

AVALIAÇÃO DE PROGRAMAS SOCIAIS NO BRASIL: REPENSANDO PRÁTICAS E METODOLOGIAS DAS PESQUISAS AVALIATIVAS*

Paulo de Martino Jannuzzi**

A ampliação do escopo e da cobertura da política social no Brasil, processo este dinamizado a partir da Constituição Federal de 1988 (CF/88), vem ensejando esforços crescentes do campo de monitoramento e avaliação de programas sociais no país. Contudo, em que pesem tais esforços de aprimoramento, é preciso reconhecer que a proposição de sistemas de monitoramento para acompanhar as ações e os programas governamentais e a especificação de pesquisas de avaliação destes últimos ainda padecem de uma série de dificuldades.

Essas dificuldades no monitoramento e na avaliação de programas no Brasil – e as frustrações decorrentes disto – são consequência, em boa medida, de desenvolvê-los seguindo modelos e prescrições muito particulares e padronizados, sem a devida “customização” que deveriam ter em função do estágio de maturidade dos programas – ou do momento do “ciclo de vida” em que eles se encontram –, ou, ainda, de conduzi-los sem uma análise prévia do grau de avaliabilidade dos programas.

Este texto procura, assim, trazer alguma contribuição de natureza conceitual e metodológica para auxiliar na estruturação de sistemas de monitoramento e na especificação de pesquisas de avaliação, que realmente se prestem ao objetivo de permitir o acompanhamento contínuo e o aprimoramento efetivo da ação governamental e, portanto, a gestão de programas.

Como na pesquisa acadêmica, não existe uma receita única, pronta e acabada para responder a qualquer tipo de demanda avaliativa. Frente à complexidade do objeto de estudo, e sem prejuízo do rigor metodológico, é preciso que o estudo de avaliação seja conduzido com certa maleabilidade e pluralismo metodológico, triangulação de abordagens investigativas e complementaridade de técnicas quantitativas e qualitativas. Ainda que um tanto óbvio, não é bem assim que os estudos têm sido conduzidos no Brasil.

Palavras-chave: Monitoramento; Avaliação; Políticas Públicas; Programas Sociais.

* Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de produtividade concedida e aos meus ex-alunos do Curso de Especialização em Avaliação de Programas na Escola Nacional de Administração Pública (ENAP), Maria Rosângela Machado Silva, Mariana Almeida de Faria Sousa e Leonardo Milhomem Resende, e do Mestrado em Estudos Populacionais e Pesquisas Sociais da Escola Nacional de Ciências Estatísticas, do Instituto Brasileiro de Geografia e Estatística (ENCE/IBGE), Daniel Siqueira e Caroline Constant, pela oportunidade de diálogos e reflexões que me permitiram escrever o presente texto. Uma versão para fins didáticos deste artigo foi publicada como Cadernos EIAPP em 2009 pela ENAP.

** Professor da ENCE do IBGE e pesquisador do CNPq. Mestre em Administração Pública pela Escola de Administração de Empresas de São Paulo, da Fundação Getúlio Vargas de São Paulo (EAESP/FGV-SP) e doutor em Demografia pela Universidade Estadual de Campinas (UNICAMP). *E-mail*: paulo.jannuzzi@ibge.gov.br

EVALUATION OF SOCIAL PROGRAMS IN BRAZIL: RETHINKING PRACTICES AND RESEARCH METHODOLOGIES

The expansion of the scope and coverage of Social Policy in Brazil, a process spurred by the Constitution of 1988, have demanded increased efforts in the field of Monitoring and Evaluation of Social Programs in the country. In spite of these improvement efforts, it is clear that the development of monitoring systems to monitor the actions and government programs and the specification of evaluation research of these programs still suffer from a number of difficulties. These difficulties in monitoring and evaluating programs in Brazil – and the resulting frustrations – are due, in first place, of using very specific models and standardized rules without proper customization that should have done depending on the stage of program maturity - or the moment of “life cycle” in which they meet - or even to lead them without prior examination of the degree of evaluability of programs.

This paper thus seeks to bring some conceptual and methodological contribution to assist in structuring monitoring systems and in the specification of evaluation research, which really lend to the goal of allowing continuous monitoring and effective government improvement.

As it happens in academic research, there is no single recipe, prompt and ready to respond to any demand evaluative. Due to the complexity of the object of study, without losing the scientific rigor, the evaluation studies should be based on some sort of methodological pluralism, triangulation of research approaches and complementarity of quantitative and qualitative techniques. Although somewhat obvious, is not the case that the studies have been conducted in Brazil.

Key-words: Monitoring; Evaluation; Public Policy; Social Programs.

