (Silva, 1991).

Ainda que por caminhos diferentes, as duas avaliações procuram discutir o nível de aceitação das medidas adotadas por parte dos agentes educacionais, seu grau de efetividade e os fatores favoráveis ou desfavoráveis ao desenvolvimento do regime de ciclos. Quanto à metodologia utilizada, o primeiro tipo de pesquisa possibilita que se tenha uma idéia do impacto do programa nas redes escolares relativamente aos efeitos surtidos no conjunto da população escolar (ainda que os resultados específicos a que chegou a investigação mencionada sejam passíveis de contestação). O segundo oferece as pistas para o entendimento da construção e reconstrução dos significados que os diferentes agentes educacionais e os usuários da escola atribuem às modificações pretendidas, indica os processos por meio dos quais são redefinidas as mudanças por aqueles que nelas estão diretamente envolvidos, as condições de funcionamento escolar em que isso ocorre, assim como o grau de satisfação dos atores sociais com as medidas. Em última instância, os dois estudos exploram dimensões importantes e que se complementam para o entendimento do ciclo.

Dentro da perspectiva de curto prazo que parece mover cada vez mais as decisões na área das políticas de educação, e que privilegia as infor-

mações que possam ser mais facilmente utilizadas pelos dirigentes governamentais, a tendência tem sido a de preferir dados da natureza dos produzidos pelo primeiro tipo de pesquisa, ou seja, os que dão idéia da extensão, abrangência e impacto dos fenômenos. É também usual, no afã de informar os “tomadores” de decisão, que a literatura que circula sobre as políticas públicas recorra a exemplos de estudos realizados em outros países, invariavelmente descontextualizados, para apoiar iniciativas de governo, ao mesmo tempo que se desconsideram, como não relevantes, as informações relativas aos processos de implementação de mudanças, o que leva, por tabela, a desqualificar a produção acadêmica das universidades na área da educação.

Em que pesem as fragilidades metodológicas de muitas das pesquisas desenvolvidas no âmbito universitário, é preciso convir que, embora muitas vezes dispersas, certas informações de grande relevância têm sido menosprezadas na formulação e no que concerne aos processos de implementação de programas e das políticas educacionais. Isso não significa deixar de reconhecer as grandes lacunas na pesquisa educacional, particularmente no que se refere à escassez de estudos macro-sociais.

A VOCAÇÃO POPULISTA DE ALGUNS PROGRAMAS EDUCACIONAIS

Na primeira metade dos anos 80, o programa de um dos partidos de oposição que assumiu o governo do Estado do Rio de Janeiro (PDT) teve como carro-chefe a criação de escolas de tempo integral para o ensino fundamental. Visava com elas ao atendimento das crianças das camadas majoritárias, mediante um projeto educacional audacioso, que se propunha a oferecer-lhes um currículo voltado para um processo mais abrangente de socialização, preocupado com o resgate da cultura popular e com a assistência às necessidades básicas dos alunos. Os CIEPs — Centros Integrados de Educação Pública —, nome que receberam essas escolas, compreendiam até mesmo o abrigo, na própria escola, de um pequeno número dos alunos de famílias mais carentes, os quais passavam a lá residir durante os dias de aula, voltando para casa apenas nos fins de semana.

O programa teve grande repercussão política porque, de certo modo, propunha-se a dar uma resposta ao mal-estar disseminado na sociedade pela situação de precariedade e desamparo de grande contingente de crianças pobres no país, amplamente divulgada pela grande imprensa no Brasil e no exterior, denunciada por organismos internacionais e particularmente associada à violência relacionada ao trato com os meninos de rua.

Sobre os CIEPs, embora se saiba que tenham sido objeto de avaliações sistemáticas, nas fontes consultadas apenas foram encontrados dois estudos, os quais fazem um recorte muito limitado na análise, fornecendo elementos pouco conclusivos para a sua avaliação. Um deles compara apreciações de educadores e usuários dos CIEPs com os de escolas de tempo parcial, sugerindo que há necessidade de revisão de certos pressupostos acerca da população-alvo e de suas possibilidades de emancipação (Zaluar & Leal, 1996). O outro discute o emprego da abordagem etnográfica na avaliação de projetos inovadores, com base em dois estudos de caso: um projeto de inovação educacional, o CIEP, e um projeto de inovação social, o projeto Sleep-in, de um Centro de Atendimento ao Usuário de Drogas, em Paris, este último tendo sido objeto de pesquisa avaliativa por solicitação do Ministère d’Affaires Sociales (Leonardos & Macedo, 1996).

Na esteira das repercussões políticas dos CIEPs, foi criado o Programa de Formação Integral da Criança — Profic, na rede estadual paulista, em 1986, com menor porte que o do Rio de Janeiro. Destinado a escolares de baixa renda, que deveriam receber atendimento suplementar na escola após a jornada de meio período, propunha-se a lhes oferecer uma “educação integral”, que visava mais do que à instrução, ampliando assim a função da escola de primeiro grau.

Os dados coletados e analisados por Ferretti et al. (1991), com base em amplo mapeamento das escolas que implantaram o Programa entre 1986 e 1988, e o estudo em profundidade realizado em duas dessas unidades escolares indicam que o Profic assumiu características centralistas, tanto em sua formulação como na implementação. Além disso, o processo de implantação foi precário, criando vários problemas para as escolas que o adotaram, ficando o índice de adesão a ele, pela rede escolar, muito aquém do esperado.

A pesquisa contextualiza o programa, discute criticamente a proposta, seus objetivos, os programas que a inspiraram, os pressupostos em que se baseia, procurando dar conta da abrangência e das diferenças internas na sua implementação. Aponta para o caráter psicologizante e assistencialista subjacente à abordagem dos problemas enfrentados pelas populações a que se destina, postura que teria sido influenciada pela maneira como os programas internacionais de saúde, que lhe serviram de inspiração, concebem as populações-alvo: subculturas homogêneas e marginalizadas, marcadas por carências psicológicas, dificuldade de abandonar valores tradicionais. Daí a proposta de integrá-las na sociedade urbano-industrial por meio da “participação comunitária”, o que lhes possibilitaria o acesso a cuidados médicos e o desenvolvimento de novos comportamentos sociais diante da exclusão.

PROGRAMAS DE AGRUPAMENTO DE ESCOLAS ISOLADAS NA ZONA RURAL

Há outros dois programas que foram analisados por Ramos (1991) e Vasconcellos (1993). São praticamente idênticos quanto às características e objetivos. Trata-se de Escolas Rurais Consolidadas, implantado no Estado do Paraná, entre 1977 e 1986, e Agrupamento de Escolas Rurais, implantado no Estado de São Paulo, em 1989. Consistem basicamente na reunião de várias escolas isoladas e multisseriadas em uma escola de primeiro grau maior, agrupada, com melhores condições físicas e pedagógicas. As análises mostram que ambos melhoraram as condições de escolaridade na zona rural no que se refere a oferta e utilização de equipamentos e vagas, ampliação dos anos de escolaridade, oferecimento de condições mais adequadas de trabalho e de aperfeiçoamento profissional para os professores, menor absenteísmo dos alunos.

No primeiro caso, as escolas consolidadas foram implantadas em 15% dos municípios paranaenses com ajuda federal. Além de intervenientes de ordem política, que teriam contribuído para o desvirtuamento dos objetivos, o que acarretou a diminuição do impacto, constatou-se que os custos operacionais do programa eram altos, grande a rotatividade dos professores, ao mesmo tempo que os índices de reprovação entre os

alunos praticamente não se alteraram. O dimensionamento inadequado da realidade por parte dos proponentes do Programa, como no caso do Profic, é atribuído à mera transposição de modelos do exterior.

○ artigo sobre a avaliação do Agrupamento de Escolas Rurais, de São Paulo, descreve e situa o processo de implantação do programa no contexto social, econômico, político e administrativo do Estado. Com base na experiência paulista e na de vários outros países que adotaram esse modelo, que são referidas no texto, tal artigo faz considerações sobre as questões mais relevantes suscitadas pelas medidas, como a melhoria do ensino-aprendizagem, os desafios colocados pela relação entre eficiência econômica das escolas e equidade na oferta de serviços, bem como destaca a necessidade de considerar melhor os custos de adaptação/construção de escolas e gastos com o esquema de transporte dos alunos, o que requer um planejamento cuidadoso para a sua operacionalização e uma definição clara acerca das responsabilidades dos gastos e da operação na esfera do Estado e dos municípios envolvidos.

FORMAÇÃO CONTINUADA DE EDUCADORES

Reiterando tendência já diagnosticada em outros estudos sobre a formação docente, os textos que versam sobre este tema debruçam-se sobre questões suscitadas pela formação em serviço. Essa tem sido a tônica das políticas na área, e, talvez como reflexo delas, têm rareado os estudos sobre a formação inicial de professores. Ademais, quando eles existem, tampouco costumam configurar-se como avaliações de programas, uma vez que os cursos têm características perenes.

Embora três artigos tenham focalizado programas de capacitação de profissionais que atuam no sistema público de ensino, o único avaliado de forma sistemática foi o chamado Um Salto para o Futuro. Sobre os demais programas, existem apreciações mais ou menos fundamentadas em algumas evidências empíricas, insuficientes, entretanto, para que se tenha uma idéia relativamente precisa acerca do processo de implementação, dos resultados obtidos e/ou de repercussões na rede escolar.

Um dos textos faz um contraponto entre o Programa de Capacitação para Educadores de Apoio da Secretaria Estadual de Educação de Pernambuco (1987 a 1991) e o Projeto de Capacitação dos Coordenadores Pedagógicos da Secretaria Municipal de Educação de São Paulo (1995 a 1996), destacando as orientações básicas que lhes deram suporte (Silva, 1996). As considerações sobre o Programa de Formação de Capacitação de Recursos Humanos da rede estadual paulista (1984 e 1994) referem-se à atuação das universidades públicas na oferta dos cursos e evocam dados da apreciação feita pelos professores (Martins, 1996).

O texto sobre o Programa Um Salto para o Futuro é resultado de pesquisa de avaliação realizada em 1992 pela Fundação Carlos Chagas, por solicitação do Ministério da Educação. Trata-se de programa de caráter nacional, dirigido aos professores das séries iniciais do primeiro grau, que visa levar o professor a formular um referencial teórico-metodológico adequado e também a ampliar o seu nível de conhecimento geral. Promove capacitação a distância via satélite, combinando recursos multimídia e utilizando transmissão monitorada. Quando submetido à primeira avaliação externa — que é relatada no artigo —, contava para a sua viabilização com o Ministério da Educação, por meio da Secretaria Nacional de Educação Básica (Seneb), a Fundação Roquette Pinto, responsável pela televisão educativa do Ministério e as Secretarias de Estado da Educação, no que diz respeito às ações de campo.

A pesquisa analisa, mediante apreciação de juízes, os pressupostos teórico-metodológicos e o formato do programa, focalizando o material produzido; registra a opinião dos atores sociais envolvidos e atenta para a interação dos cursistas com os agentes educacionais e destes entre si, o que é feito por meio de entrevistas, questionários e observação de telepostos em amostra representativa; verifica o aproveitamento dos cursistas com base em questões abertas sobre o conteúdo dos programas; recolhe dados sobre as condições de implementação. Aponta para aspectos que podem ser melhorados no que diz respeito ao projeto educativo e à exploração dos recursos tecnológicos; encontra resultados positivos quanto à compreensão, pelos professores, dos conceitos básicos veiculados pelo curso; chama a atenção, porém, para o reduzido número da audiência ante as estimativas iniciais de cobertura do

programa, destacando a necessidade de uma análise de custos que leve em conta os recursos envolvidos nas diferentes instâncias governamentais parceiras. Reconhece, por fim, que o programa constitui uma proposta inovadora do ponto de vista educacional, que potencialmente poderá tornar-se importante instrumento de atualização docente (Barretto, 1997).

Estudos dessa natureza costumam apenas situar sumariamente o objeto que se propõem a avaliar. Concentram-se nas enormes tarefas operacionais necessárias à coleta e análise das informações referentes às várias dimensões do programa a serem pesquisadas, seja em amostras grandes, em vista da exigência de representatividade, seja em estudos qualitativos, que analisam determinados aspectos do programa em profundidade. Considerando-se o exíguo tempo que dispõem para apresentar os resultados do trabalho — que deriva de complexos desenhos de investigação —, terminam pecando pela falta de contextualização mais abrangente, por citar outros estudos apenas quando a referência se torna imprescindível, sem que haja condições de discussão mais alentada acerca de referenciais teórico-metodológicos empregados, ou comparação com resultados obtidos em investigações congêneres. Se tais pesquisas ganham em agilidade e prestam-se a subsidiar diretamente a tomada de decisões, o tempo que dispõem para a maturação das análises fica a dever em relação aos estudos de caráter mais acadêmico, o que, de certo modo, as fragiliza no que se refere ao seu poder de explicação.

ESTUDOS SOBRE CUSTO/RESULTADO E CUSTO/BENEFÍCIO

Otras avaliações, apoiadas em metodologia específica, analisam apenas aspectos parciais de determinados programas. As informações sobre o programa, no caso, são bastante genéricas, destinando-se, sobretudo, a situar a questão a ser investigada, de sorte que, por meio delas, não é possível ter uma idéia clara do conjunto de ações nele envolvidas. Dois estudos se desenvolvem nesses termos, com base em dados colhidos em razão de outras avaliações e voltados para a análise de custos, modalidade de avaliação particularmente escassa no país. Um deles refere-se

ao Programa Aceleração de Aprendizagem, destinado a corrigir o fluxo escolar das quatro primeiras séries do ensino fundamental das redes públicas de ensino, mediante a formação de classes de aceleração para atender alunos com defasagem idade-série.

Concebido em 1994, tal programa foi inicialmente implantado em 1995, no Estado do Maranhão, com recursos do Ministério da Educação e apoio técnico do Centro de Ensino Tecnológico de Brasília. Em 1996, foram desenvolvidos e testados materiais didáticos consistentes com a proposta do programa e, em 1997, iniciados dois novos projetos dentro da mesma concepção, um pela Secretaria de Estado da Educação de Minas Gerais e outro pelo Instituto Ayrton Senna/Petrobrás, disseminado por todo o país. Na sua avaliação, destacam-se a economia gerada e a efetividade do programa, com base nos resultados do teste a que foram submetidos alunos das classes de aceleração de aprendizagem.

Os resultados alcançados por alunos, antes considerados multirrepetentes, confirmam a proposta do programa, ou seja, a de que houve aquisição de conhecimentos equivalentes aos do aluno médio de 4ª série, demonstrando que é possível e viável corrigir o fluxo escolar, acabar com a reprovação e erradicar a cultura da repetência no país. O estudo também demonstrou que os resultados estão associados aos investimentos. Comparando-se os três projetos, concluiu-se que aquele em que os insumos foram insuficientes, no caso, o do Estado do Maranhão, obteve piores resultados quanto às séries que o aluno conseguiu saltar (Oliveira, 1998).

O segundo estudo refere-se ao Edurural — Programa de Expansão e Melhoria do Ensino Rural —, implantado, no período de 1981 a 1987, nas zonas rurais de 218 municípios do nordeste, com o objetivo de melhorar as condições do ensino fundamental na região.

Embora o programa tenha abrangido todo o nordeste, sua avaliação e acompanhamento foram realizados em três estados, a saber, Piauí, Pernambuco e Ceará. Recorrendo aos dados coletados durante o período em que durou a avaliação (Gomes Netto et al., 1993), faz-se uma análise de custo-benefício dos insumos educacionais fornecidos pelo programa, com base em estimativas sobre o número de estudantes/ano e tendo

como referência a progressão escolar dos alunos. Os autores chegam à conclusão de que os investimentos em insumos oferecidos às escolas participantes (infra-estrutura, materiais de ensino aprendizagem, qualificação do professor) aumentaram a eficiência do sistema, tendo em vista a melhoria do fluxo dos alunos, o que, por sua vez, produziu um retorno financeiro em razão da economia obtida. Nesse sentido, os avaliadores ponderam que os investimentos se justificam plenamente, pois o custo para se obter um aluno formado foi menor do que os investimentos.

ARTIGOS COM PREOCUPAÇÕES DE CARÁTER METODOLÓGICO

Embora no *Estado da arte* tenha se constatado a predominância de artigos que abordam a avaliação sob o prisma conceitual, tentando elucidar seus pressupostos, finalidades, teorias, modelos, o eixo dessa discussão gira em torno da avaliação da aprendizagem, da escola e de seu contexto institucional, mencionando-se a avaliação de programas educacionais só nas retrospectivas históricas que se remetem às influências do exterior no pensamento brasileiro.

Dentre o conjunto de textos tomados para análise, apenas um autor preocupou-se com a discussão metodológica sob o enfoque das ciências políticas, o qual é muito pouco conhecido entre os educadores. Sousa (1994) procura estabelecer uma relação entre os princípios teóricos que fundamentam a formulação e implementação de políticas públicas e a estruturação de procedimentos avaliativos, com base nos trabalhos de Cerych e Sabatier, pesquisadores que aplicam conceitos de *policy analysis* e de *implementation analysis* a estudos de reformas do ensino superior na Europa. Contudo, a abordagem do artigo tem mais um caráter introdutório do que analítico.

Além desse texto, vale a pena registrar, entre as reflexões de Gatti sobre a experiência de avaliação de desempenho do sistema educacional no nível de primeiro grau adquirida na Fundação Carlos Chagas, aquela realizada no âmbito do Programa Edurural como uma das mais abrangentes na década de 1980 (Gatti, 1994).

Em parceria com a Fundação Cearense de Pesquisa, a pesquisa avalia-

tiva consistiu em estudo longitudinal, desenvolvido ao longo de cinco anos (1981-1985), com dados colhidos em amostra aleatória de 603 escolas nos três estados já mencionados. O estudo incidiu sobre diversas variáveis: condições das escolas, perfil dos professores, impacto de treinamentos, material escolar, merenda escolar, condições das famílias e rendimento escolar dos alunos. Foi ainda complementado por seis estudos de caso, dois em cada um dos estados. A quantidade de análises derivadas do projeto foi muito significativa, sendo disseminada por meio de um número representativo de publicações, sob a forma de livros, artigos em periódicos e teses acadêmicas, tendo resultado em contribuição das mais relevantes para o conhecimento das condições do ensino rural no país.

Quanto aos aspectos metodológicos, a primeira questão a considerar, como lembra Gatti, é que a avaliação deve levar em conta as características do ambiente educacional em que se desenvolverá, “evitando a imposição de modelos nem sempre ajustáveis aos vários aspectos do sistema educacional, ainda que bem estruturados do ponto de vista teórico” (op. cit., p.77). No processo de avaliação, isso ficou evidente, pois um dos problemas enfrentados foi justamente a impossibilidade de se conciliar as exigências do desenho de pesquisa proposto pelo Banco Mundial com a realidade política e cultural da região.

Como a tradição empirista de pesquisa social propõe os modelos experimentais e quase-experimentais para a avaliação do impacto dos programas, os estudos financiados pelas agências multilaterais quase que invariavelmente colocam como exigência a adoção desses modelos, visto que estão interessadas sobretudo na efetividade das ações programadas. A lógica do desenho tem como pressuposto que não basta a constatação de que as mudanças propostas aconteceram para se inferir o sucesso de um programa; é necessário demonstrar que elas não ocorreriam na sua ausência.

No caso do Edurural, a avaliação deveria, portanto, ter um desenho experimental, em que os municípios que recebessem insumos do programa fossem comparados com outros, que deles estivessem excluídos. No decorrer da pesquisa, percebeu-se, entretanto, que o ensino rural recebia

recursos de vários projetos, que se sobrepunham e eram distribuídos pelo conjunto dos municípios segundo critérios variados e sob influência clientelista, fato que inviabilizou a escolha dos grupos de controle.

Ainda comentando problemas enfrentados na avaliação do Edurural, a autora reporta-se às questões decorrentes do emprego de amostras grandes na avaliação educacional, tais como, imprecisões e representatividade de dados estatísticos, tempo demandado para a coleta de dados, custos, além da baixa credibilidade dos estudos amostrais perante a população. No caso do programa, a seleção aleatória de escolas levou à constatação de que muitas eram “fantasmas”, o que comprometeu a representatividade da amostra. Em aplicações posteriores, terminou-se confirmando que muitas escolas não eram perenes, o que obrigou à escolha intencional de escolas equivalentes.

Sobre as dificuldades mencionadas, acrescentaríamos que, onde os aparatos burocrático-institucionais de prestação de serviços ainda estão precariamente instalados, como acontece nas redes municipais de ensino que atendem as áreas rurais do nordeste — as mais pobres do país —, os esquemas de controle concebidos em realidades com alto nível de institucionalização de serviços, e originalmente criados para situações de laboratório, tendem a mostrar grandes fragilidades. Mesmo assim, eles têm sido freqüentemente priorizados em detrimento de outras abordagens que, igualmente, podem trazer contribuições valiosas. Os estudos de caso, introduzidos na avaliação do Edurural após insistência dos pesquisadores brasileiros, trouxeram à tona uma série de questões subjacentes aos problemas detectados no programa, que o uso do modelo experimental não conseguiu esclarecer.

Outros aspectos discutidos no texto de Gatti dizem respeito à necessidade de se considerar, nas pesquisas de avaliação, o preparo do pessoal, considerando-se que determinados estudos exigem uma formação acurada dos que elaboram os instrumentos e dos que os aplicam em campo. Por vezes, em certas regiões e locais, faltam pessoas qualificadas, até mesmo para aplicar instrumentos mais simples de avaliação e registrar seus resultados. Além disso, em certas circunstâncias, é difícil encontrar aplicadores que colem as informações com um mínimo de

confiabilidade, uma vez que a situação de avaliação costuma ser ameaçadora e que as pessoas temem receber sanções em decorrência de resultados que possam advir desses processos de aferição.

A isso, conforme remarca a autora, acrescentem-se os custos de avaliação de programas amplamente disseminados nos sistemas de ensino em um país com as dimensões continentais, como o Brasil, o que os torna por vezes quase inviáveis em termos orçamentários, apesar de tecnicamente bem delineados como projetos de avaliação.

CONCLUSÕES

Em primeiro lugar, a pequena proporção de artigos sobre avaliação de políticas educacionais no conjunto dos textos que tratam da avaliação no ensino básico, e, em particular, dos que se referem à avaliação de programas, sugere que, acerca de medidas que têm ampla repercussão na vida de milhões de brasileiros, tendem a prevalecer grandes vazios de informação e reflexão.

Por sua vez, observa-se que a maior parte das avaliações de programas divulgadas nos periódicos compõe-se de ações cuja implementação ocorreu ou pelo menos foi iniciada durante as gestões da década de 1980. Ainda que se considere necessária a decorrência de um tempo razoável desde o começo de um programa até que os seus resultados tenham condição de ser avaliados, chama atenção o fato de que, sobre as profundas alterações pelas quais vem passando o sistema educacional brasileiro, especialmente nos anos 90, haja volume excessivamente pequeno de dados trabalhados circulando nos principais meios difusores ao final da década.

Isso não significa, porém, que se endossem as políticas de resultados que têm predominado no modo hegemônico de fazer política no setor, não só no país, como em outras partes do mundo. Sabe-se que, em educação, muitos resultados significativos costumam mostrar-se após um prazo bem maior do que o que comportam os delineamentos de

pesquisas de controle. Além disso, por ser a educação uma atividade com características eminentemente formadoras, os processos envolvidos na implementação de programas educacionais ganham especial importância na medida em que chegam a conformar os próprios resultados e afetam tanto os sujeitos-alvo quanto os agentes educacionais, que precisam apossar-se de certas propostas institucionais como suas, condição para que mudanças duradouras se operem. Tais considerações não invalidam, contudo, o pleito de que as ações governamentais sejam mais sistematicamente acompanhadas e avaliadas.

Sobre os níveis de ensino estudados, tal como constatou levantamento congênere (Azevedo & Aguiar, 1999), os estudos incidem majoritariamente sobre o ensino fundamental, e até os programas de formação continuada de docentes são dedicados basicamente aos educadores do ensino obrigatório. De acordo com outra tendência encontrada, observa-se também que a maior parte dos programas analisados foi levada a cabo em estados da região Sudeste, onde estão concentrados os sistemas de ensino de maiores recursos, bem como a maior quantidade de pessoal tecnicamente qualificado para realizar as avaliações. Poucos são os programas de âmbito nacional analisados; apenas um deles tem caráter regional, havendo tão-somente uma referência a um programa desenvolvido por município.

Do ponto de vista metodológico, observa-se que, entre o reduzido número de textos que abordam o tema, uma série deles não se constitui em avaliações propriamente ditas; trata-se de meras apreciações sem maior rigor na abordagem.

Nas avaliações de programas realizadas com metodologia própria, verifica-se, na maior parte dos casos, a preocupação de contemplar tanto as dimensões quantitativas quanto os processos qualitativos detectados no decorrer da implementação. Pode acontecer que avaliações de programas, que tendem a ser descontextualizadas quando feitas sob demanda, sejam publicadas sob a forma de estudos e textos de natureza mais acadêmica, que levam em conta outros aportes.

A escassez de informações sobre os programas governamentais e, mais ainda, a falta de divulgação de estudos realizados sobre o tema recol-

cam, enfim, algumas questões de natureza política. Nem sempre os gestores do setor público, que solicitam avaliações, têm a preocupação de dar conhecimento dos resultados aos diretamente envolvidos com as ações ou a públicos diferenciados. Por vezes, eles não consideram interessante ou conveniente a divulgação de dados que possam comprometer a legitimidade de suas ações. Quando os grupos de interlocução são restritos e os resultados permanecem exclusivamente no domínio dos dirigentes, a informação pode configurar-se em instrumento de poder com características bastante discricionárias, tendendo a reforçar as posições políticas dos segmentos no poder, independentemente da equidade e adequação das medidas postas em prática.

Além disso, a alternância na gestão costuma acarretar descontinuidade nos projetos e programas governamentais, podendo afetar também o interesse pela realização de avaliações, de sorte que as políticas de mais longo prazo, como as da área da educação, correm o risco de não ter prosseguimento.

A disseminação de informações entre os beneficiários diretos, bem como entre outros grupos críticos e a sociedade em geral, permite que haja um contraponto com outros tipos de conhecimento e de interesses, o que pode subsidiar os mecanismos de pressão para que o que está trazendo benefícios coletivos seja preservado, ou para fazer emergir o novo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ARRETCHE, Marta T. Tendências no estudo sobre avaliação. In: RICO, Elizabeth Melo (Org.) *Avaliação de políticas sociais: uma questão em debate*. 2ed. São Paulo: Cortez, Instituto de Estudos Especiais-PUC, 1999.
- AZEVEDO, Janete Maria Lins de, AGUIAR, Márcia Angela da Silva. Políticas de educação: concepções e programas. In: WITTMAN, Lauro Carlos, GRACINDO, Regina Vinhaes (Coord.) *O estado da arte em política e gestão da educação no Brasil*. Brasília: ANPAE, 1999, p.63-76.
- BARRETTO, Elba Siqueira de Sá. Capacitação à distância de professores do ensino fundamental no Brasil. *Educação e Sociedade (Campinas)*, v.18, n.59, p.308-29, CEDES, ago. 1997.
- BARRETTO, Elba Siqueira de Sá, PINTO, Regina Pahim (Coord.). *Estado da arte: avaliação na educação básica*. São Paulo: FCC, Brasília: INEP/COMPED, 2000.
- DEPRESBÍTERIS, Léa. Avaliação de programas e avaliação de aprendizagem. *Educação e Seleção (São Paulo)*, n.19, p.5-31, FCC, jan./jun. 1989.
- FERRETTI, Celso João, VIANNA, Cláudia Pereira, SOUZA, Denise Trento de. Escola pública em tempo integral. *Cadernos de Pesquisa (São Paulo)*, n.76, p.5-17, FCC, fev. 1991.
- FIGUEIREDO, Marcus Faria, FIGUEIREDO, Argelina Maria Cheibub. Avaliação política e avaliação de políticas: um quadro de referência teórica. *Análise e Conjuntura (Belo Horizonte)*, v.1, n.3, p.107-27, set./dez. 1986.
- FLETCHER, Philip R. Propósitos da avaliação educacional: uma análise das alternativas. *Estudos em Avaliação Educacional (São Paulo)*, n.11, p.93-112, FCC, jan./mar. 1995.
- GATTI, Bernardete A. Avaliação educacional no Brasil: experiências, problemas, recomendações. *Estudos em Avaliação Educacional (São Paulo)*, n.10, p.67-80, FCC, jul./dez. 1994.
- GOMES NETTO, João Batista F. et al. Investimentos autofinanciáveis em educação. *Cadernos de Pesquisa (São Paulo)*, n.85, p.11-25, FCC, maio 1993.
- LEONARDOS, Ana Cristina, MACEDO, Sandra. A abordagem etnográfica na avaliação de projetos inovadores. *Ensaio: Avaliação e Políticas Públicas em Educação (Rio de Janeiro)*, v.4, n.11, p.109-20, Fundação CESGRANRIO, abr./jun. 1996.
- MARTINS, Angela Maria. Formação contínua de professores: parceria universidade e rede estadual de ensino. *Revista da Faculdade de Educação (São Paulo)*, v.22, n. 2, p.189-205, FEUSP, jul./dez. 1996.
- NEUBAUER da SILVA, Tereza Roserley, DAVIS, Cláudia. É proibido repetir. *Estudos em Avaliação Educacional (São Paulo)*, n.7, p.5-44, FCC, jan./jul. 1993.
- OLIVEIRA, João Batista de Araújo e. Análise de custos e resultados do Programa de Aceleração de Aprendizagem. *Ensaio: Avaliação e Políticas Públicas em Educação (Rio de Janeiro)*, v.8, n.18, p.23-56, Fundação CESGRANRIO, jan./mar. 1998.
- RAMOS, Lilian Maria P. de C. Escolas rurais consolidadas: experiência paranaense. *Cadernos de Pesquisa (São Paulo)*, n.77, p.19-24, FCC, maio 1991.
- SILVA, Aida Maria Monteiro. A Formação continuada do professor: a relação constitucional entre Secretaria de Educação e Universidade. *Revista da Faculdade de Educação (São Paulo)*, v.22, n.2, p.178-88, FEUSP, jul./dez. 1996.
- SILVA, Zoraide Inês Faustini. Um estudo avaliativo sobre o ciclo básico de alfabetização em quatro escolas de São Paulo. *Estudos em Avaliação Educacional (São Paulo)*, n.3, p.49-70, FCC, jan./jun. 1991.
- SOUSA, Edson Machado de. Políticas públicas e a questão da avaliação. *Ensaio: Avaliação e Políticas Públicas em Educação (Rio de Janeiro)*, v.1, n.2, p.51-60, Fundação CESGRANRIO, jan./mar. 1994.
- VASCONCELLOS, Eduardo Alcântara de. Agrupamentos de escolas rurais: alternativa para o impasse da educação rural? *Cadernos de Pesquisa (São Paulo)*, n.86, p.65-73, FCC, ago. 1993.
- VIANNA, Heraldo Marelin. Avaliação educacional: uma perspectiva histórica. *Estudos em Avaliação Educacional (São Paulo)*, n.12, p.7-24, FCC, jul./dez. 1995.
- _____. Avaliação: considerações teóricas e posicionamentos. *Estudos em Avaliação Educacional (São Paulo)*, n.16, p.5-36, FCC, jul./dez. 1997.
- ZALUAR., Alba, LEAL, Maria Cristina. Cultura, educação popular e escola pública. *Ensaio: Avaliação e Políticas Públicas em Educação (Rio de Janeiro)*, v.4, n.11, p.159-76, Fundação CESGRANRIO, abr./jun. 1996.

NOTAS

1 Embora o ciclo básico não possa, a rigor, ser rotulado como um programa, por se tratar de iniciativa que anuncia apenas o início da reestruturação do ensino de primeiro grau como um todo, decidiu-se considerá-lo como tal para fins de análise, uma vez que, em relação a ele, foram tomadas uma série de medidas com finalidades específicas, que permitiram distinguir a sua trajetória enquanto inovação na rede escolar.

2 É neles que se efetua a maior parte das pesquisas sobre educação nas universidades brasileiras.

* Trabalho apresentado no II Congresso Luso-Brasileiro de Política e Administração da Educação. Braga, Universidade do Minho, de 18 a 20 de janeiro de 2001.

Maria do Carmo Brant de Carvalho
Denise Blanes

Sistema de
monitoramento
e avaliação
implementado por
uma organização
do terceiro setor:
o caso da Ação
Comunitária
do Brasil

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.

DENISE BLANES é pesquisadora do Instituto de Estudos Especiais — IEE, da PUC-SP. Doutoranda do curso de pós-graduação em Serviço Social da Pontifícia Universidade Católica de São Paulo, é autora de vários trabalhos publicados.

A Ação Comunitária do Brasil — ACB já possui um longo currículo de serviços prestados às comunidades carentes da zona sul de São Paulo. Nasceu em 1967 e desde lá vem ampliando e diversificando suas atividades. Atua em quarenta comunidades, tendo como parceiros as próprias lideranças e agentes locais. Este é um grande diferenciador do projeto da ACB: suas ações partem de interesses e demandas locais e são implementadas com suas lideranças. Está fincada na comunidade.

Seu público principal (cerca de 5 mil pessoas) é formado por crianças, adolescentes e jovens provenientes de famílias de baixa renda, habitantes de comunidades da periferia da cidade. Esses dados por si só apontam para a dimensão e importância de suas ações na esfera pública.

Em junho de 1999, a Ação Comunitária do Brasil nos solicitou a implementação de um sistema de monitoramento e avaliação de seus serviços. É preciso destacar que tal solicitação — partindo de uma organização filantrópica — é uma iniciativa inovadora e corajosa. São poucas as organizações não-governamentais, no campo da ação social, que produzem avaliações de resultados e impactos. Nesta área, as organizações, em geral, não percebem a importância de aferir resultados. É como se “para pobres, qualquer ação fosse lucro”. Não há compromisso com a

transparência. O que move essas organizações filantrópicas é a preocupação com a expansão quantitativa de seus serviços, mais do que a busca de sua qualidade e eficácia.

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.

Num país e cidade como a nossa, onde os universos populacionais são enormes, os recursos precisam ser otimizados e melhor distribuídos.¹ A transparência e socialização de resultados e impactos produzidos pelas organizações governamentais e não-governamentais também se tornou tarefa urgente no fortalecimento da democracia.

A montagem do sistema de monitoramento e avaliação foi realizada com intensa participação dos técnicos e coordenadores da ACB. Esta participação é importante, pois permite aos sujeitos da ação — gestores, técnicos e agentes — apropriarem-se do seu fazer e lhe conferir significado, ou seja, compreenderem que a avaliação sistemática permite aferir processos/resultados/impactos e referenciar as decisões políticas e técnicas da instituição. A participação é também imprescindível, pois se trata de uma avaliação contínua, o que exige da organização o domínio do seu fazer.

A CONSTRUÇÃO DO SISTEMA

Realizamos uma série de oficinas com técnicos e coordenadores para que compreendessem o significado da avaliação e participassem da construção do Sistema de Monitoramento e Avaliação. Este momento foi extremamente rico. Refletiu-se o percurso histórico da ACB, analisando seu presente e projetando o seu futuro. Foi necessário melhor focalizar a missão da entidade. Quais ações e estratégias constituem o “núcleo duro” da missão institucional ou expressam melhor a identidade

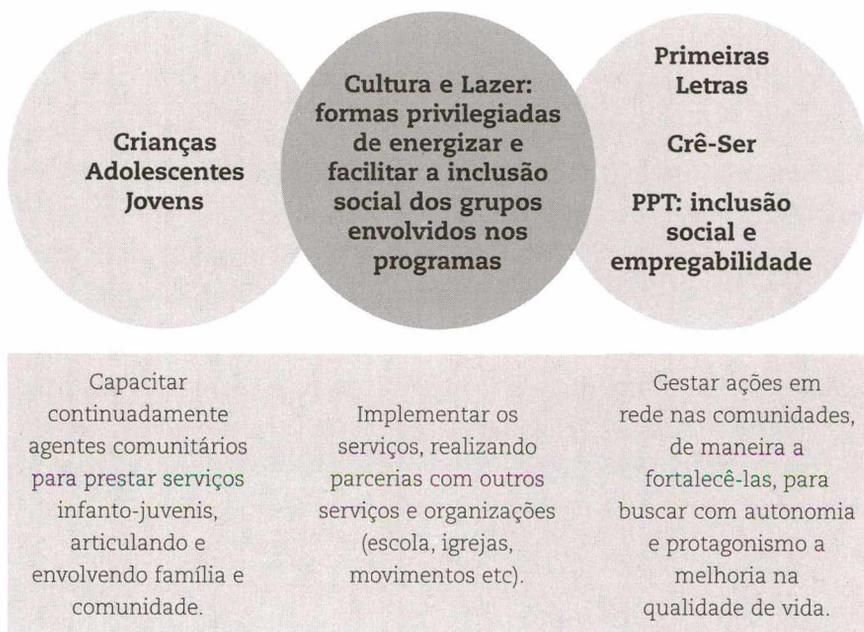
da instituição? Quais ações conferem à ACB maior legitimidade e reconhecimento na comunidade? Quais públicos-alvo das ações são preferenciais?

Definiu-se mais claramente o seu público-alvo: crianças, adolescentes, jovens. Como diretriz, ficou fortalecida a idéia de ações (programas e serviços) implementadas com a comunidade, articulando-se de forma mais intensa as lideranças, agentes comunitários e movimentos de bairro. As ações lúdicas e culturais assumem um papel de animação comunitária, sendo desenvolvidas de forma articulada aos programas nucleares.

Estabeleceram-se, assim, os programas considerados “carros-chefes” da instituição:

- Primeiras Letras
- Crê-Ser
- Iniciação para o Trabalho

Essas definições mostram que a montagem de um Sistema de Avaliação impõe, necessariamente, uma revisão à missão da organização e portanto, uma avaliação institucional.



Com esses desdobramentos, o Sistema de Monitoramento e Avaliação construído centrou-se nos programas Primeiras Letras, Crê-Ser e Iniciação para o Trabalho.

Na seqüência, refletimos com a equipe técnica de cada programa sobre a coerência interna entre objetivos, estratégias e resultados desejados, retirando desta reflexão os aspectos/eixos a serem avaliados; indicadores e instrumentos quali-quantitativos de aferição.² A avaliação contemplou agentes comunitários/educadores e crianças/adolescentes.

Foram definidos indicadores, tomando-se por base os grandes eixos educativos desenvolvidos pela ACB com crianças (Primeiras Letras), crianças e adolescentes (Programa Crê-Ser) e jovens (Programa de Iniciação para o Trabalho), quais sejam:

- *Conhecer*: desenvolvimento da capacidade cognitiva;
- *Fazer*: desenvolvimento de habilidades para a vida;
- *Ser*: desenvolvimento da identidade;
- *Conviver*: desenvolvimento da convivência.

Para cada eixo, foram debatidos e construídos, com a equipe técnica, os indicadores respectivos. A aferição ocorreu apenas sobre os indicadores correspondentes a objetivos efetivamente perseguidos. No ano de 1999, os Programas não desenvolviam, na sua totalidade, todos os processos e conteúdos necessários à obtenção plena dos objetivos implicados em cada eixo.

O desenho do sistema de monitoramento e avaliação foi todo construído com os gestores e a equipe técnica da ACB. Os testes para avaliação das aprendizagens das crianças e adolescentes, por exemplo, foram elaborados e propostos pelas equipes técnicas de cada programa, o mesmo ocorrendo com a implementação do sistema, que envolveu até mesmo os agentes comunitários/educadores. Por fim, construiu-se uma ficha de matrícula com registro mais completo de informações sobre as crianças/adolescentes e suas respectivas famílias, necessária ao sistema de monitoramento.

É importante afirmar que esses primeiros produtos avaliativos espelham um “estado das artes” ou, ainda, uma avaliação diagnóstica. É o ponto de partida do Sistema.

Para efeito deste relato, suprimimos a avaliação do programa voltado mais especificamente aos jovens (PPT — Programa de Preparação para o Trabalho).

PRIMEIROS RESULTADOS AVALIATIVOS

A personalidade da criança, seus “raciocínios” e seus comportamentos, suas ações e reações são incompreensíveis fora das relações sociais que se tecem, inicialmente, entre ela e os outros membros da constelação familiar, em um universo de objetos ligados às formas de relações intrafamiliares. De fato, a criança constitui seus esquemas comportamentais, cognitivos e de avaliação através das formas que assumem as relações de interdependência com as pessoas que a cercam com mais frequência e por mais tempo, ou seja, os membros de sua família. Ela não “reproduz”, necessariamente e de maneira direta, as formas de agir de sua família, mas encontra sua própria modalidade de comportamento em função da configuração das relações de interdependência no seio da qual está inserida. Suas ações são reações que “se apóiam” relacionalmente nas ações dos adultos que, sem sabê-lo, desenhavam, traçam espaços de comportamentos e de representações possíveis para ela.³

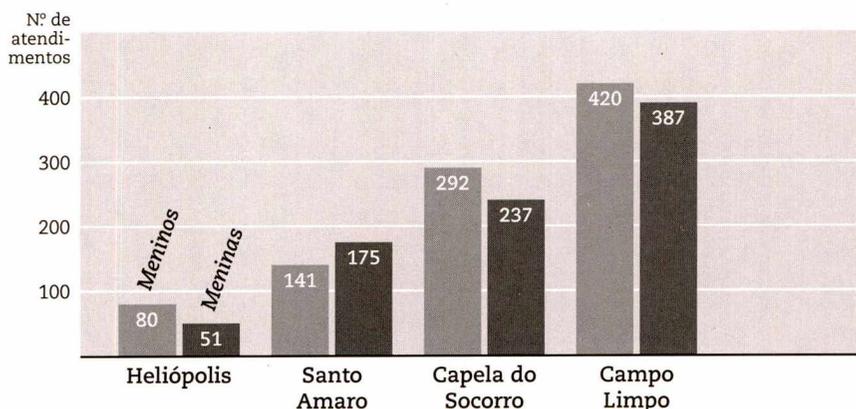
AVALIAÇÃO DO PROGRAMA CRÊ-SER

O programa está implantado em 26 comunidades, do município de São Paulo, com 53 turmas no total, beneficiando 1.783 crianças e adolescentes entre 7 e 14 anos, no ano de 1999.

- 13 comunidades com 24 turmas na região de Campo Limpo
- 8 comunidades com 16 turmas na região de Capela do Socorro
- 4 comunidades com 9 turmas na região de Santo Amaro
- 1 comunidade com 4 turmas em Heliópolis

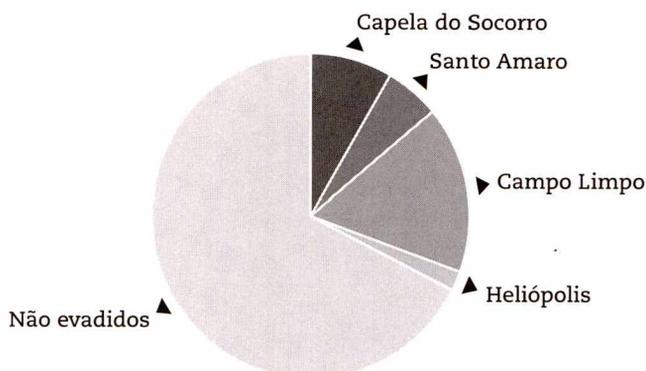
Campo Limpo: 807	Capela do Socorro: 529
Meninos: 420	Meninos: 292
Meninas: 387	Meninas: 237
Desistentes: 298	Desistentes: 151
Santo Amaro: 316	Heliópolis: 131
Meninos: 141	Meninos: 80
Meninas: 175	Meninas: 51
Desistentes: 92	Desistentes: 33
Total de meninos: 933	
Total de meninas: 850	
Total geral: 1.783	
Total de evasão: 574 (32%)	

Distribuição do atendimento por sexo e por região



Em março de 1999, foram contabilizadas 1.256 crianças/adolescentes, o que indica o ingresso durante o ano de mais 527 crianças/adolescentes. Esse dado, combinado ao dado de evasão, sugere uma certa rotatividade de participantes que não se fixam no programa durante todo o ano letivo.

Distribuição da evasão por região, no conjunto de crianças atendidas



O número de evadidos ou desistentes é alto. Na média total, está na ordem de 32%. Há um conjunto de fatores que o explicam. Um deles é o fato de os espaços serem pouco apropriados para conter uma programação com atividades diversificadas e rica em estímulos (esporte, cultura etc.), que respondam aos interesses de um grupo etário que vai de 7 a 14 anos. A programação fica, assim, mais restrita a aprendizados escolares e não a aprendizados complementares à escola, o que sem dúvida pode gerar desmotivação.

Outro fator é a falta de parceria com as famílias. Não há compromissos e ações desenvolvidas em conjunto, o que torna o vínculo com os pais e com suas crianças bastante frágil. O mesmo ocorre com a escola,⁴ não se registrando relações de reconhecimento mútuo e, menos ainda, ações de parceria.

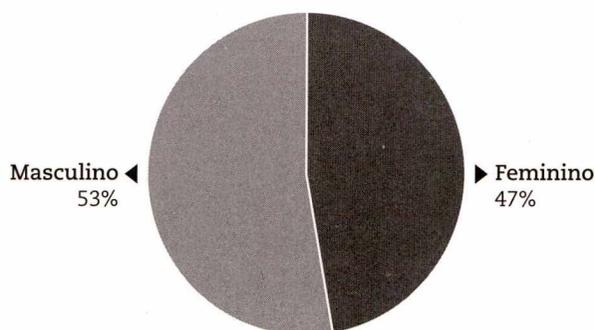
Esses fatores acabam por fixar no agente/educador toda a possibilidade/responsabilidade de “segurar” e motivar crianças e adolescentes no programa.

O fato de a turma ser grande e a faixa etária ser extensa (de 7 a 14 anos) determina maior evasão, pois não permite ao agente comunitário oferecer uma relação interpessoal densa, combinada com atividades que atendam a interesses muito diversos. As turmas grandes têm uma evasão acima de 32%, havendo exceções, como, por exemplo, uma turma

no Paulo VI, com 41 matriculados e 5% de evasão. Assim, pode-se dizer que outro fator é a própria competência do agente/educador.

No caso das turmas que envolvem crianças e adolescentes (de 7 a 14 anos), pode ocorrer que os pais considerem “pouco seguro manter seus filhos menores em contato com os maiores”. Em menor escala, ocorrem desistências de meninas que ficam em casa para dar conta dos afazeres domésticos quando a mãe trabalha fora, ou para serem preservadas da violência das ruas. Há, enfim, evasões por mudanças de domicílio. Esses últimos fatores sugerem um nível de evasão não controlável pelo Programa, isto é não dependente do seu grau de desempenho.

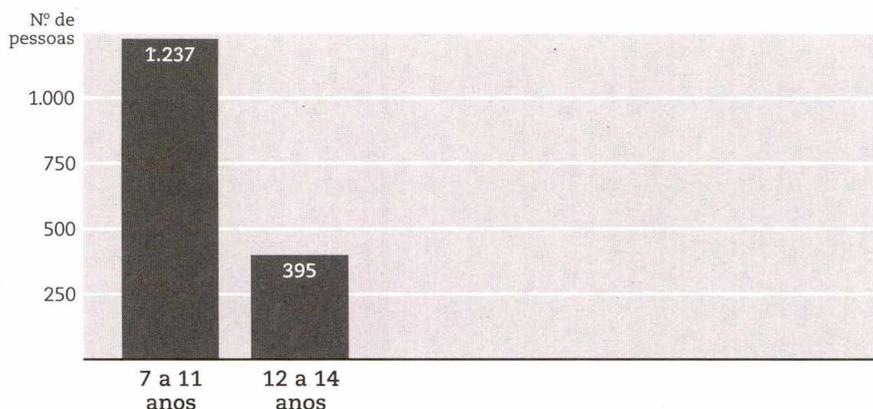
Distribuição de crianças e adolescentes por sexo



O gráfico acima destaca uma predominância de meninos, o que pode sugerir uma nova preocupação das famílias da periferia: manter os filhos distantes de gangues e drogas, constante ameaça na socialização desejada para os filhos homens.

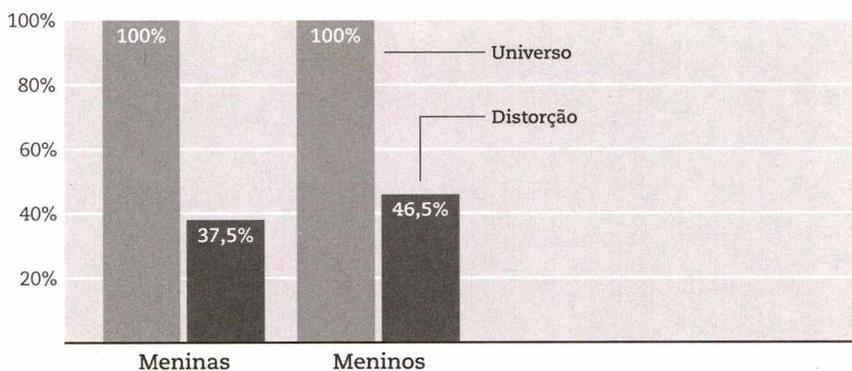
O gráfico a seguir, por sua vez, confirma a predominância do atendimento das crianças de 7 a 11 anos. Além desse grupo, há apenas uma criança com 6 anos, três jovens com 15 anos e 77 que não informaram a idade.

Distribuição de crianças e adolescentes por faixa etária



Esse dado permitiria repensar as opções político-programáticas referentes ao programa Crê-Ser: atuar melhor e mais focadamente nos 7 a 11 anos ou manter a extensão da faixa etária dividindo as turmas e diversificando os conteúdos e estratégias programáticas.⁵

Distorção idade/série x universo



Fica claro que há um percentual significativo de meninos com distorção de idade/série escolar (46,5%), acima da média nacional para esse índice (44%). Já o índice de distorção idade/série escolar das meninas é de 37,5%, refletindo uma situação melhor. Se considerarmos que a concentração maior

das crianças é em torno de 7 a 11 anos, é preciso repensar até mesmo os aprendizados, de forma a prevenir ou corrigir precocemente tais distorções. É sabido que essa distorção entre crianças provenientes de famílias de baixa renda tem menos a ver com reforço escolar *stritu senso* e mais com ações incisivas de ampliação de seu universo informacional e cultural.

Avaliação de aprendizados das crianças/adolescentes

O que fizemos?

Elaboramos, em conjunto com os técnicos, os instrumentais para aferição dos aprendizados das crianças e adolescentes, referenciados nos eixos já descritos. Tais instrumentais foram aplicados pelos agentes/educadores, sob a supervisão dos técnicos do Programa. Esses realizaram leitura de todo o material, análise das produções e dos discursos, tomando como referência as orientações do processo de avaliação. Sistematizamos estatisticamente os dados avaliados. Numa amostra de casos, aprofundamos a análise do material produzido pelas crianças e pelos adolescentes.

É importante reafirmar que, em se tratando de uma primeira avaliação sistemática de desempenho das crianças, os testes refletem as perspectivas de aprendizado esperadas pelos técnicos. Desse modo, a análise recai sobre o desempenho das crianças e, simultaneamente, do programa.

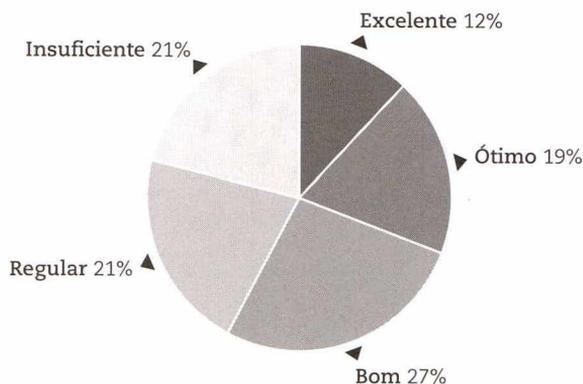
Dos 1.209 crianças/adolescentes participantes em dezembro/99, 1.075 realizaram algum tipo de avaliação de aprendizado, o que é extremamente positivo.

Nos testes relativos ao eixo Conhecer (1), foram verificadas as capacidades relativas a :

- Expressar-se de forma espontânea;
- Usar diversas formas de expressão;
- Expressar sentimentos;
- Elaborar e expressar idéias e pensamentos.

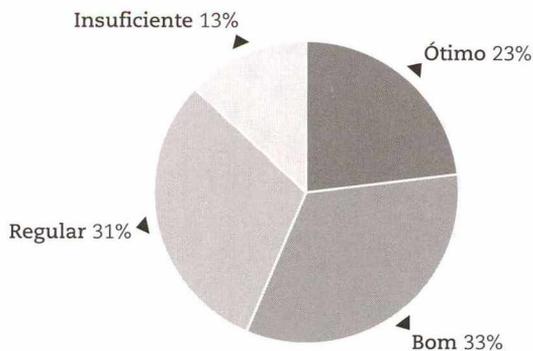
A escala utilizada foi do ótimo ao insuficiente, especificando-se as capacidades atingidas dentro de cada graduação. Utilizamos classificação excelente neste aprendizado por termos identificado um grupo de crianças com desempenho muito acima da média.

Conhecer (1)



Utilizou-se um segundo instrumental de avaliação relativo ao eixo Conhecer, pois a equipe técnica desejava aferir duas capacidades trabalhadas em parceria com outra organização (Cenpec). Assim, o atributo *interesse pela leitura e capacidade de localizar, acessar e processar informações* foram verificadas no que se denominou Conhecer (2).

Conhecer (2)



Os instrumentais utilizados revelam um forte acento em aprendizagens escolares. Por esses resultados, é possível afirmar que o desempenho é bom para 56% das crianças e adolescentes participantes.⁶

Não houve uma avaliação do desenvolvimento de habilidades para a vida, denominadas aqui como Fazer. A equipe técnica considerou a formação do educador e o próprio conteúdo programático como insuficientes para permitir a observação e o registro desse eixo.

Essas mesmas razões justificaram o fato de serem agrupados os eixos Conviver e Ser no momento de sua aferição entre as crianças. A pouca ênfase demonstrada revela um déficit no Programa, que deverá ser superado. É importante destacar, ainda, que tais dimensões são distintas, referindo-se à sociabilidade e à construção de identidade. Era necessário isolá-las para uma melhor apreciação.

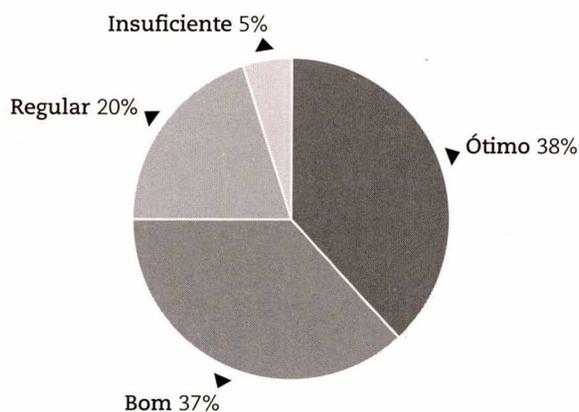
As capacidades e os valores considerados no eixo Conviver foram:

- respeito a si mesmo e ao outro;
- capacidade de diálogo e negociação;
- capacidade de fazer relações;
- capacidade de verbalizar sentimentos.

No eixo Ser foram consideradas as capacidades de:

- perceber-se, identificando suas necessidades e limites, valores pessoais, dificuldades e potencialidades;
- cuidar da aparência pessoal;
- reconhecer e valorizar seus espaços e grupos de pertencimento.

Conviver e Ser



Essas classificações são resultantes basicamente da observação e da convivência dos agentes/educadores com as crianças/adolescentes com quem atuavam. Sem dúvida, existe aí uma certa subjetividade, mas de toda forma revela-se a percepção dos agentes/educadores. O dado permite

inferir que a percentagem de crianças/adolescentes (5%) com algum problema de comportamento ou sociabilidade é pouco significativa.

Esse dado é coerente com a percentagem de crianças/adolescentes com registro de comportamentos⁷ como agressividade, agitação, desobediência, entre outros, nas folhas de observação preenchidas pelos agentes comunitários.

Que análise mais acurada foi possível produzir?

Destacamos algumas turmas para realizar uma avaliação mais acurada do desempenho das crianças e adolescentes, valendo-nos mesmo da colaboração de especialistas⁸ da área.

Como já afirmado, parece haver uma forte preocupação com aprendizados escolares, sem que se possa inferir que o Programa tenha impacto direto no sucesso escolar desse grupo de crianças/adolescentes. Há observações de um número significativo de crianças que apresentam problemas de alfabetização.⁹

○ Programa reflete uma tensão entre dois enfoques distintos presentes em programas dessa natureza: um voltado para aprendizados escolares e outro para ampliação do repertório cultural. O campo da complementação escolar põe maior acento na promoção da convivência; no desenvolvimento dos fazeres/habilidades exigidos na vida cotidiana privada e pública; na ampliação do universo cultural; na sociabilidade dos educandos; na oferta de oportunidades lúdicas e esportivas. Esse enfoque mais abrangente vem ganhando maior consenso sobre programas complementares de educação voltados a crianças e adolescentes.

Os instrumentos de avaliação utilizados (fichas de observação e testes) têm limites. Nessa primeira avaliação, não foi possível contemplar algumas seqüências diárias de atividades realizadas com as crianças e adolescentes para evitar possível redução na análise. O trabalho com Arte, por exemplo, poderia estar representado por seqüência de produções num determinando espaço de tempo, o mesmo ocorrendo em relação à produção e interpretação de textos. Quanto aos jogos, deveríamos pensar em abranger uma diversidade de objetivos e conteúdos, observando os fazeres não apenas do ponto de vista do desempenho individual.

Crê-Ser — alunos de 1.^a a 4.^a série

Durante o ano de 1999, na comunidade de Mocaph, foram inseridos 34 alunos no projeto, tendo havido uma evasão de 8 participantes. Dos dezoito que realizaram os testes e foram avaliados, a maior parte está cursando a terceira série.

	Conviver	Conhecer	Ser	Vida Pública
1. ^a a 4. ^a série	4%	51%	38%	8%

Temos uma concentração no eixo “Aprender a Conhecer”, o que concorre com as observações anteriores em relação à complementação escolar X escola formal e é reforçado pela priorização de critérios de avaliação para esse eixo.

Considerando o conjunto de testes e dos indicadores contemplados, podemos inferir que o desempenho desse grupo encontra-se entre bom e regular. As crianças apresentam vocabulário restrito e repetitivo, o que aparece como estereotipia nos desenhos. O aprendizado, do ponto de vista das crianças, aparece centrado na questão da disciplina. Entre as crianças que não escreveram nos testes, algumas apresentam traços que permitem supor que sabem desenhar letras, outras, sabem ler e escrever, segundo o relato da educadora.

Crê-Ser — alunos de 1.^a a 8.^a série

Durante o ano de 1999, na comunidade de São Bernardo, numa turma foram inseridas 37 crianças e adolescentes, tendo havido uma evasão de 16 participantes. Das 24 que permaneceram, 70% está cursando entre a 4.^a e 6.^a séries, numa faixa etária entre 10 e 13 anos. É importante notar que há uma diversidade maior entre as idades dos educandos neste grupo (ressaltando a existência de 30% de crianças e adolescentes entre 7 e 9 anos).

	Conviver	Conhecer	Ser	Vida Pública
5. ^a a 8. ^a série	20%	43%	37%	0,5%

Neste caso, vemos um certo equilíbrio entre os três eixos: Conviver, Conhecer e Ser. Estamos considerando como aprendizado relativo à vida pública ações com a comunidade, seja na forma de relação com seus moradores, seja na forma de ações na comunidade ou pesquisa para conhecimento desta.

Considerando o conjunto de testes, podemos inferir que o desempenho desse grupo encontra-se entre bom e ótimo. Pelo conteúdo do texto dos educandos e as observações da educadora, é possível perceber uma riqueza nos trabalhos desenvolvidos e maior precisão na observação do desenvolvimento das crianças e adolescentes.

Quanto às “reproduções artísticas”, parece que a atividade foi organizada de forma a permitir maior criatividade das crianças e adolescentes; as observações da educadora expressam clareza de objetivos da atividade.

Algumas considerações gerais

Os educadores trabalham com objetivos e entendimentos muito diferenciados das atividades que realizam e da expectativa em relação aos educandos. Ao mesmo tempo, parecem ter entendimentos diferenciados sobre a função e os critérios de avaliação.

Apesar do eixo mais presente nas observações das educadoras ter sido o Conhecer, não são destacadas considerações sobre a curiosidade e a pesquisa das crianças e adolescentes, questões imprescindíveis quando se trata desse eixo.

As avaliações sinalizam, finalmente, um traço institucional direcionado mais para o pedagógico *stritu senso*.

Dados sobre os agentes/educadores

São 38 os agentes comunitários que atuam no Programa Crê-Ser, a avaliação de seu desempenho — realizada pela equipe técnica — indica um desempenho ótimo de 24% (nove educadores); bom de 45% (dezesete educadores) e regular de 29% (treze educadores).

Do total de educadores, cinco possuem ou estão cursando o ensino superior; seis possuem o magistério e doze o ensino médio; os demais

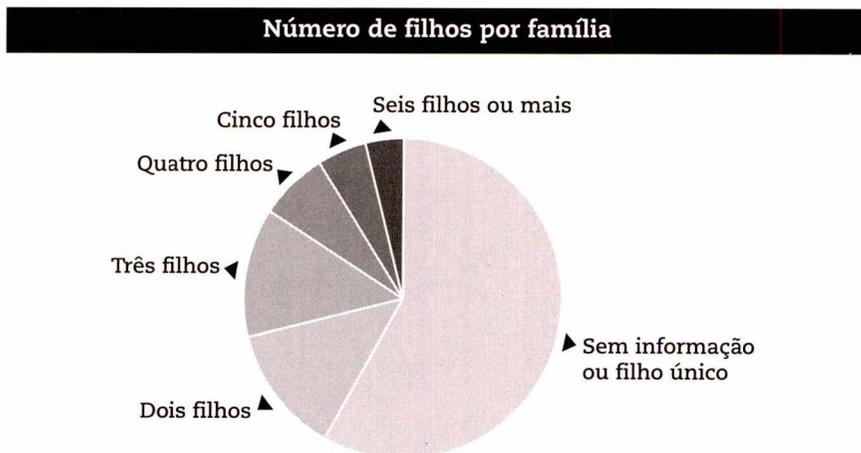
possuem o ensino médio incompleto (ou em curso) ou o ensino fundamental completo. Há ainda três agentes com o ensino fundamental incompleto. Essas diferenças de escolaridade não rebateram no desempenho avaliado pela equipe técnica, ou seja, há ótimos desempenhos entre agentes com ensino superior, como também entre aqueles que têm apenas o ensino fundamental.

Dados sobre as famílias do Crê-Ser

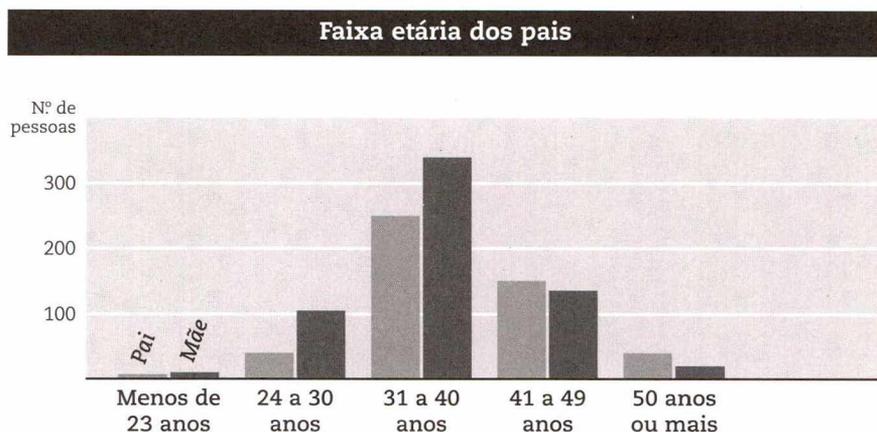
Indica-se aqui, fundamentalmente, a ausência de registro (ou desconhecimento) de dados sobre as famílias das crianças/adolescentes. Essa ausência pode significar falta de relação próxima com os familiares. É possível inferir que o trabalho restringe-se às crianças, não havendo envolvimento dos pais.

Ficará clara esta despreocupação na leitura dos gráficos abaixo, que retratam o não-preenchimento das informações solicitadas ou, por exemplo, a distorção dos dados da composição familiar. De qualquer forma, as informações existentes apontam tendências.

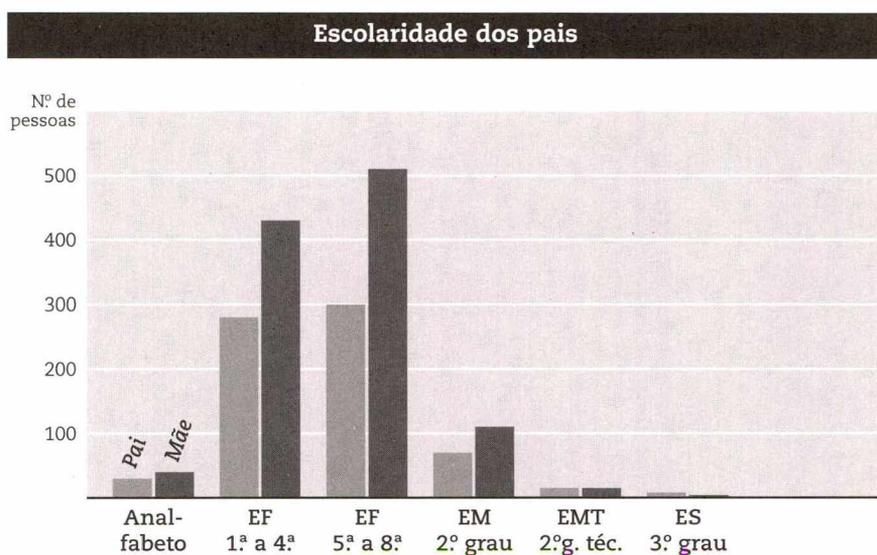
Observa-se, igualmente, que há mais informações registradas sobre as mães do que sobre os pais. Esse dado tanto pode significar um percentual significativo de famílias monoparentais, chefiadas por mulheres, quanto maior proximidade dos agentes/educadores com a figura materna.



Há concentração de pais na faixa etária de 31 a 40 anos, o que significa famílias jovens.

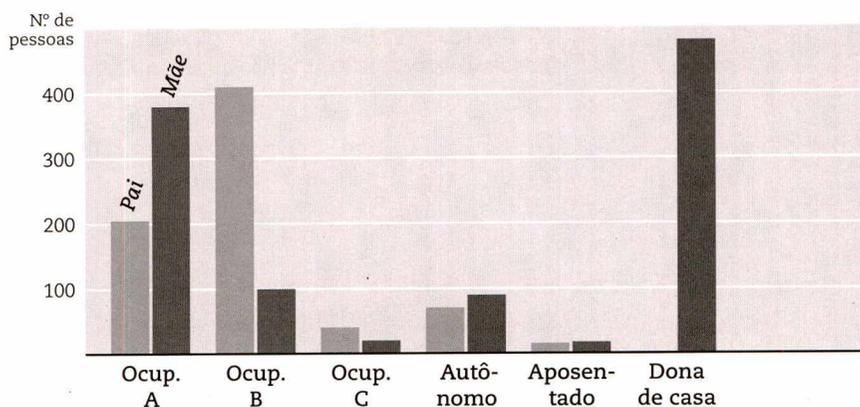


Não há registro de informações quanto à faixa etária de 72% dos pais e 65% das mães.



Não há registro de informações quanto à escolaridade de 59% dos pais e 33% das mães. Das informações obtidas, é possível visualizar maior concentração de pais com escolaridade acima de 5ª série.

Ocupação dos pais



Não há registro de informações quanto à ocupação de 57% dos pais e 37% das mães. Há um número expressivo de mulheres “donas de casa”, ou seja, que não trabalham fora do lar. Este dado sugere a possibilidade de envolver essas mulheres nos Programas da ACB (esporte, cultura e maior participação no cotidiano do Programa). É interessante notar, igualmente, que, entre as que trabalham, há uma preponderância de mulheres inseridas em ocupações manuais e de homens em ocupações manuais que exigem qualificação.¹⁰

AVALIAÇÃO DO PROGRAMA PRIMEIRAS LETRAS

O programa é extenso, abrangendo hoje oitenta turmas. Está implantado nas regiões de Campo Limpo, Santo Amaro, Capela do Socorro e Heliópolis, num total de 41 turmas de PRÉ, 38 turmas de JD e uma turma BRINQ (Brinquedoteca).

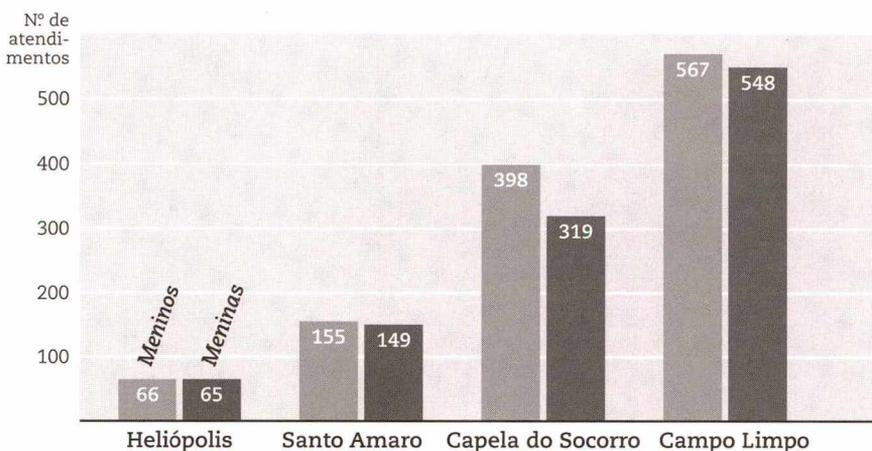
Foram incluídas, na avaliação geral, 71 turmas, divididas por regiões conforme o quadro apresentado, sendo atendidas, no conjunto, 2.267 crianças na faixa etária entre três anos e sete anos.¹¹

- 34 turmas na região de Campo Limpo
- 23 turmas na região de Capela do Socorro
- 10 turmas na região de Santo Amaro
- 4 turmas em Heliópolis

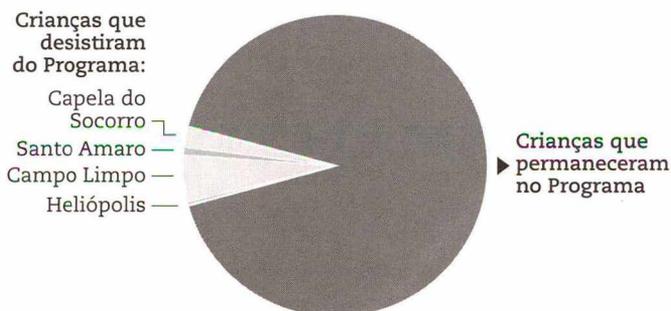
Campo Limpo: 1.115	Capela do Socorro: 717
Meninos: 567	Meninos: 398
Meninas: 548	Meninas: 319
Desistentes: 103	Desistentes: 42
Santo Amaro: 304	Heliópolis: 131
Meninos: 155	Meninos: 66
Meninas: 149	Meninas: 65
Desistentes: 17	Desistentes: 5

São 1.186 meninos e 1.081 meninas, havendo, portanto, uma pequena predominância de meninos. O índice de evasão é bastante baixo (167 crianças), menos de 10%, observando-se, porém, em algumas turmas, a ausência de informações sobre crianças, o que pode indicar uma evasão um pouco superior.

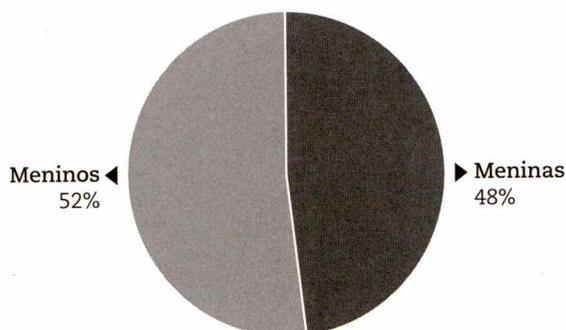
Distribuição do atendimento por sexo e por região



Distribuição da evasão por região, no conjunto de crianças atendidas

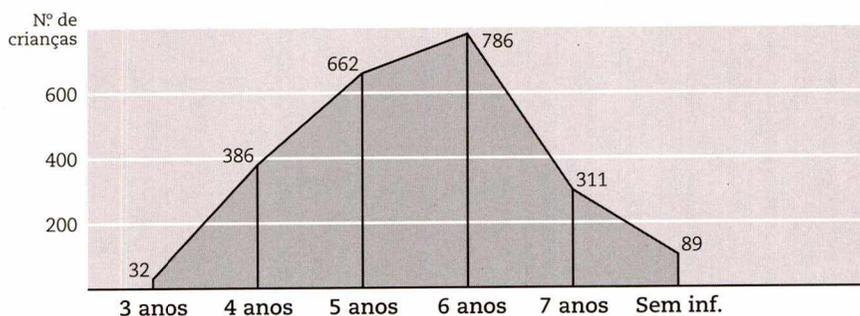


Distribuição de crianças por sexo



No quadro abaixo, visualizamos a distribuição das crianças por idade e a grande concentração do atendimento na faixa dos cinco a seis anos.

Distribuição de crianças por idade



Nas turmas desse programa, observou-se a ausência de limites quanto ao número de crianças, embora o acordo entre as lideranças e a ACB estabeleça um mínimo de vinte e um máximo de trinta crianças por turma. Apesar disso, há muitas turmas com mais de trinta crianças. Ora, este programa exige turmas com no máximo 25 crianças, pois envolve, além de ações voltadas ao desenvolvimento infantil, cuidados e proteção.

Avaliação de aprendizados das crianças

As crianças foram avaliadas nos quatro eixos já mencionados: Conhecer, Fazer, Conviver e Ser. Da mesma forma, foram utilizados, aqui, fichas

de observação e testes. A escala utilizada foi do ótimo ao insuficiente, especificando-se as capacidades atingidas dentro de cada graduação. No eixo Fazer, utilizou-se uma graduação de satisfatório e não satisfatório.

Conhecer:

Comunica-se de forma a compreender e ser compreendido;
Utiliza e opera com símbolos, idéias, imagens e representações;
Enriquece seu repertório de palavras.

Fazer:

Explora o ambiente e o universo de relações;
Sabe manipular os objetos do cotidiano.¹²

Conviver:

Demonstra interesse para ouvir, perguntar, pedir ajuda, colaborar com o outro;
Utiliza o diálogo como forma de lidar com os conflitos;
Utiliza alguns princípios elementares de convívio social.

Ser:

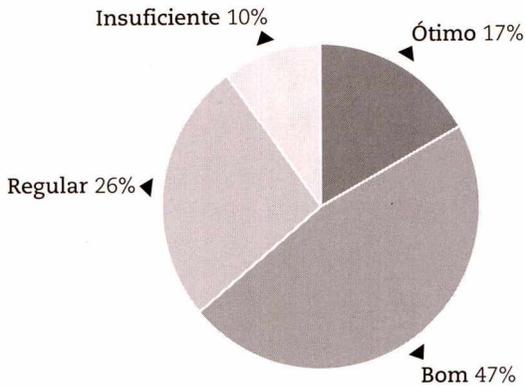
Demonstra cuidados com sua higiene e aparência pessoal;
Tem consciência de sua identidade (nome, características físicas e pessoais);
Demonstra iniciativa e interesse para envolver-se em situações de brincadeiras dirigidas ou livres, coletivas ou individuais.

O que fizemos?

Sistematizamos, estatisticamente, os dados avaliados pelos profissionais do projeto, ao mesmo tempo que os adjetivamos com base em uma amostragem de material produzido pelos alunos. Na amostragem, procedeu-se à leitura de todo o material, à análise das produções e dos discursos, tomando-se como referência as orientações do processo de avaliação.

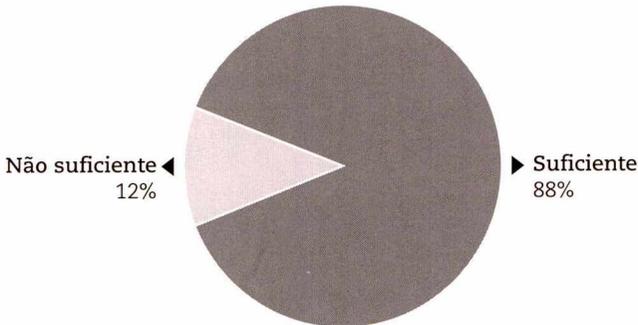
É importante reafirmar que, em se tratando de uma primeira avaliação sistemática de desempenho das crianças, os testes foram produzidos com base nas perspectivas de aprendizado esperadas pelos técnicos, nas particularidades que esse Programa assume (desenvolvido por agentes comunitários, nas periferias da cidade, em espaços geralmente precários etc.).

Conhecer



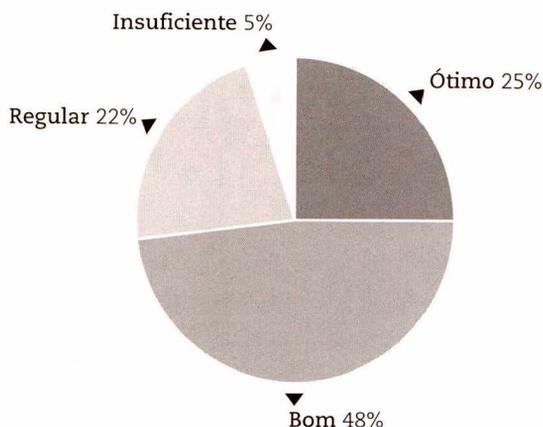
Embora a equipe técnica tenha enfatizado a não-preocupação com a alfabetização precoce ou mesmo com a prontidão para a escrita, desejou-se captar essa dimensão. A avaliação reflete um desempenho entre bom e ótimo em 64% das crianças.

Fazer



A avaliação sugere que a grande maioria tem um desempenho satisfatório no que se refere ao domínio de habilidades esperadas nessa faixa etária. Nas informações colhidas no registro das observações realizadas pelos educadores não há, porém, ênfase sobre o Fazer. Desta forma, fica uma interrogação: serão essas habilidades trabalhadas pelos agentes/educadores ou pelas famílias?

Conviver e Ser



Essas classificações são resultantes basicamente da observação e da convivência dos agentes/educadores com as crianças com quem atuavam. Como já afirmado, existe aí uma certa subjetividade, mas de toda forma revela-se a percepção dos agentes/educadores. A percentagem de crianças (5%) com algum problema de comportamento ou sociabilidade, tal qual no Programa Crê-Ser, é pouco significativa.

Que análise mais acurada foi possível produzir?

Escolhemos algumas turmas, aleatoriamente, para realizar uma avaliação mais acurada do desempenho das crianças, valendo-nos mesmo da colaboração de especialistas¹³ da área.

Jardim da Infância — alunos com quatro anos

O grupo de alunos esteve em atendimento na ACB-Sede, durante o ano de 1999; apenas uma criança passou a compor o grupo no segundo semestre. Podemos, assim, pressupor que o “professor-avaliador” conhece os alunos em seu desenvolvimento.

Quanto ao teste *prático*, a performance apresentada parece indicar certa homogeneidade no que se refere ao desenvolvimento de habilidades dos alunos. Com relação a algumas questões de coordenação motora (usar talheres, amarrar o calçado e abotoar as roupas), seria necessário realizar

atividades lúdicas que permitissem a alguns alunos exercitar essas habilidades, dando um sentido a essas tarefas que não só o da necessidade.

	Conviver	Conhecer	Ser	Fazer
4 anos	32%	22%	30%	15%

Apresentam certo equilíbrio com relação aos quatro eixos que norteiam a educação, porém, necessita-se dar maior ênfase ao Fazer, considerando a faixa-etária. Essa consideração é reforçada pela avaliação das habilidades, na qual alguns alunos mostram precisar de mais situações de aprendizado no que se refere aos fazeres.

Quando olhamos a *performance geral dos desenhos*, podemos notar um traço de ansiedade no grupo (sobreposição de traços), ao mesmo tempo que aparecem traços de disciplina (respeito ao enquadre fornecido aos alunos). Aqui também podemos perceber pouca presença de movimento (pouca diversidade de cores). Não nos é possível precisar as motivações, apenas indicar que tal fato se soma ao baixo índice dos fazeres.

A organização do material nos conta sobre um trabalho respeitoso com os alunos, bastante individualizado. Seria importante desenvolver um olhar para o grupo, pois esse permite que o professor tenha uma posição inclusiva, reconhecendo o pertencimento de cada um ao grupo.

Jardim da Infância — alunos com cinco anos

O grupo de alunos esteve em atendimento na “Comunidade de Provisão”, durante o ano de 1999, alguns deles desde o ano anterior. Podemos assim pressupor que o “professor-avaliador” conhece os alunos em seu desenvolvimento e que alguns são também conhecidos por outros agentes da instituição (estiveram presentes no ano anterior).

Quanto ao teste *prático*, a performance apresentada parece indicar certa incongruência no que se refere às habilidades (amarrar o calçado e abotoar as roupas), o que pode ser um sinal de não-investimento nesse aprendizado, sendo este apenas avaliado. Isso ocorre, muito provavel-

mente, por um pressuposto de que ele deve se dar em outro ambiente e sob a responsabilidade de outra instituição, a família.

Outro índice significativo refere-se à pouca utilização do diálogo para resolver conflitos, denotando, portanto, um cotidiano exigente de exercício de autoridade.

	Conviver	Conhecer	Ser	Fazer
5 anos	19%	65%	11%	5%

Apresentam certo desequilíbrio em seus índices. A ênfase recai no eixo Conhecer; o Ser e o Fazer aparecem comprometidos por falta de investimento. Quando olhamos as narrativas (impressões, falas e fatos), podemos notar uma forte referência aos familiares, o que adjetiva nossa hipótese relativa à análise dos índices do teste prático.

Com relação ao índice de Conviver, aparecem na narrativa traços que nos indicam que ele foi avaliado tomando-se como parâmetro a relação do aluno com o professor (“ele veio me dizer que...”) e não entre os alunos. Esse fato corrobora a hipótese que formulamos com relação ao teste prático, no qual a emergência de exercício de autoridade acaba produzindo um contra-efeito, que é uma relação marcada pela dualidade professor/aluno, desencadeadora de uma relação conflitiva entre os pares.

Gaberia reordenar o trabalho pedagógico-didático com a finalidade de proporcionar situações em que a produção em grupo — mobilizadora da complementaridade entre os alunos (convivência) — fosse o eixo, em vez do conteúdo da atividade.

A *performance geral dos desenhos* expressa movimento (diversidade das cores), aparecendo, porém, uma certa contenção (sobreposição com pressão nos traços) e certa agressividade/irritabilidade (traços fortes fora do contorno). Essas características reforçam as hipóteses já apresentadas: relação com excesso de individualização; conflitos resolvidos pelo professor e não mediados por ele; poucas atividades que exijam ação grupal e complementaridade entre os membros do grupo.

A organização do material pelo professor denota uma valorização do seu trabalho, porém, com pequena ênfase em seus efeitos na produção dos alunos (escrita sobre os desenhos).

Pré-escola — alunos com seis anos

O grupo de alunos esteve em atendimento na “Comunidade de Provisão”, durante o ano de 1999, apenas duas crianças tendo entrado no segundo semestre. Uma grande parte participou em 1998 também. Podemos assim pressupor que o “professor-avaliador” conhece os alunos em seu desenvolvimento e que alguns são também conhecidos por outros agentes da instituição.

Com relação ao teste prático, podemos observar uma grande heterogeneidade na performance de habilidades do grupo. Essa observação nos causou estranheza, considerando-se o fato de estarem sob os cuidados da instituição há um ano e 48% dos alunos há dois anos. Passamos, portanto, a olhar o material buscando hipóteses que pudessem evidenciar essa heterogeneidade, sem desconsiderar o que pode ser intrínseco à faixa etária.

	Conviver	Conhecer	Ser	Fazer
6 anos	28%	36%	10%	27%

Os índices apresentam certo equilíbrio, porém, o eixo Ser expressa pouco reconhecimento do professor-avaliador em relação aos alunos; as narrativas estão norteadas por um olhar dicotômico — fácil/difícil; saber/não saber –, o que supomos comprometer até mesmo o eixo mais valorizado, o Conhecer.

Os desenhos apresentam movimento (diversidade de cores), uma certa falta de enquadramento (desordenação dos traços), tendo, porém, conteúdos criativos.

A organização do material apresenta um desordenamento nas narrativas e confusão entre essas e os eixos avaliados.

Retomando a estranheza do início da análise e considerando que 48% dos alunos estiveram na instituição no ano anterior, podemos ler a heterogeneidade como efeito dos desequilíbrios anteriores nos índices de cada eixo. Quanto menos situações vivenciamos para desenvolver nossas habilidades, mais passamos a vivenciá-las como dificuldades.

Parece haver necessidade, portanto, de maior instrumentação pedagógica do professor-avaliador, pois muito provavelmente ele acaba oscilando entre a omissão e o autoritarismo em seu cotidiano. Indícios disso já estão presentes no material.

Dados sobre os agentes comunitários

Os agentes comunitários que atuam no Programa Primeiras Letras totalizam 79 pessoas. A avaliação de desempenho desse grupo, realizada pela equipe técnica, sinaliza para um desempenho ótimo de 22% (17 educadores); bom de 43% (34 educadores) e regular de 31% (25 educadores).

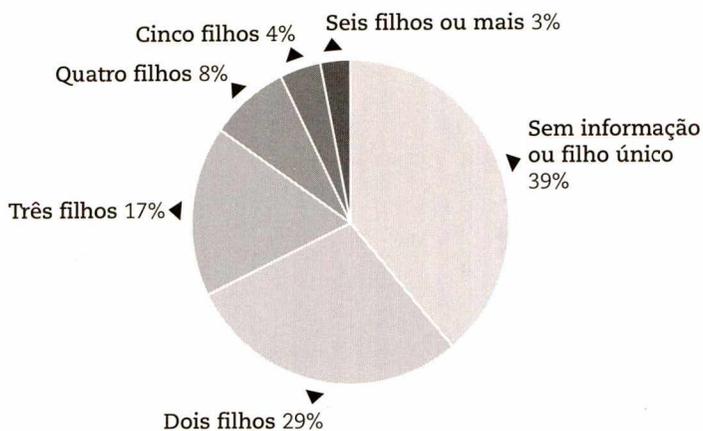
Do total de educadores, dez possuem ou estão cursando o ensino superior; 21 possuem o magistério; quatro o estão cursando; 27 tem o ensino médio completo (incluindo alguns casos de ensino técnico); os demais possuem o ensino médio incompleto (ou estão cursando) e/ou o ensino fundamental completo. Há, ainda, dois agentes com o ensino fundamental incompleto.

Essas diferenças de escolaridade refletem-se no desempenho avaliado pela equipe técnica, ou seja, o nível de escolaridade tem a ver com o desempenho. Nesse sentido, há que se investir na capacitação desses agentes, assim como exigir o 2º grau completo ou magistério, como já vem sendo solicitado pelo MEC.

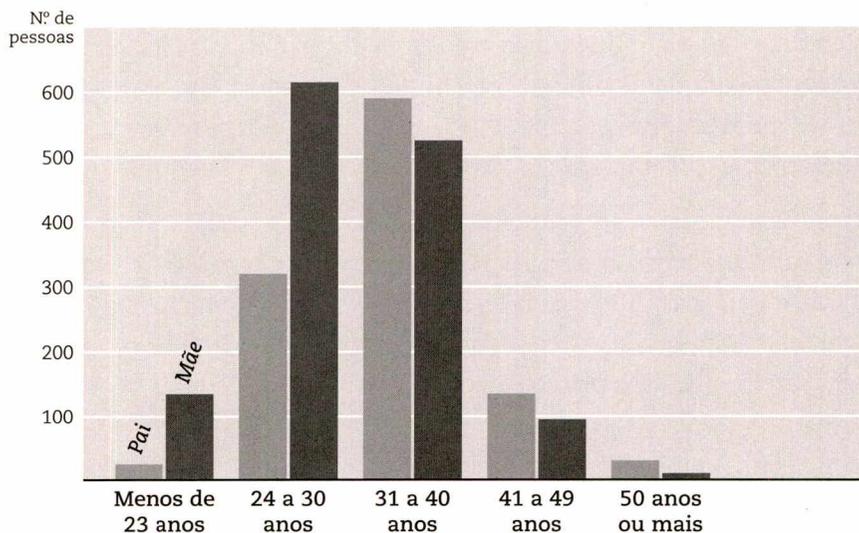
Dados sobre as famílias do Primeiras Letras

Como já constatado no Programa Crê-Ser, indica-se aqui, fundamentalmente, o precário registro (ou desconhecimento) de dados sobre as famílias das crianças. Da mesma forma, é possível inferir que o trabalho restringe-se às crianças, não havendo envolvimento dos pais.

Número de filhos

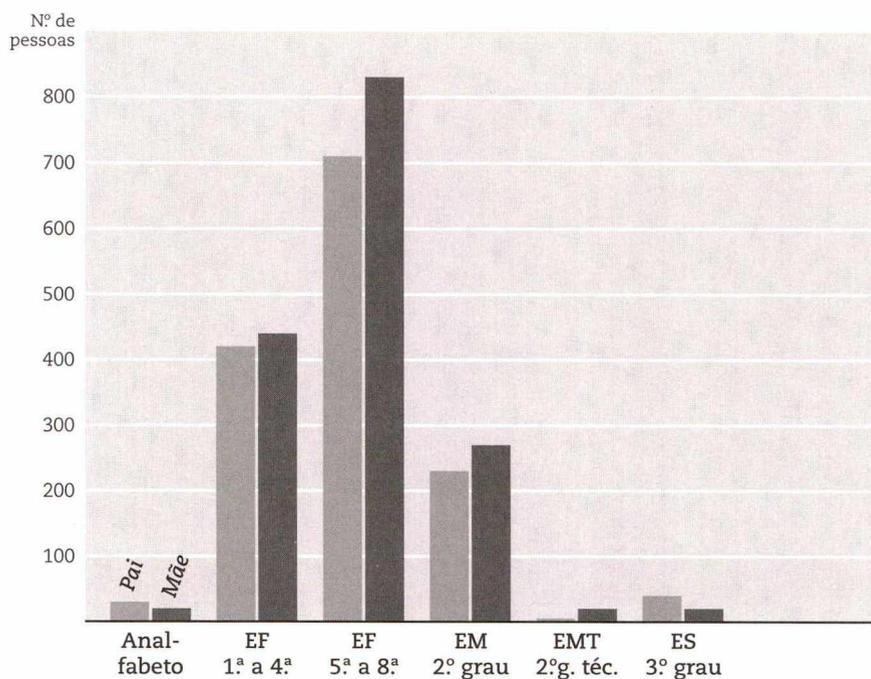


Faixa etária dos pais



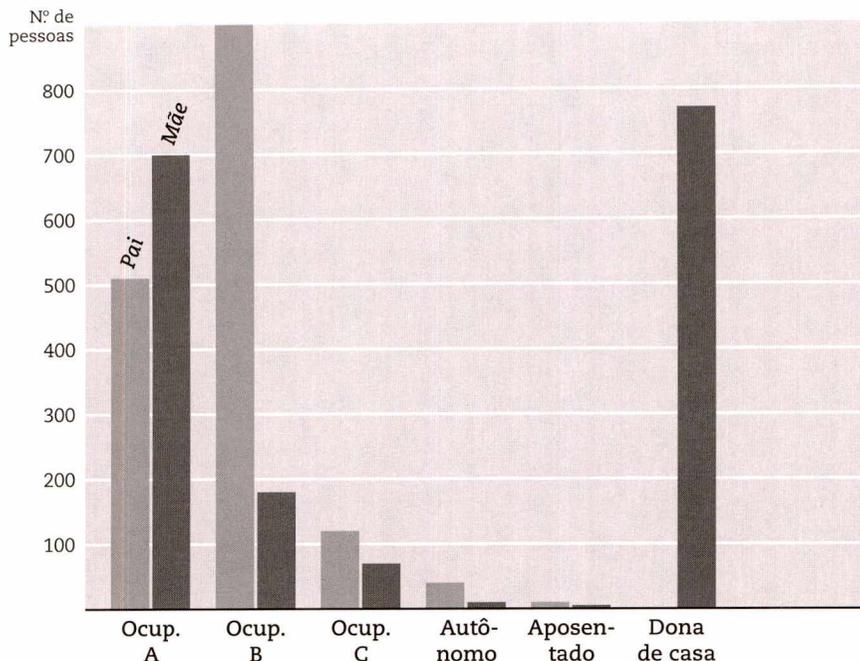
Não há registro de informações quanto à faixa etária de 52% dos pais e 39% das mães. Como era de se esperar, as famílias são jovens.

Escolaridade dos pais



Não há registro de informações quanto à escolaridade de 36% dos pais e 28% das mães. Nota-se certa elevação, se comparada à escolaridade dos pais das crianças do Programa Crê-Ser. Há um número expressivo de pais com 2º grau e até mesmo com ensino universitário.

Ocupação dos pais



Não há registro de informações quanto à ocupação de 30% dos pais e 24% das mães. É interessante notar que a maior escolaridade não alterou ainda a inserção da maioria dos pais, que se encontram no mercado de trabalho exercendo ocupações manuais com ou sem qualificação. É notória, também aqui, a presença de mães “donas de casa”.¹⁴

RECOMENDAÇÕES

1. Aceitar turmas com no máximo 30 educandos no Crê-Ser e 25 educandos no Primeiras Letras. Essa é uma quantidade *máxima* com que um educador consegue trabalhar satisfatoriamente.
2. No caso do Programa Crê-Ser, separar as turmas por faixa etária, ou seja, de 6/7anos a 11 anos incompletos e de 11 anos a 14 anos. Os

interesses e necessidades e mesmo o grau de maturidade/sociabilidade são diferentes para esses dois grupos etários.

3. Realizar um trabalho sistemático e denso com as famílias e com a comunidade, de forma a se tornarem parceiros no projeto de desenvolvimento das crianças e dos adolescentes. É possível assegurar uma participação pró-ativa desses parceiros na condução de atividades programadas em “salas de aula”, e não apenas em eventos festivos. Esse fluxo mais intenso de relações entre famílias, comunidades e programas é necessário não apenas para assegurar padrões de qualidade, mas também para atender à missão da ACB, ou seja, sua maior inserção e ação *na e com as comunidades*.
4. Investir sistematicamente na formação dos agentes. Esta formação envolve necessariamente supervisão *in locu*, com demonstrações concretas dos técnicos na ação educativa com as crianças em “sala de aula”. Propõe-se, no mínimo, uma supervisão semestral *in locu* para cada turma do Crê-Ser e do Primeiras Letras. Essa estratégia é fundamental para redirecionar os programas, a fim de atingirem com maior eficácia seus objetivos. É igualmente fundamental para assegurar um sistema efetivo de monitoramento e avaliação de processos e resultados.
5. Diversificar a equipe técnica, criando maior interdisciplinaridade, com a finalidade de melhor desenvolver as quatro dimensões (Conhecer, Ser, Fazer, Conviver). Hoje, há um predomínio de pedagogos.
6. Na consolidação do processo de avaliação, aferir as atividades socio-educativas desenvolvidas também pela produção seqüencial do educando no processo.
7. Avaliar as necessidades e possibilidades de adequação dos espaços físicos (parte física), assim como o uso de outros espaços complementares ao desenvolvimento das ações requeridas nos programas.
8. Padronizar as informações solicitadas no ato de inscrição de crianças, adolescentes e jovens, de modo a comparar dados interprogramas, bem como alimentar de forma mais satisfatória o planejamento, monitoramento e avaliação. Os dados de inscrição — quando bem formulada a chamada “ficha de matrícula” — permitem um diagnóstico contínuo acerca do público-alvo, ou seja, é ferramenta fundamental na ação.

CONSIDERAÇÕES FINAIS

A ACB tem como missão, desde o seu início, “Promover de forma contínua a qualidade de vida e o autodesenvolvimento de pessoas e de lideranças em comunidades urbanas e carentes”. Com o tempo, foi se afirmando no mercado social como prestadora de serviços à população infanto-juvenil. A mobilização de lideranças e agentes comunitários foi sendo canalizada para essas ações. Nesse sentido, foi deslocando seu foco de atuação: não mais o desenvolvimento de comunidades com base no fortalecimento de suas lideranças e agentes, mas sim o engajamento de tais lideranças e agentes na implementação e desenvolvimento de serviços para o grupo infanto-juvenil.

Não se percebe, no entanto, no plano de ação, um programa estratégico relativo a lideranças e agentes comunitários, missão original da ACB. Este fato parece indicar que as lideranças e os agentes comunitários, para a ACB de hoje, constituem-se em parceiros-chave na realização dos programas Primeiras Letras, Crê-Ser, Lúdico/Cultural e Preparação para o Trabalho. Assim, como parceiros, investe-se em sua competência para que realizem ações voltadas à população infanto-juvenil e não para que ajam diretamente na comunidade, tomando-se como base suas demandas e prioridades.

Desenvolvimento de lideranças comunitárias voltadas para o segmento infanto-juvenil ou para as 32 comunidades onde atua, objetivando a melhoria da qualidade de vida da comunidade? Este, talvez, ainda seja um dilema presente para a equipe de coordenação técnica.

A atual opção não está equivocada, já que crianças e adolescentes das comunidades carentes continuam sendo prioridade social. Nessa opção, a ACB já ganhou competência e visibilidade. E, ainda, sua estrutura organizacional vigente é coerente a tais propósitos.

De qualquer forma, seria recomendável um certo reordenamento institucional, ou seja, *redefinir as equipes de trabalho por agrupamento de comunidades próximas, articulando-as aos programas desenvolvidos*. As equipes de trabalho deverão ter um perfil mais polivalente, pois sua atuação será mais abrangente, envolvendo as ações atuais e outras de interesse ou

necessidade das comunidades e de suas lideranças. Isto significaria novos investimentos em recursos, habilidades e ações.

No momento, os impactos a serem medidos só podem recair na qualidade de vida da criança/adolescente e indiretamente na família. A aferição de impactos na melhoria de vida da comunidade é quase impossível, considerando-se o foco atual das ações da ACB.

As ações hoje realizadas têm, indubitavelmente, impacto indireto na vida da comunidade, mas seria leviano estabelecer uma relação direta, pois existem inúmeras variáveis intervenientes que extrapolam ações voltadas ao grupo infanto-juvenil.

Se as ações da ACB caminhassem na direção do desenvolvimento das comunidades carentes, mesmo que priorizando ações voltadas a crianças e adolescentes, outras ações e prioridades precisariam ser contempladas e, certamente, ganhariam densidade. Nesse caso, outras variáveis poderiam ser monitoradas, permitindo aferir impactos, no tempo, sobre a melhoria da qualidade de vida da comunidade.

NOTAS

- 1 Cf. ARRETCHE, Marta. Tendências no estudo sobre avaliação. In: *Avaliação de Políticas Sociais: uma questão em debate*. São Paulo, Cortez, 1998.
- 2 Segue, anexo, o detalhamento desses instrumentais (roteiros, fichas de observação e testes de desempenho aplicados em crianças do Primeiras Letras, Crê-Ser etc.).
- 3 In: LAHIRE, Bernard. *Sucesso escolar nos meios populares: as razões do improvável*. São Paulo: Ática, 1997. p.17.
- 4 Embora haja uma diretriz da importância da parceria com a escola, esta ação é extremamente difícil de ocorrer nas grandes cidades. Nesse contexto, a ACB tem como nova diretriz priorizar somente as escolas com maior concentração de crianças.
- 5 A coordenação e a equipe técnica da ACB tomaram a decisão de, no ano 2000, atuar com a faixa etária de 7 a 14 anos, separando-a em duas turmas, de acordo com as idades, e oferecendo atividades diferenciadas (teatro, *street dance*, informática, atividades físicas). Para as comunidades com apenas uma turma, será priorizado um ou outro grupo etário, de acordo com a demanda.
- 6 No intuito de observar os impactos do Programa no desempenho escolar, foi solicitado aos agentes/educadores que, — numa amostra casual de quatro crianças, de séries escolares diferentes, consultassem a professora, na escola, quanto ao desempenho escolar desses alunos. O que se pôde observar é que as relações entre os educadores do programa e os professores de escolas são bastante frágeis. Há um desconhecimento mútuo entre programa e escola.
- 7 Como observa Bateson, esses adjetivos são aplicáveis às relações entre o indivíduo e seu meio, material e humano. Ninguém é esperto ou dependente ou fatalista no vazio. Cada traço que atribuímos ao indivíduo não é seu, mas corresponde ao que acontece entre ele e alguma outra coisa (ou alguma outra pessoa). Assim, não é possível rotular ou reificar comportamentos; eles são produtos das condições de coexistência. In: Bernard Lahire, op. cit., p.18.
- 8 Maria Júlia Azevedo e Cláudia Sintoni.
- 9 A equipe está procedendo a um diagnóstico sobre essa questão, a fim de definir as formas de enfrentamento.
- 10 Os códigos de identificação das ocupações são:
 - A: Ocupações manuais sem qualificação — Vigilante, segurança, ajudante-geral, doméstica, arrumadeira, porteiro, guarda, diarista, balconista, ajudante de manutenção, auxiliar de limpeza.
 - B: Ocupações manuais que exigem qualificação — Pedreiro, azulejista, encanador, serralheiro, pintor, cobrador, carpinteiro, polidor, fachadista, açougueiro, cozinheira, operadora de caixa, feirante, mecânico, padeiro, manicure, ajudante de terapia, ajudante de fotógrafo.
 - C: Ocupações técnicas/chefias — Contador, farmacêutico, professora, fotógrafo, gerente de padaria, encarregado, comerciante, motorista, torneiro mecânico.
- 11 Existe apenas uma criança de dois anos.
- 12 Os itens observados para assegurar uniformidade na avaliação foram: usar talheres; colocar e amarrar calçados; colocar e abotoar roupas.
- 13 Maria Júlia Azevedo e Cláudia Sintoni, já citadas.
- 14 Os códigos de identificação das ocupações são os mesmos discriminados na nota 10.

Mariangela Belfiore Wanderley
Denise Blanes

Monitorando a
implantação de
um programa
social

MARIANGELA BELFIORE WANDERLEY é doutora em Serviço Social pela Pontifícia Universidade Católica de São Paulo e pós-doutorada em Sociologia pela École des Hautes Études en Sciences Sociales de Paris. Diretora do Instituto de Estudos Especiais, Vice-coordenadora do Programa de Pós-Graduação em Serviço Social e professora associada do Departamento de Metodologia do Serviço Social da Pontifícia Universidade Católica de São Paulo, é autora de vários trabalhos publicados.

DENISE BLANES é pesquisadora do Instituto de Estudos Especiais — IEE, da PUC-SP. Doutoranda do curso de pós-graduação em Serviço Social da Pontifícia Universidade Católica de São Paulo, é autora de vários trabalhos publicados.

A parceria do IEE/PUC-SP com a Secretaria de Cidadania e Ação Social da Prefeitura Municipal de Santo André desenvolveu-se por meio de um conjunto amplo e diversificado de ações em diferentes níveis, sempre na perspectiva da articulação de programas, projetos e ações no contexto da gestão social municipal.

A direção que orientou as prioridades políticas das ações conduzidas pelo governo municipal de Santo André tem se revelado um campo altamente profícuo para o trabalho que vimos desenvolvendo, considerando-se a perspectiva estratégica de intervenção territorializada, que concentra um conjunto de ações decorrentes de políticas setoriais. Estas buscam, com sua integração, o enfrentamento da pobreza em suas múltiplas expressões, tendo em vista sua redução e a inclusão social da população local.

Essa concepção de ação social, municipal, especializada e integrada não se resume a um conjunto de orientações que visam articular as ações técnicas para melhorar o desempenho do trabalho voltado às comunidades. A perspectiva é a de co-gestão com a população e a ampliação qualitativa da participação. Mediante estímulo a processos de auto-

organização e representação popular, as comunidades locais vão se apropriando dos programas e serviços como direitos sociais, rompendo o círculo de troca de favores por lealdade política, tão enraizado na cultura política brasileira.

Nesse sentido, o conjunto de estratégias¹ mobilizadas pela gestão municipal precisa ser situado nesse contexto, o qual imprime sentido e direção política às ações sociais integradas no campo da *gestão do social*.

O Programa de Renda Mínima de Santo André — PRMSA² é parte dessa estratégia participativa global que orienta a gestão pública municipal, o que permite uma leitura integrada e integral das necessidades do cidadão e das múltiplas faces que constituem a situação de exclusão social.

Quando concebidos numa perspectiva articulada ao conjunto das ações municipais, os Programas de Renda Mínima podem assumir uma função desencadeadora de um processo de emancipação da população, pois abrem espaço para múltiplas inserções em outros programas e projetos, estimulando, ao mesmo tempo, a participação popular em espaços públicos de debate e decisões sobre os destinos da cidade e as alternativas de inclusão social que oferecem aos seus cidadãos.

O desafio de uma gestão social integrada está em impulsionar programas sociais que, pela sua própria natureza, implicam ações conjugadas com outras políticas setoriais, com outras secretarias (Saúde, Educação, Desenvolvimento Urbano e Habitação e Desenvolvimento Econômico e Saneamento Ambiental) e com diversas assessorias e núcleos que desempenham (no caso de Santo André) um papel político articulador e estratégico, como, por exemplo, a Assessoria de Participação Popular (no contexto do PIIS).

Cabe situar essas reflexões no conjunto de transformações que vêm se processando na sociedade brasileira, mais especificamente a partir dos anos 90, para fazer frente ao crescimento das demandas por serviços sociais públicos — indicadoras do estrangulamento do Estado —, que vem exigindo um profundo reordenamento institucional dos organismos governamentais, a incorporação de uma nova cultura pública nas rela-

ções internas entre as diferentes instâncias governamentais e destas com os destinatários das ações públicas. Daí a ênfase em ações integradas intergovernamentais e entre governo e sociedade civil.

A partir da Constituição Federal de 1988, o conjunto de textos legais (LOAS, ECA, LOS, LDB), os quais regulamentam as políticas sociais em seus diferentes setores, vem configurando um novo campo de exigências quanto à incorporação de novos modelos de gestão social, em relação aos quais a estrutura governamental de prestação de serviços sociais se vê impelida a promover a reorganização de seus procedimentos e a redefinição de competências, sempre na perspectiva de consolidar práticas participativas e democráticas de gestão social compartilhada.

O PROGRAMA DE RENDA MÍNIMA DE SANTO ANDRÉ NO PROCESSO DE ACOMPANHAMENTO E AVALIAÇÃO

Um processo de monitoramento e avaliação de Programas Sociais exige a identificação do desenho do programa, das formas de gerenciamento, dos procedimentos metodológicos adotados, do acompanhamento dos resultados e dos diversos fatores intervenientes durante sua implementação.

Assim, serão apresentados, de forma breve, as principais características do programa de Renda Mínima de Santo André — PRMSA, uma síntese histórica de sua implantação e as reflexões relevantes decorrentes do trabalho desenvolvido.

O QUE É O PROGRAMA?

O Programa de Renda Mínima de Santo André — Família Cidadã foi instituído pela Lei nº 7.472, de março de 1997, e regulamentado pela Lei nº 7.567, de 1 de dezembro de 1997.

Tem como **finalidade** a inclusão social das famílias, entendida como resgate das condições de cidadania, concretizada em mudanças objetivas nas condições de emprego e renda e na melhoria dos níveis educacionais, socioculturais e políticos dos membros da família.

São **objetivos** do PRMSA:

a) Objetivo geral:

- Ampliar os níveis de inclusão social dos membros das famílias, no contexto local e da cidade.

b) Objetivos específicos:

- Apoiar economicamente as famílias, num período entre 6 e 18 meses;
- Garantir a permanência das crianças e dos adolescentes na escola e o progressivo sucesso em seus resultados escolares;
- Incluir os jovens e adultos das famílias nos programas de alfabetização e na qualificação e requalificação profissional;
- Incluir os adultos das famílias em programas de geração de emprego, de renda, de proteção e de fomento a formas cooperativadas de trabalho;
- Ampliar os vínculos relacionais da família, aumentando suas trocas culturais e seu acesso a novas informações;
- Fortalecer o grupo familiar, de modo a gerir seu processo de desenvolvimento e inclusão social de forma autônoma.

A **população-alvo do Programa** é constituída de famílias com renda *per capita* inferior a 1/2 salário mínimo por mês; com filhos e dependentes (0 — 14 anos) em situação de risco pessoal e social; que residam em Santo André há três anos — moradoras, no momento da implantação do Programa, das microrregiões definidas para a intervenção.

O Programa prevê a mobilização, articulação e integração de um conjunto de serviços públicos e da comunidade, aos quais as famílias serão conectadas, de acordo com suas necessidades, visando ao seu fortalecimento e permitindo sua inclusão social.

O **gerenciamento** é feito pela Secretaria de Cidadania e Ação Social, à qual cabe apontar as diretrizes, normas, critérios e procedimentos operacionais de acordo com o escopo do Programa.

As **ações articuladas** integram, além da Secretaria de Cidadania e Ação Social, a Secretaria de Educação e Formação Profissional, a Secretaria de Saúde, a Secretaria do Desenvolvimento Urbano e Habitação, a Secretaria de Desenvolvimento Econômico e Emprego, o Serviço Municipal de Saneamento Ambiental de Santo André e o Núcleo de Participação Popular e, ainda, os serviços e equipamentos da sociedade civil.

A **equipe técnica** que trabalha no PRMSA é composta por duplas de profissionais, assistentes sociais e psicólogos, que desenvolvem com as famílias ações de acompanhamento individual e em grupos socioeducativos.

BREVE HISTÓRICO

O processo de implantação do Programa teve início em fevereiro de 1998, após definição das diretrizes da administração da cidade, para o enfrentamento da exclusão social no município. O Programa de Renda Mínima — Família Cidadã configurou-se como uma estratégia para esse enfrentamento, sendo suas concepções, linhas operativas e coor-

denação dadas pela Secretaria de Cidadania e Ação Social, com o apoio consultivo e assessoria do IEE. O processo de implantação desenvolveu-se em etapas complementares:

- discussão e elaboração do desenho do Programa, envolvendo as secretarias de governo da área social;
- treinamento e capacitação da equipe técnica que operaria o Programa;
- preparação das estratégias de operacionalização;
- supervisão e acompanhamento da equipe após o início do trabalho com as famílias.

O primeiro cadastramento teve início em maio de 1998, tendo sido realizado no núcleo habitacional favela Sacadura Cabral, onde o projeto de urbanização e o desadensamento populacional já haviam sido iniciados pela Secretaria de Desenvolvimento Urbano e Habitação. Esta favela localiza-se na zona noroeste do município de Santo André e abriga 780 famílias.

O Programa Renda Mínima foi estendido, progressivamente, para outras áreas de favelas, onde se processavam urbanizações, a saber: Tamarutaca (zona noroeste), Capuava (zona norte) e Quilombo (zona noroeste).

A inserção das famílias de Sacadura Cabral, Tamarutaca, Capuava e Quilombo se deu conforme demonstrado no quadro abaixo:

QUADRO-RESUMO • INSERÇÃO DAS FAMÍLIAS POR ÁREA/MÊS											
Mês de inserção das famílias	Jun/98	Set/98	Abr/99	Jan/00	Mar/00	Abr/00	Maio/00	Total Famílias	Total Pessoas	Total Famílias/Área	Total Habitantes
Sacadura	105	62	0	0	66	0	0	233	1.187	780	3.020
Tamarutaca	0	168	131	0	82	0	84	465	2.369	1.300	5.200
Capuava	0	0	0	77	0	67	0	144	767	1.327	7.000
Quilombo	0	0	0	0	0	0	21	21	104	230	822
Total	105	230	131	77	148	67	105	863	4.427	3.637	16.042

Fonte: Gabinete da Secretaria de Cidadania e Ação Social — PMSA.

A meta de atendimento para 1998 era de 360 famílias, tendo sido efetivamente inseridas 335 no Programa. Para 1999, a meta era de 600 famílias, tendo sido atendidas efetivamente 466. Até junho de 2000, o Programa atendeu 863 famílias.

O orçamento do Programa para o exercício de 1998 foi de R\$ 1 milhão, para o exercício de 1999 foi de R\$ 1,35 milhões, ao passo que a previsão para o exercício de 2000 foi de R\$ 1,3 milhões.

A equipe inicial para a implantação do PRMSA era composta por três duplas de profissionais, formadas cada uma por um assistente social e uma psicóloga. Em junho de 2000, eram cinco duplas atuando junto às famílias das quatro áreas de favelas, desenvolvendo um acompanhamento quinzenal sistemático por meio de atendimentos individuais e em grupos socioeducativos.

O trabalho grupal é uma das marcas do PRMSA, o que lhe garante diferencialidade e qualidade no atendimento. O objetivo é proporcionar um espaço de discussão e reflexão com as famílias a respeito de temas do cotidiano, que despertem e motivem a luta pela melhoria da qualidade de vida, oportunizem possibilidades de acesso a outros programas e serviços, ampliem seus vínculos relacionais e contribuam para o fortalecimento do grupo familiar em seu papel de ir autogestionando seu processo de inclusão social e busca de autonomia.

AS REFLEXÕES SOBRE O PRMSA

As estratégias para monitoramento e avaliação são diversificadas e atendem às demandas específicas do Programa. Vão desde a supervisão sistemática da equipe técnica até encontros públicos de reflexão sobre os rumos, os resultados e os impactos do programa na comunidade local, na relação com outros programas sociais e, ainda, na qualidade de vida de cada indivíduo inserido no PRMSA.

Para refletir sobre o Programa de Renda Mínima de Santo André como estratégia de inclusão social, no contexto da gestão social do município,

e iniciar uma reflexão sobre as possibilidades e os indicadores de emancipação de famílias que foram alvo desse tipo de programa, apresentamos, de forma resumida, pontos relevantes que emanaram dos momentos de supervisão técnica, de apresentações e discussões públicas ocorridas durante o processo de monitoramento.

- a renda domiciliar *per capita* na região do ABC é alta, com um dos menores graus de indigência na região metropolitana de São Paulo;³
- o município de Santo André tem liderança regional, acarretando uma capacidade de expansão das políticas públicas na região do Grande ABC. É preciso, no entanto, que estas sejam articuladas entre as cidades e com o município de São Paulo, para que não haja transferência de demandas, sobrecarregando certos municípios;
- é um programa de baixo custo direto (863 famílias atendidas desde a implantação, a um custo global de R\$ 3,4 milhões por um período de 24 meses);⁴
- com o orçamento atual, o Programa atinge 10% da população indigente do município;⁵ o valor do benefício médio por família (R\$ 218,00) é considerado elevado, quando comparado a outros programas de complementação de renda e/ou de benefício continuado. Essa questão merece estudos, pois, aparentemente, o Programa poderia estar atingindo um número maior de famílias, com aportes menores;
- o custo administrativo do Programa,⁶ *a priori*, é elevado, pois seu desenho prevê um acompanhamento sistemático e intensivo das famílias, acrescido de capacitação continuada dos técnicos;
- o PRMSA é parte do PIIS, o que possibilita uma leitura integrada da cidade e da população que reside nas áreas de intervenção. Permite, também, a integração das estratégias de inclusão, pois, isolado, o PRMSA representaria apenas um “alívio” da pobreza, porém, quando articulado a outras políticas, insere-se no processo conjunto de combate à pobreza; quebra a dicotomia entre a focalização do Programa e a universalidade de acesso;

- é visível a intersecção entre o PRMSA e o PIIS, mas os resultados são ainda modestos. É preciso maior ousadia, com destaque para a questão da escolaridade, da educação de jovens e adultos, incrementando-se a educação fundamental, essencial ao processo de inclusão social;
- a articulação com outros programas e políticas assegura às famílias, durante o processo, mesmo quando desligadas, o acesso à escolarização, saúde, habitação e meio ambiente, geração de renda etc.;
- quando integrado, os programas de renda mínima têm função alavancadora, e, no caso específico de Santo André, tem também função articuladora, pois o trabalho dos grupos socioeducativos acaba se refletindo nos outros programas e no repensar destes. Tal processo leva, de um lado, ao avanço da noção de direito e, de outro, à credibilidade do poder público;
- a participação é uma estratégia do PRMSA e do PIIS, na medida em que se tem a perspectiva de trabalhar em co-gestão com as comunidades; estas se apropriam do conjunto de ações como direito, possibilitando abrir o espaço para outras ações, com a autonomia da comunidade, que vai se gestando no processo;
- o PRMSA tem sua estratégia participativa inserida no processo mais amplo de participação, que transcende o âmbito de cada Programa (plenárias abertas do PIIS, por exemplo). As comunidades vão gradativamente se inserindo num debate mais geral sobre a cidade (Orçamento Participativo, Conferências da Cidade), tendo o PIIS como referência para a reflexão sobre a inclusão social;
- o impacto se faz sentir para além das áreas territoriais de intervenção, visto que tanto o Orçamento Participativo como o Projeto Cidade do Futuro têm recebido a reivindicação da extensão do PRMSA e do PIIS para outras regiões do município. As atuais quatro áreas de intervenção do PIIS/PRMSA representam uma experiência piloto, com base na qual está se gestando uma proposta para toda a cidade, resguardadas as particularidades de cada situação.

A CONSTRUÇÃO DE INDICADORES DE ACOMPANHAMENTO E AVALIAÇÃO DO PRMSA

A importância dos indicadores não está apenas na possibilidade de medir e avaliar os avanços de um dado processo, mas devem também ser um estímulo à reflexão, ao debate, à informação dos sujeitos envolvidos, ou seja, seu próprio processo de formulação deve ser um elemento de transformação das relações existentes.

A participação das populações envolvidas é um importante pressuposto do processo avaliativo. Quanto maior for a apropriação dos indicadores pelos sujeitos implicados no processo, maior será o potencial de transformação e de publicização das ações empreendidas, sobretudo quando se está tratando de políticas públicas.

Nesse sentido, a avaliação é concebida como *produção de conhecimento público*, preservando-se assim o direito à informação e ao conhecimento, por parte dos cidadãos, sobre o uso dos recursos públicos, a eficiência e a eficácia das políticas, os programas e as ações implementados.

A construção de indicadores é eivada de intencionalidades; daí a necessidade de se trabalharem consensos e conflitos entre os diferentes atores (organizações governamentais e não-governamentais, universidade etc.) envolvidos no monitoramento e avaliação. Os processos de monitoramento e avaliação exigem a sistematização das informações sendo cada vez mais, imprescindível a existência de bancos de dados, de sistemas de informações locais e regionais que alimentem esses processos.

A pouca tradição de avaliação das políticas sociais faz o processo de construção de indicadores acontecer de forma mais lenta e no decorrer do acompanhamento.

A questão dos indicadores talvez seja o aspecto mais desafiante no contexto da avaliação e monitoramento do Programa de Renda Mínima de Santo André, exigindo o aprofundamento de estudos e pesquisas qualitativas. Os indicadores foram definidos à medida que o PRMSA foi

se consolidando. Sua construção é um processo que exige familiaridade com a realidade a ser avaliada, maturação e sucessivas aproximações ao longo do trabalho.

Desta forma, alguns dos indicadores não tiveram acompanhamento desde a implementação do Programa, ao passo que, para outros, apesar de serem relevantes, não foi ainda possível o estabelecimento de estratégias de monitoramento. É o caso, por exemplo, do sucesso escolar de crianças e adolescentes das famílias beneficiadas, previsto nos objetivos, que não está sendo acompanhado e mensurado sistematicamente.

A ausência de um sistema informatizado de informações, que permitisse o registro e acompanhamento das famílias de forma contínua e sistematizada, é outra questão que está sendo equacionada, em parte, pelo próprio governo municipal, com a construção do Banco de Dados do Cidadão. Está sendo desenvolvida também uma parceria, com apoio da Fapesp, para a construção de uma ferramenta informatizada que auxilie na gestão do PRMSA, em seus aspectos internos e na sua correlação com os demais programas do PIIS.

Indicadores

Foram criados indicadores para o acompanhamento das famílias, sistematizados em dois eixos:

Dados sociais e econômicos

- composição familiar (nº de membros na família, grau de parentesco, idade, sexo, tipo de família etc.);
- educação: escolaridade (jovens, adultos, crianças, crianças pequenas x creche etc.); cursos profissionais, cursos de aceleração etc.;
- renda;
- qualificação profissional;

- formas de inserção no mercado de trabalho (formal, informal etc.);
- participação em outros programas da PMSA (Geração de Renda, Cooperativas, Mova, Seja, Saúde da Mulher etc.);
- valor do benefício;
- utilização do benefício (medicamentos, alimentação, móveis, utensílios, reformas, construção de moradia, prestações atrasadas, poupança etc.).

Dados qualitativos

Trata-se de dados colhidos em entrevistas, depoimentos, relatórios etc., articulados com dados quantitativos:

- interesses iniciais (projetos de vida, expectativas)
- melhorias na qualidade de vida
 - conforto
 - consumo
 - aparência
- construção de cidadania (âmbito pessoal/profissional/comunitário)
 - discurso
 - participação
 - iniciativa (espírito reivindicador)
- acessibilidade
 - grau de conhecimento dos recursos disponíveis
 - grau de inclusão em outros programas
- condições de auto-sustentabilidade
 - inserção no mercado/renda/outras estratégias
 - empregabilidade (cursos profissionais e escolares, qualificação, auto-estima)

Algumas considerações

As considerações que se seguem são iniciais, uma vez que o processo está em curso:

- A totalidade de crianças e adolescentes envolvidos freqüenta a escola desde início da inserção das famílias no Programa. A escola é um valor para a população em geral e, em especial, para as famílias pobres, que a vêem como uma possibilidade de alteração das precárias condições de vida: *“O Programa ajudou com os meus filhos... voltaram a estudar... [o Programa] empurrou a ir para a frente”*.
- O fortalecimento da auto-estima é a manifestação mais visível e rapidamente perceptível do sucesso do PRMSA. Expressa-se em atitudes e comportamentos dos participantes do Programa: *“Precisava buscar novas alternativas para o futuro... Estou trabalhando na cooperativa... Pude estudar, o marido não queria, não entendia... mas insisti, e hoje já fiz vários cursos”*.
- A concentração de ações num mesmo território, e com a presença de várias secretarias de governo, potencializa o conjunto de ações voltadas para a melhoria das condições de vida. Por exemplo, no tocante à melhoria das condições habitacionais, é comum as famílias pouparem, com o benefício recebido, a fim de comprarem materiais de construção ou pagarem serviços de pedreiro para melhorar as condições da casa. Também no que se refere às melhorias ambientais, atividades relacionadas a coleta de lixo, saneamento básico, higiene, entre outras, contribuem para a melhoria do meio ambiente. A articulação entre essas ações é, no entanto, um processo em construção. O próprio PRMSA contribui para o fortalecimento dessas ações, enquanto cria espaços de reflexão sobre as condições de vida, propicia a aquisição e a troca de saberes. A falta de uma efetiva articulação entre as ações setoriais das diversas secretarias não considera/utiliza plenamente essa potencialidade do PRMSA.

- O PRMSA propicia uma alavancagem para a inserção das famílias em outros serviços e programas da rede de proteção social, mediante levantamento de necessidades e encaminhamento aos recursos existentes. Pelo fato de trabalhar com a noção do direito, cria também novas demandas por parte da população, sobretudo no que se refere aos serviços e, mais remotamente, à participação social.
- Um certo descompasso na implantação do PIIS e de seus programas repercute na dinâmica da aplicação dos recursos financeiros pela população. Assim, por exemplo, se no início do PRMSA a população beneficiada tinha como desejo a inserção em programas de geração de renda (40%), hoje, efetivamente, participam deles 13%, o que não, significa desinteresse da população, mas, provavelmente, pouca oferta. O desejo por atividades de geração de renda continua sendo predominante por ocasião do desligamento das famílias do PRMSA.
- A busca por capacitação/qualificação se expressa de várias formas: pela real inserção dos jovens e adultos nos cursos oferecidos (alfabetização, supletivos e cursos profissionalizantes; pela inserção dos titulares e/ou familiares em atividades de geração de renda e em cooperativas; pelo depoimento dos titulares, durante as entrevistas de desligamento, como um dos ganhos mais importantes da participação no Programa.
- O perfil da população atendida pelo PRMSA evidencia uma multiplicidade de carências, o que exige investimentos de várias ordens (financeiros e econômicos, de apoio psicossocial etc.), num prazo médio de tempo.
- O quadro da situação de trabalho/ocupação dessas famílias demonstra que estão distantes do mercado formal; é possível observar movimentos nesse quadro, mas é certo que a transformação dessa situação não será alcançada pelo Programa, pois trata-se de questões estruturais.
- As conquistas por parte das famílias desligadas são reais, mas não há elementos para afiançar sua persistência no tempo; um acompanhamento longitudinal permitirá observar a consistência dessas aquisições e a pertinência do desenho do Programa.

- As repercussões sobre a qualidade dos vínculos sociais são claramente expressas nas avaliações dos técnicos e dos beneficiados, com reflexo na vida cotidiana. A melhoria das relações conjugais e familiares se deve, em parte, pelo acesso das mulheres (a maioria dos titulares do Programa) aos recursos financeiros, atribuindo-lhes uma nova qualidade em sua posição na configuração familiar. A novidade está, também, na possibilidade das mulheres freqüentarem novos espaços, adquirirem novos relacionamentos e participarem de novas atividades.
- O PRMSA trabalha algumas questões relacionadas à vida societária, à participação comunitária, embora não suficientemente, pois o perfil das famílias e o tempo de permanência no programa exigem que o foco do trabalho permaneça nos aspectos individual e familiar.
- O acompanhamento realizado não permite ter uma visão de conjunto do uso das ações da assistência pelas famílias beneficiadas. Sabe-se, porém, que o Programa de Complementação à Escola atende 56% das crianças e jovens entre 7 e 13 anos, o que é bastante significativo; e que houve uma queda na procura ao serviço de plantão da própria Secretaria de Cidadania e Ação Social.

CONSIDERAÇÕES FINAIS

Apesar do desmonte das políticas sociais, conseqüência das políticas neoliberais, os governos locais têm implantado, cada vez mais, programas de complementação de renda, como o PRMSA, que se apresenta como alternativa a ser ampliada.

A abrangência em que o PRMSA foi implantado, decorrente da opção política do governo local em focalizar o conjunto das ações do PIIS em áreas pilotos, não permite a visualização de um impacto social significativo para o município como um todo, embora sejam patentes os impactos nas condições de vida das famílias atingidas.

É fundamental incrementar a interlocução e a definição de pautas comuns de proposição e intervenção, tendo em vista superar a fragmentação das políticas sociais no município. O fortalecimento de espaços e fóruns de articulação das políticas públicas municipais é um caminho a ser perseguido.

Quanto ao desenho e à gestão dos programas e serviços, a pesquisa identificou que, de modo geral, a Política Municipal de Assistência Social de Santo André estabelece prioridades, desenha estratégias e prioriza os segmentos populacionais típicos da atenção desta área de política pública, estabelecidos pela própria Constituição Federal e pela LOAS.

Há, no entanto, dois desafios a serem enfrentados: a consolidação de um efetivo trabalho em rede, que articule de forma orgânica o conjunto de programas e serviços, tanto privados quanto governamentais; a potencialização da gestão descentralizada e participativa das instâncias da política de assistência social (Conselhos, Fundo e Plano).

São fundamentais, portanto, a definição de objetivos comuns e a adesão a procedimentos normativos e processuais pactuados em esferas públicas democráticas, que articulem as diferentes redes sociais: movimentalis-tas, espontâneas, de serviços sociais governamentais, sociocomunitárias, com a finalidade de construir a rede de proteção socioassistencial à população local.

A questão da construção e consolidação do trabalho em rede é, pois, um dos aspectos centrais a ser focado na análise das estratégias de gestão social, no âmbito das políticas públicas municipais.

NOTAS

- 1 Com destaque para o Programa Integrado de Inclusão Social — PIIS. O PIIS, no qual se insere o Programa de Renda Mínima Família Cidadã, é uma estratégia da Prefeitura Municipal de Santo André, pela qual as políticas setoriais são concentradas espacialmente, visando à redução da pobreza e à ampliação da inclusão social das populações abrangidas.
- 2 Programa de Renda Mínima — Família Cidadã, do município de Santo André-SP, período 1997-2000, gestão Celso Daniel.
- 3 A renda domiciliar *per capita* do ABC paulista é cerca de 12% superior dos níveis encontrados no Brasil metropolitano, e a proporção de pobres é cerca de 38% menor daquela encontrada no Brasil metropolitano. O município de Santo André é considerado de Baixa Indigência Relativa (4,78%). Quando comparados os períodos 1995-1997 e 1992-1993, observa-se uma variação significativa nas taxas de indigência no ABC paulista de 13,34% para 6,49%, enquanto a renda domiciliar *per capita* sobe de R\$ 276,00 para R\$ 378,00. Dados extraídos de André Urani, *Mapa da Pobreza no ABC Paulista: diagnóstico e algumas recomendações de políticas*. Abril, 2000. (Mimeogr.)
- 4 Se tomarmos os valores *per capita*/mês no PRMSA — R\$ 43,00 aproximados — e, por um simples exercício comparativo, cotejarmos com aquele de uma creche no município de São Paulo (R\$ 64,00) ou de um Programa de Erradicação do Trabalho Infantil no país (R\$ 50,00), o PRMSA tem um menor investimento.
- 5 Na área geográfica atendida pelo PIIS, o PRMSA atinge à totalidade da demanda.
- 6 Um estudo preliminar demonstra que o custo administrativo é de aproximadamente R\$ 84,50 família/mês.

Nabil Bonduki

Uma metodologia
para avaliar
programas de
habitação

NABIL BONDUKI, é arquiteto e urbanista, é doutor pela Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo — USP. Professor do Departamento de Arquitetura e Planejamento da Escola de Engenharia de São Carlos — USP, é autor de vários trabalhos publicados. É coordenador-geral do presente projeto.

APRESENTAÇÃO

Este texto apresenta de forma sintética a proposta preliminar de *Metodologia de avaliação dos programas de habitação financiados com recursos do Fundo de Garantia por Tempo de Serviço — FGTS*,¹ que está sendo desenvolvida por uma equipe vinculada à Escola de Engenharia de São Carlos, da Universidade de São Paulo (EESC-USP). Resulta de uma Resolução do Conselho Curador do FGTS, que, objetivando “dar maior transparência às ações empreendidas com os recursos do trabalhador”, determinou que fossem “desenvolvidos estudos visando à definição de uma metodologia a ser aplicada na avaliação dos programas de aplicação lastreados em recursos do FGTS”.

Com esse objetivo, foi desenvolvida uma abrangente metodologia a fim de avaliar os programas Pró-Moradia, Carta de Crédito Associativo, Carta de Crédito Individual, Apoio à Produção e Pró-Saneamento, financiados pelo FGTS, num amplo leque de aspectos, que serão apresentados e conceituados neste texto. Trata-se de uma metodologia de avaliação inédita no país, uma vez que os programas de habitação e saneamento,

ao contrário dos programas sociais (saúde, educação e assistência), não têm sido objeto de avaliações globais, que ficam restritas, se tanto, a aspectos específicos ou a estudos de caso.

Para a EESC-USP, particularmente para o Departamento de Arquitetura e Urbanismo, a elaboração dessa metodologia representou uma possibilidade extremamente rica de realizar um trabalho articulado de pesquisa e extensão universitária, que exigiu a articulação entre a experiência adquirida em pesquisas acadêmicas, particularmente na área de Ciências Sociais Aplicadas e de Tecnologia do Ambiente Construído, como o conhecimento e a reflexão sobre a gestão e implementação de políticas públicas. Foi uma oportunidade de utilização prática do desenvolvimento científico e tecnológico, permitindo que a Universidade pudesse colaborar com o poder público e com os vários segmentos da sociedade, representados no Conselho Curador do Fundo de Garantia por Tempo de Serviço (entidades empresariais e centrais sindicais, além do próprio governo), na avaliação e aperfeiçoamento dos programas de habitação e saneamento.

A formação do Grupo de Acompanhamento do Projeto — composto por técnicos da Caixa Econômica Federal e do Grupo de Assessoria Permanente do Conselho Curador do FGTS, que, ao longo dos quase dois anos de desenvolvimento da metodologia, se reuniu periodicamente para discutir a proposta em cada uma das suas etapas — permitiu, por sua vez, uma profícua troca de experiências e de pontos de vista entre a equipe e os diferentes setores envolvidos e interessados na avaliação e aperfeiçoamento dos programas do FGTS. Além disso, o seminário realizado na Escola de Engenharia de São Carlos, em novembro de 1999, possibilitou que a proposta preliminar fosse debatida com especialistas de todo o país, incorporando-se contribuições importantes para o aperfeiçoamento da metodologia. Assim, pode-se dizer que a proposta apresentada já passou por um processo de ampla discussão e crivo dos principais setores envolvidos com esses programas.

O trabalho realizado exigiu enorme criatividade da equipe envolvida, sobretudo por inexistir no Brasil outras metodologias semelhantes que

pudessem servir de referência. Os poucos trabalhos de avaliação em habitação e saneamento que têm sido desenvolvidos no país, como já foi mencionado, têm se limitado a analisar empreendimentos, especialmente em relação à qualidade do produto, sem contudo inseri-los em uma avaliação de programas, nos seus diversos aspectos. Em outros casos, alguns programas foram avaliados sem que os empreendimentos produzidos fossem abordados conjuntamente.

O desenvolvimento da proposta, que gerou uma complexa metodologia que permitirá avaliar os programas de aplicação do FGTS, mostrou a necessidade de se aprofundar a pesquisa nesta área, pois, para muitos indicadores propostos, inexistem parâmetros de avaliação consagrados. Isto torna necessário o desenvolvimento de estudos para a criação de critérios objetivos, capazes de permitir referências específicas de avaliação para programas de habitação e saneamento. Malgrado este problema, a equipe estabeleceu os parâmetros indispensáveis para realizar a avaliação, deixando claro que a proposta pode ser ainda muito aperfeiçoada por estudos posteriores e após sua primeira aplicação integral.

Caberia ainda destacar que, embora a proposta de avaliação tivesse sido solicitada para os cinco programas supracitados, ao longo de seu desenvolvimento foi ficando clara a necessidade de, para efeito da formulação da metodologia, serem desdobrados alguns dos programas nas modalidades que apresentam especificidades significativas, decisão que foi tomada em reunião com o Grupo de Acompanhamento do Projeto, que defendeu com ênfase esta opção. Assim, o programa Pró-Moradia — PPM foi dividido em dois (PPM-Conjunto Habitacional e PPM-Urbanização de Áreas); o programa Carta de Crédito Individual — CCI, em três (CCI-Aquisição, CCI-Construção e CCI-Cesta de Materiais); e o Pró-Saneamento, em todas as suas modalidades. Esta opção acarretou uma multiplicação do trabalho, pois, embora pudesse haver indicadores comuns aos vários programas, tornou-se necessário analisar cada caso em detalhe, verificando-se a pertinência e os parâmetros, bem como construir propostas específicas para cada um dos indicadores, em todos os programas/modalidades definidos.

Além disso, deve-se ressaltar que a metodologia desenvolvida tem enorme potencialidade e versatilidade, podendo ser adaptada para avaliar outros programas de habitação e saneamento, seja no âmbito do FGTS, seja no âmbito de outras esferas do poder público federal, estadual e municipal. Espera-se que isto possa ocorrer, criando-se uma tradição de avaliação de programas de habitação e saneamento que, sem dúvida, poderá contribuir para uma melhor aplicação dos recursos neste setor.

A IMPORTÂNCIA DE AVALIAR

No quadro de escassez de recursos que caracteriza todas as esferas do poder público no país e, em contrapartida, diante da enorme demanda existente por serviços e equipamentos sociais, o desenvolvimento de um processo contínuo de avaliação das políticas públicas é absolutamente indispensável. Trata-se de um dos principais instrumentos para o planejamento das ações de governo, para a racionalização na utilização de recursos e a revisão das diretrizes e normativas que orientam os programas públicos, podendo garantir um processo permanente de aperfeiçoamento dos programas sociais.

Como afirmam Cohen & Franco (1998), “não é possível que os programas e projetos sociais sejam eficazes e eficientes se não forem avaliados os resultados de sua aplicação. Por isto, dispor de avaliações *ex post* de projetos em curso ou já realizados é fundamental para melhorar os modelos”. Apesar da obviedade desta afirmação, não só no Brasil, como em toda a América Latina, a avaliação dos programas sociais não é frequente, sendo mesmo excepcional; em geral, os atores sociais envolvidos são refratários a avaliar o que realizam. Não cabe aqui discutir as razões dessa resistência, mas sobretudo destacar, neste quadro, a importância da iniciativa do Conselho Curador do Fundo de Garantia por Tempo de Serviço — FGTS em propor a montagem de uma metodologia de avaliação dos seus programas, que este relatório consolida.

No Brasil, como fruto da ausência sistemática de preocupação em avaliar, existe um atraso no desenvolvimento de metodologias adequadas para analisar políticas, programas e projetos sociais, ao contrário do que ocorre em relação à avaliação de custos e benefícios correspondentes a projetos de investimentos, em que predominam critérios economicistas, em geral requeridos pelos órgãos de financiamento. “Ao contrário do que ocorre nos projetos de investimentos, as decisões nesta área [programas sociais] costumam ser tomadas com as melhores intenções de atender às necessidades de uma determinada população, mas geralmente carecem de análise técnica das alternativas que permita selecionar a mais eficaz para a consecução dos objetivos procurados e a mais eficiente na utilização dos recursos que serão destinados a tal programa” (Iglesias, in Cohen, 1998).

Se, de uma maneira geral, inexistente uma tradição de avaliação dos programas sociais como um todo, no que diz respeito aos programas habitacionais a situação é ainda mais grave. Embora no período do Banco Nacional de Habitação — BNH tenha se iniciado um processo de pesquisa e análise dos programas e projetos então implementados, não foram formuladas metodologias sistemáticas de avaliação de programas como um todo. Desde o fim do BNH, as avaliações tornaram-se ainda mais escassas, limitando-se a estudos extremamente gerais, de análise do conjunto da produção, geralmente baseados em dados secundários ou pesquisas limitadas a aspectos específicos, em estudos de caso. Predominam abordagens setoriais e disciplinares, que não permitem avaliações globais e que tratam o problema da habitação e do saneamento de forma fragmentária, como foi mostrado no Relatório 2 deste projeto.

Tal verificação torna a implementação desta metodologia de extraordinária relevância. Trata-se de um trabalho de grande complexidade, pois, como será mostrado adiante, pretende-se avaliar os programas de habitação e saneamento em todos os seus aspectos, requerendo interdisciplinaridade e diferentes métodos de pesquisa e análise. De fato,

raros são os setores de política social que apresentam tantas interfaces como os programas habitacionais e de saneamento.

Por sua vez, a avaliação dos programas propostos — Pró-Moradia, Carta de Crédito Individual, Carta de Crédito Associativa, Apoio à Produção de Moradia e Pró-Saneamento — ganha ainda melhor relevância, na medida em que eles foram criados há cerca de seis anos, no processo de retomada dos financiamentos com recursos do FGTS, depois de uma longa paralisação, tendo sido orientados por novos pressupostos.

Assim, a aplicação desta metodologia permitirá a avaliação dos resultados deste que se caracteriza como um momento distinto das políticas de gestão urbana, aspecto de grande importância, pois pode permitir o aprimoramento da política de habitação e saneamento à luz de uma avaliação objetiva dos programas, identificando fragilidades, inconsistências e avanços. Nesse sentido, é preocupante o fato de três dos cinco programas analisados, embora sejam tão recentes, estarem praticamente paralisados — os programas Pró-Moradia e Pró-Saneamento, em virtude da proibição de contratar com o setor público, determinada pela Política Monetária Nacional, e o programa Apoio à Produção, por não ter logrado aceitação pelo mercado. A interrupção ou descontinuidade desses programas pode, ao menos em parte, frustrar a expectativa do projeto, que visa não só criar um sistema permanente de avaliação, mas, sobretudo, colocá-lo na perspectiva de aperfeiçoar os programas e a política de habitação e saneamento.

DIFICULDADES ENCONTRADAS PARA O DESENVOLVIMENTO DA METODOLOGIA

O desenvolvimento desta metodologia esbarrou numa série de dificuldades decorrentes da amplitude da proposta —, que buscou avaliar os programas em todos os aspectos considerados relevantes —, e, sobretudo, da grande diversidade de tipos de empreendimentos que os programas de aplicação dos recursos do FGTS podem financiar (de uma única unidade habitacional financiada pela Carta de Crédito Individual

a uma estação de tratamento de esgoto). Neste sentido, os novos pressupostos que orientaram a formulação desses programas, que são corretos como diretrizes de política habitacional, ao possibilitarem o financiamento de empreendimentos muito diferenciados entre si geram enorme dificuldade para a formulação de uma metodologia de avaliação, uma vez que criam especificidades que obstaculizam a construção de indicadores capazes de serem observados em todos os casos. Além disso, a própria diversidade de aspectos a serem considerados, que acabam por requerer métodos muito diferentes de pesquisa, criam dificuldades no cruzamento de variáveis e na construção de indicadores.

Determinados aspectos, como *Produto e Impacto na cidade e ambiente*, que têm em cada empreendimento sua unidade básica de análise e exigem uma observação *in loco*, seriam avaliados de modo mais preciso e aprofundado se fossem utilizados métodos como o estudo de caso, que requerem um levantamento mais detalhado e particularizado. Esta opção metodológica foi, no entanto, descartada — decisão tomada em conjunto com o Grupo de Acompanhamento —, pois o objetivo do projeto é a formulação de uma metodologia voltada para a avaliação geral dos programas habitacionais e de saneamento sob análise. Isso exige a construção de indicadores capazes de serem observados numa amostra representativa do universo dos empreendimentos contratados em cada um dos programas de aplicação dos recursos do FGTS. Esta perspectiva impossibilita a utilização do estudo de caso, método que propicia uma avaliação mais aprofundada de cada empreendimento, mas que não pode ser aplicado, por razões de ordem operacional, prática e de custos, num grande número de empreendimentos.

Por razões semelhantes, no aspecto *Impacto nas condições de vida dos moradores*, a avaliação do Projeto Social desenvolvido nos empreendimentos dos programas Pró-Moradia e Pró-Saneamento, modalidade Pró-Sanear, ficou prejudicada, uma vez que sua observação dependeria de entrevistas em profundidade e de análises qualitativas, métodos que não poderiam ser aplicados em um grande número de empreendimentos, representativos do universo dos contratos desses Programas.

Diante dessa impossibilidade, vários indicadores foram sendo alterados ou eliminados ao longo da elaboração da metodologia, toda vez que não era possível encontrar alternativas generalizáveis e indicadores, que só poderiam ser construídos por meio de estudos qualitativos.

Apenas na avaliação da Instância Colegiada foi proposta a utilização de entrevistas em profundidade e análises qualitativas, pois o fato de existir apenas uma instância por Unidade da Federação não permite o uso de métodos quantitativos, ao passo que uma observação de profundidade, mais específica de cada instância, é o melhor método para observar as questões a serem avaliadas.

CONCEPÇÃO GERAL DA METODOLOGIA

HISTÓRICO DO DESENVOLVIMENTO DA PROPOSTA

O Termo de Referência, elaborado pelo Grupo de Assessoria Permanente do Conselho Curador do FGTS e que deu origem a este projeto, propunha que a metodologia de avaliação a ser desenvolvida “deveria necessariamente abordar, pelo menos, os aspectos institucional, social, econômico e urbanístico/ambiental”. Ficava estabelecido, também, os Programas que seriam objeto de avaliação, a saber: Pró-Moradia (Programa de Atendimento Habitacional por intermédio do Poder Público); Pró-Saneamento (Programa de Saneamento); Programa Carta de Crédito Associativa; Programa Carta de Crédito Individual e Programa Apoio à Produção.

Com base na experiência acumulada em avaliações de programas sociais e na literatura especializada, a equipe da EESC-USP apresentou, em outubro de 1998, uma perspectiva de avaliação que, embora não deixasse de responder a todos os aspectos e dimensões sugeridas pelo GAP-FGTS, alterava substancialmente a lógica original do Termo de Referência. Assim, sugeriu-se uma avaliação que enfocasse as dimensões do *processo*, do *produto* e dos *impactos* dos Programas a serem avaliados. Tratava-se de

uma opção metodológica bastante adotada em avaliações semelhantes e que diferia significativamente da maneira como o Termo de Referência formulado pela CCFGTS propunha, preliminarmente, o desenvolvimento da avaliação, em que prevalecia um recorte disciplinar (institucional, social, econômico, físico etc.). Malgrado esta modificação, todos os aspectos previstos no Termo de Referência estavam presentes, embora não necessariamente agrupados da forma proposta de início.

Com o início do desenvolvimento da metodologia propriamente dita, em maio de 1999, buscou-se verificar quais deveriam ser, especificamente no caso dos programas em análise, os aspectos em que essas dimensões poderiam ser subdivididas. Mediante um complexo processo analítico, procurou-se identificar quais os aspectos gerais que deveriam ser avaliados, considerando-se a grande especificidade e a diversidade de modalidades presentes em cada um dos programas. Essa reflexão gerou uma grande quantidade de aspectos específicos e de indicadores.

Após um intenso processo de depuração, marcado por uma ampla pesquisa bibliográfica, maior reconhecimento dos programas, muita reflexão e exaustivas discussões — que contaram em várias oportunidades com consultores externos e com o Grupo de Acompanhamento do Projeto —, foi formulada uma primeira proposta preliminar de metodologia de avaliação, debatida em seminário com especialistas na Escola de Engenharia de São Carlos, em novembro de 1999. Foram então selecionados os aspectos gerais mais relevantes e os indicadores mais adequados para a avaliação. Naquele momento, muitos indicadores considerados importantes foram excluídos, pela dificuldade de serem construídos ou por requererem métodos de pesquisa, como o estudo de caso, incompatíveis com a dimensão da amostra e com a perspectiva de se avaliar de modo integral os programas de aplicação do FGTS.

Após essa primeira etapa de elaboração, as três dimensões estabelecidas originalmente (processo, produto e impactos) foram subdivididas em oito aspectos gerais, que permitiam uma avaliação mais precisa dos programas:

- Economicidade;
- Procedimentos de gestão;
- Desempenho operacional;
- Produto;
- Impacto nas condições de vida dos moradores;
- Impacto na cidade e no ambiente;
- Impacto macroeconômico;
- Atendimento às necessidades de habitação e saneamento.

Na versão da metodologia debatida no seminário, foram apresentados os indicadores referentes a cada um desses oito aspectos gerais, mas não sua parametrização e ponderação. O rico debate possibilitou o avanço na proposta, gerando, entre outros, a junção dos aspectos gerais *Procedimentos de gestão* e *Desempenho operacional*, além de um recorte dos indicadores selecionados.

Definida a primeira versão da estrutura básica de avaliação, com o estabelecimento dos indicadores e variáveis, foram elaboradas os instrumentos de coleta visando ao pré-teste da metodologia. De acordo com a proposta original, o pré-teste foi dividido em duas etapas, sendo a primeira no Estado de São Paulo, voltada fundamentalmente para o teste dos instrumentos, da estratégia geral da pesquisa de campo e do levantamento dos dados documentais, e a segunda nos estados da Bahia, Goiás e Distrito Federal, quando se testaria o conjunto da metodologia.

O pré-teste, realizado em São Paulo entre os meses de janeiro e fevereiro de 2000, em cerca de dezoito cidades, foi aplicado pela equipe responsável pela elaboração de metodologia, tendo sido muito proveitoso tanto para testar os instrumentos de coleta como para permitir um melhor conhecimento dos empreendimentos financiados pelo FGTS e das informações documentais existentes nos Escritórios de Negócios da Caixa. Esta etapa foi também fundamental para dimensionar as equipes de campo e o tempo necessário para o levantamento; para definir a qualificação básica dos

entrevistadores, além de permitir uma completa revisão dos instrumentos a fim de adequá-los à caracterização e diversidade dos empreendimentos; e para superar dificuldades de aplicação, como linguagem, fluxo etc.

O segundo pré-teste foi precedido de um planejamento mais cuidadoso, buscando-se superar os problemas encontrados em São Paulo e possibilitar a aplicação dos instrumentos por uma equipe externa, na perspectiva de reproduzir as condições que seriam encontradas numa aplicação sistemática da metodologia. No caso do pré-teste da Bahia, estabeleceu-se uma cooperação com a Faculdade de Arquitetura da Universidade Federal da Bahia, para a formação e treinamento da equipe responsável pelo levantamento. O mesmo se deu com a equipe que realizou o teste em Goiás e no Distrito Federal, embora a maior parte de seus membros fossem alunos da Escola de Engenharia de São Carlos e de outras universidades de São Paulo, treinados antes de serem deslocados para a área. Pôde-se, assim, testar não somente os instrumentos reelaborados após o primeiro pré-teste como sua utilização por equipes externas, que não haviam participado da formulação da proposta, bem como perceber a complexidade do planejamento, que se faz necessário para que as várias frentes de coleta funcionassem com eficiência, e também experimentar os demais passos da metodologia, incluindo a digitação e seu processamento e análise.

Este segundo pré-teste foi extremamente útil para o avanço e o desenvolvimento da metodologia. Além de testar novamente os instrumentos, aplicados por equipes externas em empreendimentos bastante diversificados e distintos dos visitados em São Paulo, foram estabelecidas rotinas detalhadas para o trabalho de pesquisa. Ademais, a digitação dos instrumentos permitiu que se obtivessem os primeiros resultados sistematizados da avaliação, representados pelo processamento das variáveis simples.

Outro resultado expressivo foi a sistematização das informações básicas e o registro fotográfico de um grande número de empreendimentos habitacionais, cerca de 54 unidades da Federação (incluindo o teste de São Paulo), o que permitiu visualizar com concretude as características e a diversidade dos produtos gerados pelos programas financiados pelo FGTS.

O principal desdobramento do pré-teste foi, contudo, a elaboração definitiva da estrutura básica da metodologia de avaliação. Nesse momento, já com o total conhecimento das características e dos produtos gerados em cada programa, iniciou-se a etapa mais difícil, demorada e importante de todo o projeto, quando foram definitivamente definidos os indicadores de cada aspecto geral, sua parametrização e ponderação, a fim de gerar a avaliação dos programas. Além disso, nesse momento, foi necessário estabelecer quais seriam as unidades (programa/modalidade) a serem avaliadas. Dada a grande diversidade existente no interior de cada programa, chegou-se a sete, na área da habitação, mediante desdobramento dos programas Pró-Moradia em dois e Carta de Crédito Individual em três.

Após de uma primeira proposta de ponderação dos indicadores para a avaliação dos programas, gerou-se um intenso processo de discussões entre a equipe e o Grupo de Acompanhamento do Projeto, visando-se analisar o conjunto da metodologia e estabelecer a parametrização de cada um dos indicadores e sua respectiva ponderação na avaliação dos aspectos gerais.

Essa série de reuniões teve como resultado final a estrutura definitiva de avaliação, apresentada neste texto, e permitiu formalizar a construção dos indicadores, com base nos indicadores simples e variáveis, para cada um dos programas/modalidades a serem avaliados. Após a definição final dos indicadores, dos indicadores simples e, conseqüentemente, das variáveis, pôde-se processar a revisão final dos instrumentos de coleta dos dados e os respectivos manuais e dicionários.

O longo processo de elaboração da metodologia e sua discussão com todos os segmentos envolvidos permitem dizer que se alcançou um elevado grau de consenso entre todos os atores presentes no processo, que inclui, além da equipe, representantes das Centrais Sindicais, das entidades empresariais e do poder público, em particular da Caixa Econômica Federal. Ademais, as inúmeras alterações que foram realizadas entre o segundo pré-teste e a versão final tornam necessário — como não poderia deixar de ser numa proposta tão complexa como a que se

formulou — que a primeira aplicação seja realizada para, além de avaliar os programas, contribuir na revisão final da metodologia.

A ESTRUTURA BÁSICA DA METODOLOGIA DE AVALIAÇÃO

A metodologia formulada parte do pressuposto de que se pretende realizar uma avaliação *ex post* e permanente, implementada por pesquisadores externos às instituições envolvidas no financiamento e promoção dos empreendimentos. Essa avaliação objetiva analisar os programas de aplicação dos recursos do FGTS como um todo, nas diferentes dimensões pertinentes, priorizando um olhar que contemple a conjunto dos resultados alcançados por programa, sem aprofundar a avaliação específica e particular de cada empreendimento financiado.

Para alcançar tal fim, estabeleceram-se, para cada um dos sete programas/modalidades habitacionais, os aspectos e indicadores a serem avaliados e o processo de construção dos indicadores, tomando-se por base os indicadores simples e as variáveis obtidas por instrumentos de coleta dos dados aplicados para empreendimentos selecionados numa amostra representativa dos programas/modalidades.

Após um longo processo de reflexão, optou-se por adotar o aspecto como unidade máxima de agregação na avaliação de cada programa/modalidade, pois se considerou artificial estabelecer para os programas uma avaliação global única, que permitiria obter uma só nota ponderando-se notas atribuídas a aspectos tão diferentes. Assim, os programas seriam avaliados em cada um dos sete aspectos, ou seja:

- Economicidade;
- Procedimento de gestão;
- Produto;
- Impacto nas condições de vida dos moradores;
- Impacto na cidade e no ambiente;
- Atendimentos às necessidades;
- Impacto macroeconômico.

Cada programa teria, a princípio, sete notas, cada uma expressando o resultado relativo a um aspecto. A avaliação de cada um deles, que resulta numa nota final, é resultado das notas ponderadas obtidas em cada um dos indicadores que os compõe. O quadro, que será apresentado mais adiante, mostra os indicadores relacionados com cada aspecto, por programa, e a ponderação proposta.

Além dessa avaliação geral por aspecto, indispensável para permitir uma leitura mais global e clara de cada programa, a metodologia permite conhecer todas as notas dos indicadores (que são no máximo 27), além dos indicadores simples e até mesmo das variáveis que geraram a nota final. Assim, ela possibilita tanto construir séries históricas por aspectos e indicadores como observar questões mais específicas, transformando-se num instrumento de gerenciamento para identificar problemas e soluções.

A nota de cada indicador é, por sua vez, resultado de um complexo processo de construção, podendo, conforme cada caso, ser resultado de complexas operações matemáticas processadas com base nas notas obtidas pelos indicadores simples, variáveis compostas e/ou variáveis. Essas notas expressam o resultado obtido mediante parametrização das variáveis. Esta construção matemática e conceitual dos indicadores e dos indicadores simples é a base metodológica para a criação do modelo informatizado, que permite, a partir dos resultados obtidos em cada variável coletada nos instrumentos, chegar às notas dos indicadores e aspectos.

Os instrumentos de coleta dos dados serão apresentados posteriormente, de modo sintético. Salvo os dados coletados diretamente no Banco de Dados da Caixa e do roteiro a ser aplicado para a avaliação da Instância Colegiada, que tem uma lógica específica, de uma maneira geral os instrumentos são aplicados para cada empreendimento selecionado na amostra representativa dos programas/modalidades (caso da ficha de processo e projeto, de observação das áreas urbanizadas ou do entorno, entre outros). São aplicados também em uma amostra de moradores ou de unidades habitacionais de cada empreendimento se-

leccionado (caso de ficha de observação das unidades habitacionais e do questionário ao beneficiário final). Isso significa que o plano amostral deve estabelecer uma amostra de empreendimentos a serem pesquisados e, no interior destes, uma outra referente a unidades habitacionais e a moradores a serem observados e entrevistados.

A amostra de empreendimentos para cada programa/modalidade deve ser calculada partindo-se de um universo de empreendimentos finalizados e ocupados no mínimo um ano e no máximo três anos antes do início do período em que se pretenda processar a avaliação. O prazo de um ano foi considerado o mínimo para que os novos moradores pudessem conhecer razoavelmente suas novas condições e responder com consistência às questões necessárias à avaliação, e também para que a situação real da moradia, do ponto de vista construtivo, pudesse ser detectada. Por sua vez, o prazo máximo de três anos foi estabelecido por não ser conveniente alargar o tempo entre a pesquisa e a ocupação do imóvel, posto que um dos aspectos que se pretende avaliar — *Impacto nas condições de vida dos moradores* — utiliza-se, mediante questionário ao beneficiário final, da memória dos moradores a fim de obter informações sobre suas condições de habitação no momento anterior à mudança. Concluiu-se que um prazo superior a três anos colaboraria para tornar mais imprecisas essas informações.

Quanto à escolha da amostra a partir da entrega e ocupação das unidades habitacionais, isso fará que os indicadores relativos aos contratos de financiamento tenham datas distintas, dependendo do momento em que foram feitos. Essa questão foi considerada na montagem dos indicadores, eliminando-se o risco de distorções.

Apesar dessa orientação geral, propõe-se que a primeira aplicação, a ser realizada no primeiro semestre de 2002, inclua todos os empreendimentos finalizados e ocupados desde o início desses programas, em 1995, até 31 de dezembro de 2000. A partir daí, a avaliação se repetiria no primeiro semestre dos anos ímpares, avaliando um período de dois anos.

CRITÉRIOS PARA PONDERAÇÃO E PARAMETRIZAÇÃO DA AVALIAÇÃO

O estabelecimento de ponderação para definir a avaliação dos indicadores e aspectos é central na metodologia e envolve questões de caráter técnico e subjetivo, vinculadas a valores, resultando de uma leitura apurada dos objetivos e diretrizes dos próprios programas, quando eles explicitam (o que raramente ocorre) com clareza seus propósitos, de modo a apontar a importância de cada indicador na avaliação dos aspectos.

Após a definição da composição dos aspectos e dos indicadores, um dos maiores problemas do trabalho foi o estabelecimento desses critérios de ponderação, assim como a parametrização dos indicadores. Essa questão assume contornos ainda mais complexos em decorrência da diversidade de métodos de pesquisa, das disciplinas envolvidas na avaliação e da necessidade de estabelecer critérios e parâmetros nacionais comparáveis num país de dimensões continentais e de diversidades evidentes.

A importância desse aspecto é inquestionável, posto que são os parâmetros (estabelecidos para cada indicador, simples ou variável) que definem, em última instância, a avaliação do programa.

Os critérios de parametrização utilizados foram os seguintes:

- critérios estabelecidos pelos próprios programas;
- critérios definidos “tecnicamente”, como as normas técnicas estabelecidas por associações técnicas (como a ABNT); os padrões internacionais criados por entidades como a Organização Mundial da Saúde; normas criadas pela legislação;
- critérios comparativos, como os estabelecidos na relação com outros países e os definidos na relação custo/benefício, confrontados com outras possibilidades de investimento correlato;
- critérios empíricos ou estatísticos, como os endógenos ou os estabelecidos por padrões estatísticos, correlações etc.;
- critérios subjetivos definidos pela opinião dos moradores, técnicos ou mesmo pelo bom senso.

São inúmeros os problemas advindos da utilização desses critérios, mesmo quando eles são estabelecidos por instituições técnicas irrefutáveis. Muitas vezes, esses padrões definidos em nível internacional são facilmente questionáveis, estando em desacordo com as condições nacionais; critérios técnicos irrefutáveis não são facilmente adaptáveis à grande diversidade regional presente no país. No entanto, esses ainda são problemas relativamente simples de serem enfrentados. Uma questão ainda mais complexa é a referente aos indicadores que não têm nenhum critério pré-estabelecido, exigindo uma definição da própria metodologia de avaliação. O caráter inovador da proposta, incluindo muitos indicadores e até mesmo aspectos que jamais foram objeto de avaliação, tornou essa tarefa muito difícil. Ao mesmo tempo abriu, como já foi afirmado, novas frentes de pesquisas, que talvez possam, brevemente, ser objeto de estudos específicos, que permitam aprofundar a análise de parâmetros para os indicadores criados por essa metodologia.

A definição de parâmetros para indicadores é um processo complexo, que pode demandar longos períodos de pesquisa específica, o que não é compatível com a formulação de uma metodologia geral de avaliação tal qual a desenvolvida neste trabalho. Espera-se que, gradativamente, essa metodologia possa ser aperfeiçoada incorporando novos referenciais que venham a ser desenvolvidos, num processo contínuo de aprimoramento metodológico.

ASPECTOS E INDICADORES A SEREM AVALIADOS, CONCEITUAÇÃO E VALORES

Tomando-se por base as dimensões de avaliação estabelecidas inicialmente (processo, produto e impacto), foram estabelecidos, após um longo processo de elaboração e discussão, os sete aspectos a serem avaliados em cada programas/modalidade, apresentados no quadro a seguir. Esses aspectos, como foi visto no item anterior, permitem a máxima agregação de indicadores, recebendo uma nota própria que identifica a situação do programa em relação a cada um deles. Isso possibilita — se

a avaliação for executada a cada dois anos, como se propõe — a elaboração de séries históricas de avaliação de cada um dos sete aspectos, verificando-se em que medida as iniciativas tomadas pelos órgãos responsáveis pela formulação e implementação desses programas estão gerando os efeitos previstos.

Para cada um dos aspectos foram definidos os valores a serem considerados na avaliação, de essencial importância uma vez que foram eles que determinaram não só a definição final dos indicadores como os critérios de parametrização que orientaram sua construção e, em consequência, sua nota. Embora, por vezes, os valores apresentados possam parecer óbvios, sua explicitação é fundamental em qualquer processo objetivo de avaliação, permitindo identificar com clareza os princípios que orientaram a formulação da metodologia e a definição da parametrização, elementos básicos ao se avaliar.

A seguir, são apresentados os aspectos gerais e os valores correspondentes:

Aspecto	Valores
Economicidade	Sustentabilidade econômico-financeira. Eficiência. Melhoria da relação custo/benefício.
Procedimento de gestão	Agilidade. Eficácia. Satisfação do usuário. Participação. Descentralização. Transparência.
Produto	Qualidade. Conforto. Racionalidade. Sustentabilidade. Satisfação do usuário.
Impacto nas condições de vida dos moradores	Melhoria na qualidade de vida, respeito aos direitos sociais e à cidadania e participação.
Impacto na cidade e no ambiente	Integração e compatibilidade com diretrizes urbanas. Sustentabilidade. Melhoria do entorno.
Atendimento às necessidades	Abrangência. Distributivismo. Eficácia. Desempenho.
Impacto macroeconômico	Geração de emprego e renda.

As definições de cada um destes aspectos, assim como dos indicadores e indicadores simples que os compõem, serão apresentados a seguir. O quadro também procura sintetizar a estrutura geral de avaliação.

Economicidade

O aspecto *Economicidade* é entendido como a capacidade dos programas em gerar, mobilizar e gerenciar os recursos aplicados em operações de crédito, garantindo seu retorno de forma adequada. Os valores que orientam a avaliação desse item são: sustentabilidade econômico-financeira, capacidade de agregação de recursos, efetividade dos recursos investidos e distributivismo.

Esse aspecto é composto pelos seguintes indicadores: *solvência financeira*; *recursos complementares*; *efetividade dos recursos investidos*.

O indicador *solvência financeira* busca verificar a real capacidade de liquidez na administração econômica dos contratos firmados com os mutuários, poder público, pessoa física ou jurídica. O indicador *recursos complementares* trata da capacidade de agregar recursos imobilizados e viabilizados pelo financiamento do programa. Baseia-se no princípio de que a possibilidade de agregar recursos das mais diversas fontes é elemento fundamental, pois otimiza o investimento, questão que se destaca ante a escassez de recursos disponíveis para investimentos em programas de habitação e saneamento. Foi utilizado como critério principal na parametrização do indicador a eficiência em conseguir maiores contrapartidas, além da mínima estabelecida na resolução. O universo de análise são os contratos da amostra.

O terceiro indicador de Economicidade, *efetividade dos recursos investidos*, é o único que utiliza a avaliação de um outro aspecto como indicador simples, ou seja, é composto pelo aspecto *Produto* e pelo indicador simples *valor médio por família atendida*. Busca aferir a relação entre o valor médio de investimento por família atendida e a avaliação dos produtos gerados. O valor médio de investimento é endógeno, construído com base nos empreendimentos da amostra selecionados.

A construção desse indicador baseou-se em Cohen & Franco (1993), no que se refere à análise custo-efetividade, considerada mais adequada aos impactos dos projetos sociais. Essa análise compara benefícios e custos de um projeto, condiciona a sua aceitação ao fato dos benefícios superarem os custos e incorpora o fato de que os benefícios dos programas sociais não se expressam necessariamente em moeda, colocando a comparação entre custos, a relação custos/benefícios, bem como os produtos e impactos resultantes.

Considerando-se a escassez de recursos disponíveis e a dimensão das demandas existentes, optou-se por valorizar o princípio de sua racionalização e otimização, ou seja, valorizar a busca de redução nos custos praticados, mantendo como valor, neste indicador, o distributivismo dos recursos relacionado com a qualidade do produto construído ou obtido.

Procedimentos de gestão

Na avaliação do aspecto *procedimentos de gestão* procurou-se aferir se os propósitos, as estratégias e a execução do programa estão sendo realizados de acordo com as metas estabelecidas. Trata-se de verificar a adequação entre meios e fins, com a utilização de critérios de eficácia e agilidade.

Avaliar a gestão de um programa, no âmbito de uma política pública, implica deter-se em uma dimensão específica da montagem da política avaliada, ou seja, implica focalizar “o processo”, como o definiu Figueiredo & Figueiredo. Para esses autores, avaliar o processo é aferir se os propósitos, estratégias e execução do programa estão sendo realizados segundo os propósitos pré-definidos.

Os indicadores que compõe este aspecto são: *desempenho operacional; participação e gestão do empreendimento e contrato; Instância Colegiada.*

Na análise do *desempenho operacional*, foi empregado o conceito de eficácia funcional, em que se verifica a adequação dos meios aos fins previstos. Os

critérios dos parâmetros adotados foram endógenos, isto é, foi levada em conta a situação presente hoje nesses programas, embora tenha sido levantada a importância de estudos comparativos para a definição de parâmetros mais precisos. Foram utilizados valores como a agilidade na tramitação, cumprimento de prazos e tempo das obras. Desse modo, procurou-se avaliar os resultados quantitativos e financeiros e a capacidade dos programas em atender ao universo dos proponentes.

Além desses valores, foram incorporados critérios qualitativos, expressos no indicador *participação e gestão do empreendimento e contrato*, que verifica a satisfação do usuário e sua participação no processo de gestão, devendo esta ser considerada apenas nos programas cujos objetivos a incluam. Os direitos do consumidor, valor de extrema relevância na atualidade, foram também incluídos na avaliação dos modelos de gestão dos programas, compondo este indicador.

A incorporação de critérios valorativos — como satisfação do usuário, sua participação na gestão e os direitos do consumidor na avaliação dos procedimentos de gestão — introduz uma dimensão qualitativa e abre um viés analítico para que sejam questionados os fins estabelecidos, sem que se fique limitado à verificação da adequação dos meios a eles.

Com base nessa conceituação, definiram-se os indicadores que medem se a estrutura montada pelo programa é eficaz, se é capaz de gerar resultados pretendidos e como o faz, isto é, qual o modelo adotado para se chegar aos objetivos, procurando-se responder a duas questões: O programa é capaz de gerar resultados compatíveis com a disponibilidade do FGTS para contratações? Qual o modelo adotado para isso: centralizador ou participativo, democrático ou ditatorial, impõe regras ou valoriza a relação com o usuário/cliente?

Para responder a essas duas questões, foram montados os indicadores: *desempenho operacional, participação, gestão do empreendimento e exercício do direito do consumidor*, que valem para todos os programas; *avaliação da Instância Colegiada*, que se restringe aos programas destinados ao setor público, ou seja, Pró-Moradia e Pró-Saneamento.

O indicador *desempenho operacional* avalia a capacidade da instituição para responder, de forma ágil e eficaz, às necessidades dos usuários. Considera-se que, desta forma, o potencial de utilização dos recursos é maximizado e que, além disso, minimiza-se o dispêndio de energia necessária para que os pretendentes ao financiamento possam alcançar seus objetivos.

Pelo indicador *participação, gestão do empreendimento e exercício do direito do consumidor* busca-se observar os principais elementos que identificam o modelo de gestão, adotado para os diferentes programas. Não há um único modelo de gestão na medida em que existem várias linhas de ação (programas), com o objetivo de atingir diferentes públicos e abarcar uma ampla parcela da demanda por habitações.

Esse indicador contempla a questão da participação, seja na concepção do projeto, seja na gestão do empreendimento e na dinâmica das relações compactuadas entre agente financeiro e usuário, que garantem seus direitos como consumidores. Com essas questões, abrangem-se três momentos distintos de uma política habitacional no que e refere às relações entre beneficiário e agente promotor: o *antes*, por meio de questões relacionadas à definição do projeto, quem participou e quem decidiu o quê; o *durante*, sobre como se deu a fase de tramitação (com acompanhamento ou sem) e de obras (com trabalho ou não, com acompanhamento ou não); e, por fim, o *após contrato*, quando o beneficiário se transforma em mutuário e cliente.

Apesar das várias conotações que a questão da participação tem tido desde os anos sessenta, neste caso, no âmbito da avaliação de gestão, participação refere-se à possibilidade de controle e acompanhamento de todo o processo por parte do futuro morador e à interferência na definição das principais características de sua unidade habitacional.

Nos programas em que a participação não faz parte dos objetivos, essa questão não será avaliada, por mais que esse conceito goze hoje de um certo consenso quanto à sua importância. Inclui-se, nesse caso, por exemplo, os programas Carta de Crédito Individual, cujo objetivo é abrir uma linha de crédito para aquisição de um imóvel produzido pelo mer-

cado ou então adquirir materiais de construção a serem utilizados pelo próprio morador. É também o caso do programa Apoio à Produção, destinado à empresa construtora. Nesses casos, não é possível avaliar a participação, pois os próprios programas não propõem tal prática. O mesmo se coloca em relação à avaliação do trabalho social, que somente é requerido no Pró-Moradia e na modalidade Pró-Sanear.

Por fim, a avaliação do *exercício do direito do consumidor* coloca, para cada programa, a necessidade de regras claras, de compromissos que devem ser respondidos por ambas as partes, de respeito pelo usuário. Essa postura implica um novo comportamento por parte dos promotores dos empreendimentos e dos gestores dos programas: implica garantir informações corretas, transparência nas relações com o usuário, na capacidade de propor alternativas e na responsabilidade quanto à produção de um bem de qualidade.

O indicador referente à avaliação da *Instância Colegiada* relaciona-se com a necessidade de se verificar em que medida a seleção entre os vários candidatos aos recursos é realizada de forma transparente, democrática, uma das principais questões pertinentes à gestão dos programas.

Nos programas destinados ao setor público, essa seleção se dá de forma descentralizada, por intermédio da Instância Colegiada, conselho estadual presidido pelo governador e com representantes dos municípios, agrupados segundo categorias, e membros da sociedade civil. Entende-se que a participação colegiada é positiva e que a maior pluralidade e representação de seus membros permitiriam uma melhor adequação entre as decisões tomadas e as necessidades da comunidade-alvo.

Produto

O *Produto* é o resultado concreto e palpável dos programas habitacionais, necessitam ser avaliados com base em de alguns valores que excedam a sua simples construção. Nesse sentido, os indicadores que compõem a avaliação desse aspecto — *habitabilidade* (urbana e da uni-

dade habitacional), *espacialidade* e *construtibilidade*, além da *satisfação do usuário* — agregam as dimensões valorativas da qualidade do projeto e da obra, em termos arquitetônico-urbanísticos, do conforto da moradia, da sustentabilidade das soluções técnicas, da racionalidade construtiva que o produto manifesta, além de sua inserção urbana e ambiental.

○ indicador referente à *habitabilidade da unidade habitacional* diz respeito ao conjunto de aspectos que interferem na qualidade de vida e comodidade dos moradores, bem como na satisfação de suas necessidades físicas, psicológicas e sociais. Os conceitos mobilizados foram: conforto ambiental luminoso (nível de iluminância e fator de luz diurna), conforto ambiental térmico (transmitância térmica, atraso térmico e fator de calor solar), conforto ambiental acústico (nível de isolamento sonoro), segurança do usuário (estrutural e relativa a incêndios, acidentes atmosféricos e utilização) e salubridade, aqui se incluindo estanqueidade (a gases e líquidos), higiene e condições atmosféricas. Em relação aos conceitos de conforto ambiental, deverão ser cruzados os dados objetivos levantados com a percepção do usuário.

○ indicador *habitabilidade urbana* parte do pressuposto de que moradia é entendida em seu sentido amplo, conjugando-se ao direito à cidade. Essa avaliação deve verificar em que medida a habitação responde a uma preocupação dessa ordem, de estar inserida na malha urbana, baseada em sua relação com a rede de infra-estrutura e a possibilidade de acesso aos equipamentos públicos. Soma-se a isso o estudo da implantação urbana dos empreendimentos, mediante análise da acessibilidade física, das barreiras arquitetônicas e das patologias e problemas construtivos que podem se manifestar no conjunto do empreendimento.

Quanto à *espacialidade* — em que são analisados os aspectos que determinam a configuração físico-espacial de um edifício ou assentamento habitacional, da infra-estrutura e dos equipamentos e mobiliário urbanos —, os conceitos empregados foram: implantação do edifício, diversidade tipológica, acessibilidade arquitetônica (ao edifício e do edifício), ergonomia e atendimento à legislação vigente. Nesse item, os aspectos

estéticos, que deveriam aqui ser considerados, revelaram-se de difícil apreensão nos levantamentos de campo realizados a título experimental e, portanto, também difíceis de serem considerados de modo mais aprofundado na metodologia de avaliação, tendo-se em vista que se trata de fatores marcados por forte subjetividade (gosto e preferências pessoais e culturais, padrões da moda etc.).

A construtibilidade refere-se ao conjunto dos aspectos construtivos de um edifício ou assentamento habitacional, da infra-estrutura, equipamentos e mobiliário urbanos, que garantem sua viabilidade, dos pontos de vista tecnológico, ambiental e econômico-financeiro. Refere-se tanto às qualidades técnico-construtivas do edifício ou assentamento habitacional quanto à consideração de práticas socioculturais locais ou regionais.

Os conceitos aqui empregados foram os de tipologias construtivas, em que se verifica o tipo de sistema construtivo adotado, isto é, seu estágio tecnológico (artesanal, tradicional, racionalizado, pré-fabricado e industrializado), para cada subsistema construtivo (fundação, estrutura, vedações, cobertura, instalações hidráulicas, sanitárias e elétricas); racionalização do produto, que inclui modulação, padronização, normatização e flexibilidade; o conceito de durabilidade, abrangendo estado de conservação e vida útil dos subsistemas; o conceito de respeito às práticas socioculturais locais ou regionais, em que são avaliados a disponibilidade de materiais locais, a utilização de materiais regionais, as peculiaridades da mão-de-obra e o grau de dependência tecnológica; o conceito de eficiência técnico-econômica, dado pela relação entre o custo de construção da unidade e o valor do financiamento, e pela produtividade, isto é, o número de unidades por valor de financiamento.

O último indicador deste aspecto é a avaliação da moradia pelo usuário, que permite um contraponto à avaliação técnica, que predomina nos demais indicadores. Sabe-se que a avaliação do usuário apresenta um alto grau de subjetividade e se altera ao longo do tempo, segundo fatores muitas vezes totalmente independentes do objeto avaliado. Apesar disso, é de grande importância considerar a opinião dos moradores;

sendo um indicador, permite tanto uma leitura isolada como agregada com os demais indicadores técnicos, gerando a nota global de *Produto*.

Impacto nas condições de vida dos moradores

A avaliação no campo social objetiva apreender os múltiplos efeitos e a produção de impactos imediatos e mediatos na vida dos cidadãos. O objetivo é o de aferir mudanças na qualidade de vida da população envolvida em determinado programa ou ação pública.

Condições de vida da população é uma expressão ampla que abarca todos os aspectos da vida social de um grupo, em um determinado contexto histórico, político, econômico e cultural. O conceito está intimamente ligado, por um lado, a uma visão de qualidade de vida e, por outro, à identificação de necessidades sociais e indicadores de pobreza. Tem sido utilizado em pesquisas e estudos como uma categoria genérica, que deve ser qualificada de acordo com o objeto e o objetivo do trabalho a ser realizado. Significa que é fundamental definir um determinado recorte, de acordo com o estabelecido no projeto de avaliação.

Nesta metodologia, *Impacto nas condições de vida dos moradores* é entendido como as mudanças geradas pelos programas de habitação e saneamento nas condições de vida dos beneficiários, tanto no que se refere à situação posterior à implantação do empreendimento como na comparação com a situação anterior. Envolve indicadores de mudança nas condições de moradia, na obtenção da propriedade e estabilização da posse, nos gastos com moradia e nas condições de sociabilidade, participação e organização do grupo social beneficiado pelos programas de aplicação dos recursos do FGTS.

Os valores que servem de referência para a avaliação deste aspecto são: *melhoria na qualidade de vida; respeito aos direitos sociais e à cidadania; participação.*

O principal desses valores, que norteia a construção da maior parte dos indicadores de impacto nas condições de vida dos moradores, é a me-

lhoria na qualidade de vida, entendida como um conjunto de aspectos que indicam mudanças qualitativas no cotidiano de vida da população, dentro de um determinado contexto histórico, político, econômico e cultural. Assim, é de fundamental importância obter dados sobre a situação anterior ao empreendimento, ou seja, ter a informação de quais eram as condições de vida antes da implantação dos programas habitacionais e de saneamento, para que se possa comparar e aferir os impactos.

O conceito de qualidade de vida envolve dois elementos: a qualidade propriamente dita das condições de vida do homem e do meio ambiente; a democratização dos acessos ao usufruto dos serviços sociais produzidos pela sociedade. Significa a possibilidade de desenvolvimento de um ambiente saudável, participativo e ecologicamente sustentável; bem como de construção da cidadania, como extensão e consolidação de direitos sociais.

Neste sentido, consideram-se como *Impacto nas condições de vida dos moradores* as mudanças que ocorrem após a implantação dos programas no que se refere a condições de moradia, estatuto de propriedade do imóvel, condições socioeconômicas das famílias e mudanças na sociabilidade, participação e organização do grupo social envolvido na intervenção. Busca-se detectar as modificações significativas, positivas ou negativas, geradas neste conjunto de aspectos por determinado empreendimento habitacional e de saneamento, com base na comparação do que se convencionou chamar de antes e depois (A/D) do programa ou ação pública.

A proposta metodológica de avaliação de impacto busca analisar as alterações de uma situação anterior em relação a uma posterior, como também detectar mudanças ocorridas posteriormente, mas que não necessariamente podem ser comparadas com uma situação anterior. A comparação dos dados de diferentes momentos possibilita apreciar os avanços e retrocessos no período considerado e verificar as mudanças ocorridas.

Observa-se que o conjunto de indicadores apresentados pode ou não se aplicar a um determinado programa ou modalidade, dependendo de suas características e especificidades.

○ indicador *mudanças nas condições de moradia* parte da compreensão de que o ato de morar não se encerra na casa utilizada, mas abrange o estar na cidade, usufruindo os serviços que ela propicia na sua diversidade, conformando no morador um sentimento de pertencimento, segurança e cidadania. Neste sentido, o conceito envolve os aspectos que dizem respeito diretamente à unidade habitacional e também os aspectos relativos ao acesso à cidade.

Ressalte-se que o impacto nas condições de saúde do morador — cuja avaliação direta foi excluída por dificuldades de levantamento dos dados e de isolar o impacto gerado especificamente pelo programa daqueles provocados por outras intervenções neste campo — pode ser verificado indiretamente por meio deste indicador, particularmente no que se refere a densidade de ocupação e conforto da unidade habitacional, instalações sanitárias, abastecimento de água e esgotamento sanitário, ou mesmo pelo acesso aos serviços de saúde, todos integrantes deste indicador. Vários estudos demonstram que as condições de moradia e de saneamento têm forte impacto nas condições de saúde dos moradores.

○ indicador *impacto da obtenção da propriedade e da posse* refere-se à mudança da condição de propriedade e de posse, como na mudança da situação de inquilino para proprietário ou na passagem da situação de irregularidade para a regularidade (no caso de favelas/empreendimento).

○ indicador *mudança nos gastos mensais com moradia* está diretamente relacionado com as alterações nas condições socioeconômicas no que diz respeito à composição da renda familiar e às possibilidades de sustentação financeira da família. A mudança nos gastos mensais com moradia expressa-se pela variação no gasto mensal com prestação ou aluguel do imóvel, tarifas e taxas de água e esgoto, energia elétrica, condomínio, IPTU e a variação do percentual da renda familiar referente ao gasto com moradia.

○ indicador *condições de sociabilidade, participação e organização* refere-se às diversas formas de relação social estabelecida entre os moradores de determinada localidade, seja internamente ao conjunto, seja na vizi-

nhança do entorno, que possibilitam o desenvolvimento pessoal, familiar e do grupo social e que garantem a sustentabilidade social do empreendimento em questão.

Cabe destacar que, especificamente, a avaliação referente a mudanças na participação e organização aplica-se nas modalidades dos programas habitacionais e de saneamento que envolvem um grupo de população e/ou que colocam a exigência de um projeto social, como é o caso do Pró-Sanear, Pró-Moradia, Carta de Crédito Associativa. No caso do Programa Apoio à Produção, Carta de Crédito Individual e demais modalidades do Pró-Saneamento, este indicador não têm aplicabilidade, dada a natureza das ações implementadas.

O indicador *avaliação do morador sobre suas condições de vida* busca captar a avaliação subjetiva do morador sobre suas condições de vida, nos aspectos relativos a relações familiares, relações de vizinhança, condições de saúde, condições de acesso ao trabalho e condições de segurança. Esta avaliação se dá pela comparação da situação anterior com a atual, para cada aspecto, identificando-se a percepção do morador quanto à melhora, piora ou manutenção da situação anterior.

Impacto na cidade e no ambiente

Neste aspecto, são avaliadas as mudanças relativas às condições urbanísticas, à dinâmica imobiliária e à qualidade ambiental. Trata-se de uma avaliação de impacto *ex post*, realizada após a ocupação do empreendimento habitacional pelos mutuários, no seu entorno.

São considerados os valores que dizem respeito às melhorias de interesse coletivo no entorno, à sustentabilidade ambiental e à adequação a diretrizes de política urbana e ambiental. Para tanto, foram utilizados os seguintes conceitos: *impacto urbanístico*, em que se considera a localização do empreendimento em relação às diretrizes e normas de ocupação urbana, estabelecidas pela administração municipal, e à disponibilidade de infra-estrutura e serviços urbanos no entorno; *geração de investimento*

público no entorno, em que se verifica a melhoria das condições de infraestrutura e de serviços urbanos decorrentes da implantação do empreendimento; *mudanças na dinâmica imobiliária do entorno*, incluindo-se aí a valorização dos imóveis e a indução de novos empreendimentos imobiliários; *interferência nas condições ambientais*, que leva em conta a adequação à legislação ambiental vigente e a alteração das condições dos corpos d'água e de saneamento.

Para relacionar a avaliação de impacto na cidade e no ambiente às diferentes situações urbanas em que se implantam os empreendimentos habitacionais, utiliza-se o conceito de potencial de impacto, dado pela relação entre o porte do empreendimento (população moradora ou número de unidades) e a população urbana do município onde este se localiza, bem como pela disparidade entre o padrão do empreendimento e o padrão das edificações do entorno.

Os indicadores que foram utilizados para avaliar este aspecto são: *atendimento à legislação*; *geração de investimento*; *mudança na dinâmica imobiliária no entorno*; *interferência nas condições ambientais do entorno*; *potencial de impacto*.

O indicador *atendimento à legislação* avalia a localização do empreendimento em relação às diretrizes e normas de ocupação urbana, estabelecidas pela administração municipal e estadual, assim como em relação à disponibilidade de infra-estrutura e serviços urbanos.

No caso do Programa Pró-Moradia, para as modalidades que impliquem remoção de população de áreas impróprias para ocupação (áreas de risco, de proteção a mananciais, não previstas para ocupação urbana) acrescenta-se, à avaliação de impacto urbanístico, a porcentagem da população removida em relação ao total da demanda a ser removida da área de origem da população beneficiária do empreendimento. Esse total deverá considerar a população removida para outros empreendimentos, sempre que o empreendimento em processo de avaliação não previr a remoção da totalidade da população da área de origem.

O indicador *geração de investimento* avalia os investimentos públicos decorrentes da implantação do empreendimento em relação ao padrão de

urbanização do entorno, que é caracterizado pelas condições de infraestrutura, de equipamentos urbanos, de acessibilidade, do tratamento dos espaços públicos e pela disponibilidade de comércio e serviços.

O indicador *mudança na dinâmica imobiliária no entorno* avalia a variação ocorrida na procura, na venda e no preço dos imóveis nas proximidades do empreendimento, além da atração de novos empreendimentos imobiliários após sua implantação.

O indicador *interferência nas condições ambientais do entorno* avalia a localização do empreendimento em relação ao estabelecido na legislação ambiental e as mudanças ocorridas nas condições do solo, dos corpos d'água e de saneamento do entorno, após sua implantação.

O indicador *potencial de impacto* visa dar parâmetros à avaliação de impacto na cidade e no ambiente, em razão das diferentes situações urbanas em que se inserem os empreendimentos financiados. A definição da maior ou da menor possibilidade de mudanças serem geradas por um empreendimento em seu entorno é estabelecida com base no porte e no padrão do empreendimento em relação à sua localização.

Atendimento às necessidades habitacionais

Nesse aspecto geral, pretende-se avaliar a capacidade dos programas em atender ao perfil da demanda, constituída pelas famílias com necessidade de moradias e de saneamento básico. Com este objetivo, espera-se estabelecer relações entre os resultados alcançados pelos programas e as diferentes categorias de demanda que podem ser identificadas e agregadas. Nessa perspectiva, os indicadores que compõem esse aspecto são: *consecução das metas físicas; atendimento ao perfil do déficit; atendimento às solicitações e permanência das famílias atendidas.*

A mais ampla categoria de demanda que pode ser definida é aquela representada pela quantificação do déficit de moradias e de saneamento, que se manifesta de forma diferenciada segundo regiões do país e faixas de renda. Esse aspecto — atendimento ao déficit ou mais pro-

priamente impacto no déficit — constitui o horizonte mais palpável a ser perseguido pela Política de Habitação e de Saneamento. Ademais, o déficit é uma das variáveis que compõem a equação para a distribuição dos recursos entre as Unidades da Federação, devendo servir ainda como baliza importante para avaliar os programas.

Entende-se que a avaliação do impacto produzido no déficit resultante da produção de uma certa quantidade de unidades implementadas pelo conjunto dos programas torna-se um dos principais indicadores de desempenho da política. Considerou-se, no entanto, que não é objetivo do FGTS atender ao conjunto do déficit, posto que se trata de um recurso oneroso e que não responde exclusivamente pela política de habitação e saneamento. Desse modo, optou-se por avaliar apenas em que medida os programas atendem ao *perfil do déficit*, de acordo com as resoluções do CCFGTS. Na perspectiva, este indicador é composto por dois indicadores simples: *atendimento ao perfil do déficit até três salários mínimos* e *atendimento ao perfil do déficit até cinco salários mínimos*.

Outro indicador que permite avaliar os resultados dos programas é a *consecução das metas* estabelecidas para cada um deles. Estas quantificam, ano a ano e em um prazo mais extenso de tempo, os recursos e os objetivos físicos que se pretende alcançar. Ao serem estabelecidas, dimensionam a contribuição que determinado programa pretende dar para a eliminação do déficit. Obviamente, essa avaliação só seria possível se anualmente fossem estabelecidas as metas orçamentárias e físicas, por programa, o que se propõe que seja adotado.

A relação entre metas e resultados, aspecto usual de avaliação de políticas públicas, coloca-se aqui como uma das análises sobre a capacidade que os programas estabeleceram para atendimento das necessidades da demanda. É, desta forma, um indicador de desempenho dos programas, na medida em que estabelece um balanço entre as propostas (metas) e a efetiva realização (resultados). Assim, o único indicador simples relacionado com este indicador é a *relação entre metas físicas e resultados*.

O indicador *atendimento às solicitações* permite apontar em que medida a demanda propriamente dita do FGTS, que se identifica como a quantidade de interessados em obter financiamento para um dos programas, é atendida pelo programa. Este conjunto da demanda não é atendido em sua totalidade por vários motivos, entre os quais incapacidade de responder aos quesitos técnicos exigidos, renda inadequada etc.; de qualquer forma, a relação entre os resultados e demanda identificável dos programas cria outro indicador de desempenho, complementar aos demais. O indicador simples que compõe este indicador é a *relação entre solicitações e resultados*.

O indicador *permanência das famílias atendidas* trata da mesma problemática — atendimento às necessidades de habitação —, mas partindo de um outro ponto de vista: em vez de avaliar a entrada nos programas, busca verificar em que medida o programa tem sido capaz de evitar que o esforço de atender a uma família se perca pela conhecida e intensa rotatividade que atinge comumente as famílias beneficiadas por financiamentos habitacionais

Impacto macroeconômico

A metodologia de avaliação do *Impacto macroeconômico* — que apresenta grande especificidade em relação à avaliação dos demais aspectos — foi desenvolvida procurando atender a dois objetivos complementares. Em primeiro lugar, fornecer informações que fossem relevantes para as decisões de alocação dos recursos do FGTS pelos seus gestores; em segundo lugar, estimar as variações em agregados macroeconômicos que tivessem relação com as condições de reprodução da classe trabalhadora.

A metodologia para avaliação do impacto macroeconômico foi **deseenhada, portanto, para capturar, em primeiro lugar, o impacto que o investimento realizado com os recursos do FGTS ocasiona no mercado de trabalho. Atendia-se à condição de relevância para a classe trabalhadora, quantificando-se um importante efeito que o investimento com mora-**

dia e saneamento provoca. O segundo critério, a importância para as decisões alocativas dos gestores dos recursos, foi contemplado ao se desenhar uma metodologia de quantificação do emprego gerado que identificasse esse resultado por programa e por modalidade das aplicações do FGTS.

A possibilidade de estimar o emprego por programa e modalidade permite que se avaliem alternativas de composição dos investimentos, adicionando-se uma nova dimensão à análise das metas e dos valores alocados. Permite, também, que se quantifiquem, por agregação, o emprego associado ao total de recursos aplicados e sua contribuição à criação e manutenção de postos de trabalho na economia como um todo. A metodologia adotada possibilita, ainda, que se obtenha, como um subproduto relevante, a estimativa da receita tributária associada às aplicações realizadas.

A natureza macroeconômica das modificações no mercado de trabalho a serem mensuradas exige a utilização de metodologia capaz de capturar, simultaneamente, a totalidade dos efeitos que ocorrem como contrapartida aos investimentos. Vale dizer, torna-se necessário quantificar variações na utilização de todos os fatores de produção ocasionadas por essa despesa exógena. A metodologia disponível mais adequada é, consensualmente, a estimativa, mediante a utilização da matriz de insumo produto, do emprego direto, indireto e induzido gerado. Quantificam-se, assim, emprego e receita tributária originários das atividades diretas de produção de habitações e programas de saneamento, da oferta adicional de insumos e serviços dos equipamentos, bem como a demanda de bens finais ocasionada pela elevação do valor adicionado total.

A utilização da matriz de insumo produto, para estimativa dos efeitos diretos, indiretos e induzidos no mercado de trabalho, constitui metodologia já consolidada, que vem sendo aplicada aos diversos setores da economia brasileira.

Os cálculos das variações de emprego e receita tributária até agora disponíveis na literatura, entretanto, referem-se sempre aos grandes setores da economia brasileira, entre eles o setor da Construção Civil. A de-

finição desse setor, segundo a classificação do IBGE, agrupa atividades bastante distintas como, por exemplo, construção, reforma, ampliação, reparo de edificações e montagem de estruturas, de pré-moldados, inclusive plataformas flutuantes. Como consequência, os estimadores de emprego e receita tributária obtidos para o setor resultam em um valor representativo de um conjunto de atividades bastante diversas, com tecnologia, relação capital-trabalho e produtos bastante diferenciados. Devem diferir, portanto, dos multiplicadores associados ao conjunto de produtos viabilizados pelos recursos do FGTS, edificação de habitações populares e programas de saneamento.

Para atingir o objetivo do trabalho — fornecer estimadores de emprego e receita tributária que sejam a contrapartida das atividades e produtos viabilizados pelo FGTS —, foi desenvolvida uma metodologia de cálculo dessas variáveis, utilizando-se a matriz de insumo produto elaborada pelo IBGE e os protótipos dos produtos de cada programa: Pró-Moradia, Carta de Crédito Individual, Carta de Crédito Associativa, Apoio à Produção e Pró-Saneamento, em suas diferentes modalidades.

Entende-se por protótipo os produtos representativos de cada uma das modalidades dos cinco programas do FGTS; sua concepção é traduzida pela descrição da quantidade e da qualidade dos fatores de produção e insumos utilizados para sua obtenção, pressupondo, portanto, tecnologia dada e constante. As unidades-padrão são obtidas aplicando-se o vetor de preços de mercado dos insumos e os fatores a cada protótipo, gerando-se, assim, os orçamentos de produção. Para cada uma das modalidades dos programas existentes será definida, desta forma, uma ou mais unidades-padrão, representando os produtos financiados sob estas rubricas. Definem, no seu conjunto, um setor de atividades produtivas, específico e restrito aos programas do FGTS.

A matriz de insumo produto constitui uma extensão da teoria da interdependência das atividades e decisões econômicas, representando a economia como um todo. Uma matriz de insumo produto-padrão contém colunas e linhas dispostas em forma de um tabuleiro de xadrez,

em que cada linha e respectiva coluna representa um particular setor de atividade: indústria automobilística, têxtil, gráfica, construção civil e todos os demais. Com base nesta representação, é possível obter os coeficientes técnicos de produção de cada setor, bem como a interdependência entre todos os setores da economia.

Buscando-se na matriz de insumo produto, elaborada pelo IBGE com os dados das Contas Nacionais, obtêm-se os multiplicadores de emprego e receita tributária. A matriz para o Brasil tem, como linhas e colunas, os setores de atividade também definidos pelo IBGE, em que o setor de Construção Civil tem o caráter abrangente já mencionado anteriormente. Cada um dos setores pode ser desagregado em um conjunto de produtos a ele associado.

As unidades-padrão permitem definir um novo setor de atividade, definido como o conjunto de produtos utilizados para a obtenção das edificações e programas de saneamento do FGTS. Como consequência, são obtidos os coeficientes técnicos para esses produtos e sua inter-relação com todos os demais setores da economia. Podem-se, desta maneira, estimar de forma precisa os multiplicadores de emprego e receita tributária associados às unidades-padrão, evitando-se os desvios das quantificações obtidas mediante os multiplicadores da Construção Civil.

Os multiplicadores, quantificam, portanto, o emprego e a receita tributária que a produção de cada uma das unidades-padrão irá provocar. São números que se mantêm inalterados enquanto são julgadas válidas as duas hipóteses centrais na definição da metodologia: tecnologia utilizada na construção dos produtos inalterada e representatividade do universo de produtos pelas unidades-padrão.

O emprego efetivamente gerado pelas aplicações realizadas pelo FGTS consiste no produto dos multiplicadores das unidades-padrão pelo total de recursos alocado em cada uma das modalidades. O somatório do emprego gerado pelas modalidades de cada programa representa o emprego e a receita tributária associados a cada um deles, que, agregados, representam o emprego e receita tributária gerados pelo total das aplicações do FGTS.

INSTRUMENTOS DE COLETA DAS INFORMAÇÕES

Os instrumentos de coleta dos dados necessários para a implementação desta metodologia são consequência da sua concepção geral, ou seja, dos aspectos e dos indicadores que se pretende avaliar. Para obter as variáveis necessárias para a construção dos indicadores, foram criados diferentes tipos de instrumentos de coleta, adotando-se vários métodos de pesquisa: pesquisa documental, pesquisa de campo, entrevistas em profundidade, questionários, vistorias, registro fotográfico etc.

Esses instrumentos são a porta de entrada de qualquer informação que será processada na avaliação. Cada uma das informações coletadas nos instrumentos, chamadas de variáveis, recebe um código, constituindo-se a base de dados que alimenta o sistema informatizado responsável pela geração da avaliação.

Os instrumentos utilizados são os seguintes:

- **Solicitação aos Bancos de Dados da Caixa (BD).** Este instrumento sistematiza a solicitação do conjunto de informações pertencentes aos sistemas da Caixa, que são necessárias para implementar a metodologia. Essas variáveis são utilizadas para a definição da amostra, em vários aspectos, como *economicidade, procedimentos de gestão e impacto macroeconômico*.
- **Questionário de Tramitação (QT).** Este questionário, aplicado aos solicitantes do financiamento, garante informações básicas para a avaliação dos aspectos *procedimentos de gestão*.
- **Instrumento de Solicitação ao Conselho Curador.** Este instrumento é bastante simples; objetiva registrar os dados relativos às metas estabelecidas pelo CCFGTS, *necessárias para o aspecto atendimento às necessidades de habitação e saneamento*.
- **Questionários ao Secretário-executivo e a Membros da Instância Colegiada (QI).** Tais questionários são resultado de entrevistas abertas

dirigidas aos membros das instâncias colegiadas e objetivam a avaliação deste indicador, que compõe o aspecto *procedimentos de gestão*.

- **Ficha de Leitura de Processo e Projeto (FP).** Esta ficha é preenchida no Escritório de Negócios da Caixa, tendo como fonte básica de informações o processo. Seus objetivos são, por um lado, sistematizar informações relativas ao contrato e ao empreendimento necessárias aos aspectos *procedimentos de gestão* e *economicidade* e, por outro, sobre o projeto, gerando variáveis necessárias para a avaliação de indicadores de *produto*.
- **Ficha de Observação da Unidade (FU).** Esta ficha, preenchida por vistoria em campo, permite levantar dados relativos à unidade habitacional, visando à avaliação de indicadores de *produto*, em todos os programas habitacionais.
- **Ficha de Observação das Áreas Coletivas (FC).** Esta ficha, preenchida por vistoria em campo, permite levantar dados relativos às áreas coletivas dos empreendimentos, visando à avaliação de indicadores de *produto* nos programas habitacionais coletivos.
- **Ficha de Observação das Áreas Urbanizadas (FA).** Esta ficha, preenchida por vistoria em campo, permite levantar dados relativos às áreas urbanizadas, visando à avaliação de indicadores de *produto* no programa Pró-moradia, modalidade Urbanização de Áreas.
- **Ficha de Observação do Entorno (FE).** Esta ficha, preenchida por vistoria em campo no entorno do empreendimento, gera as informações necessárias para a avaliação de *impacto na cidade e no ambiente*.
- **Questionário com Imobiliárias (QIm).** Este questionário objetiva levantar informações das imobiliárias sobre as transformações geradas pelo empreendimento no processo imobiliário, indicador do aspecto *impacto nas condições de vida*.
- **Questionário com o Beneficiário Final (QB).** O questionário com o morador é um instrumento de pesquisa de campo tradicional, sendo

o mais extenso de toda a metodologia. Seu principal objetivo é coletar os dados para a avaliação do *impacto nas condições de vida*, gerando informações sobre as mudanças ocorridas entre a situação anterior e posterior à implantação do empreendimento. O QB também é utilizado para avaliar alguns dos indicadores do aspecto *produto*, como satisfação do usuário (na sua totalidade), *habitabilidade urbana* e *habitabilidade da unidade habitacional* (em algumas de suas partes), e também alguns indicadores de *procedimentos de gestão* e *atendimento às necessidades habitacionais*.

CONCLUSÃO

Os programas públicos de habitação não têm sido avaliados de modo objetivo no Brasil. Embora envolvam, de uma maneira geral, investimentos significativos, até o momento não se desenvolveram metodologias sistemáticas e abrangentes capazes de avaliar esses programas, inexistindo uma cultura nos órgãos promotores e financiadores que dê prioridade à avaliação como um instrumento de aperfeiçoamento de suas políticas. Isto é particularmente grave quando se considera que a área da habitação é particularmente propícia à implantação de práticas de caráter clientelista e assistencialista, em que critérios objetivos de avaliação são desconsiderados.

Neste sentido, a criação de um sistema permanente de avaliação dos programas de habitação financiados pelo FGTS — realizada por iniciativa do Conselho Curador do Fundo e com a participação de Caixa Econômica Federal (agente operador e financeiro do FGTS) e dos vários atores representados no Conselho, como centrais sindicais e entidades empresariais da área da construção civil — é de grande importância, podendo representar um ponto de inflexão e colaborar na formação de uma cultura de avaliação do desempenho dos programas habitacionais no país.

NOTA

1 Projeto resultante da Resolução 259 de 24.4.1997, item 1º do Conselho Curador do FGTS, que gerou um contrato firmado entre a Escola de Engenharia de São Carlos da Universidade de São Paulo, por intermédio da FIPAI — Fundação de Incremento à Pesquisa e Aperfeiçoamento Industrial e a Caixa Econômica Federal. Este texto é uma síntese da metodologia formulada de modo

coletivo por uma equipe de pesquisadores do Departamento de Arquitetura e Urbanismo da EESC-USP e de outras instituições acadêmicas. Os coordenadores da equipe responsável pela metodologia de avaliação dos programas habitacionais são os professores Ângela Amaral, Miguel Buzzar, Ricardo Martucchi, Rosângela Paz, Rossella Rossetto, Sara Feldman e Sílvia Shor.

PUC-SP — Pontifícia Universidade Católica de São Paulo

Rua Monte Alegre, 984 — Perdizes

05014-001 — São Paulo — SP

Telefone: (11) 3670 8000

Site: www.pucsp.br

Reitor: *Antonio Carlos Caruso Ronca*

**IEE/PUC-SP — Instituto de Estudos Especiais da
Pontifícia Universidade Católica de São Paulo**

Rua Ministro Godoy, 1.213 — Perdizes

05015-001 — São Paulo — SP

Telefax: (11) 3871 4429

E-mail: iee@pucsp.br

Diretora: *Mariangela Belfiore Wanderley*

**Cenpec — Centro de Estudos e Pesquisas em
Educação, Cultura e Ação Comunitária**

Rua Dante Carraro, 68 — Pinheiros

05422-060 — São Paulo — SP

Telefax: (11) 3816 0666

Site: www.cenpec.org.br

E-mail: info@cenpec.org.br

Diretora-Presidente: *Maria Alice Setubal*

Coordenadora Geral: *Maria do Carmo Brant de Carvalho*

Os textos aqui apresentados são fruto do **Seminário de Avaliação de Políticas Públicas e Programas Sociais**, realizado em 2000, em São Paulo, promovido conjuntamente pelo **IEE/PUC-SP** — Instituto de Estudos Especiais da Pontifícia Universidade Católica de São Paulo, pelo **Genpec** — Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária e pelo **Cedac** — Centro de Estudos e Documentação para a Ação Comunitária.

Pesquisadores dedicados à investigação e à intervenção avaliativa na área social examinam, por diferentes pontos de vista, o referencial teórico-metodológico relacionado à avaliação de políticas públicas. O resultado constitui um mosaico de reflexões que certamente contribuirá para o aprofundamento do debate em torno da gestão de políticas públicas na área social.

Mariangela Belfiore Wanderley