EVALUACIÓN DE PROGRAMAS SOCIALES EN BRASIL: CONSIDERACIONES SOBRE LA NECESIDAD DE REPENSAR PRÁCTICAS Y METODOLOGÍAS DE INVESTIGACIÓN

La ampliación del alcance y la cobertura de la política social en Brasil, un proceso impulsado por la Constitución de 1988, han exigido un mayor esfuerzo en el ámbito de Seguimiento y Evaluación de Programas Sociales en el país. A pesar de estos esfuerzos de mejora, es evidente que la proposición de sistemas de monitoreo para supervisar las acciones y programas de gobierno y la especificación de la investigación evaluativa de estos programas todavía adolecen de una serie de dificultades.

Estas dificultades en el seguimiento y evaluación de los programas en Brasil – y las frustraciones resultantes – son el resultado, en gran parte, del uso de modelos y prescripciones muy específicos y estandarizados, sin adaptación adecuado que debería tener en respecto a la etapa de madurez de los programas – o el momento del ciclo de vida en la que se encuentran – o incluso para conducirlos sin examen previo del grado de evaluabilidad de programas.

Este trabajo pretende, pues, traer algún beneficio de naturaleza conceptual y la metodología para ayudar en la estructuración de los sistemas de seguimiento y la especificación de la investigación evaluativa, que realmente se prestan con el objetivo de permitir el monitoreo y mejora continua de la acción gubernamental.

Como en la investigación académica, no hay una receta única, rápida y dispuesta a responder a cualquier demanda de evaluación. Frente a la complejidad del objeto de estudio, sin perder el rigor metodológico, e necesario que la evaluación de programas siga basada en cierta maleabilidad e pluralismo metodológico, con triangulación de los enfoques de investigación e complementariedad de las técnicas cuantitativas y cualitativas. Un tanto obvio, pero no es esto que se practica en la evaluación de programas en Brasil.

Palabras-clave: Seguimiento; Evaluación; las Políticas Públicas; Programas Sociales.

ÉVALUATION DES PROGRAMMES SOCIAUX AU BRÉSIL: CONSIDERATIONS SUR LA NECESSITÉ DE REPENSER LES PRATIQUES ET METHODES DE RECHERCHE

Le élargement de la couverture et de le domaine de la politique sociale au Brésil, processus qui a avancé depuis la Constitution de 1988, a exigé des efforts accrus dans le matière du suivi et évaluation des programmes sociaux dans le pays. En dépit de ces efforts d'amélioration, il est clair que la proposition de systèmes de suivi les actions et les programmes gouvernementaux et la définition approprié des études d'évaluation de ces programmes souffrent toujours d'un certain nombre de difficultés.

Ces difficultés dans le suivi et l'évaluation des programmes au Brésil – et les frustrations résultant – sont le résultat, en grande partie, pour le usage des modèles et prescipcions très spécifiques et standardisés, sans adaptation appropriée qui aurait dû, selon le stade du maturité des programmes –ou le moment du cycle de vie dans laquelle ils se trouvent – ou même de les conduire sans examen préalable du degré d'évaluation les programmes.

Le présent document tente d'apporter une contribution de nature conceptuelle et méthodologique pour aider à la structuration des systèmes de suivi et evaluation, qui se prêtent vraiment à l'objectif de permettre une surveillance continue et l'amélioration de l'action gouvernementale.

Comme dans la recherche scientifique, il n'y a pas de recette unique, rapide et prêt à répondre à toute demande d'évaluation. La complexité de l'objet de l'étude demande, sin le manque de rigueur méthodologique, que la evaluacion de programmes compte sur la flexibilité et pluralisme méthodologique, en la triangulation des méthodes de recherche e la complémentarité des techniques quntitatives e qualitatives. De toute évidence, pas tout à fait que des études ont été menées au Brésil.

Mots-clés: Suivi; Évaluation; Politiques Publiques; Programmes Sociaux.

1 INTRODUÇÃO

A ampliação do escopo e da cobertura da política social no Brasil, processo este dinamizado a partir da Constituição Federal de 1988 (CF/88), vem ensejando esforços crescentes do campo do monitoramento e da avaliação de programas sociais no país. Contudo, em que pesem tais esforços de aprimoramento, é forçoso reconhecer que a proposição de sistemas de monitoramento para acompanhar as ações e os programas governamentais e a especificação de pesquisas de avaliação destes últimos ainda padecem de uma série de dificuldades. Não é tão comum encontrar, nos gabinetes e nos escritórios de gestores e técnicos do setor público, especialmente em nível estadual e nos grandes municípios, sistemas de monitoramento que permitam acompanhar, por meio de um conjunto consistente e atualizado de indicadores, a ação governamental, da alocação do gasto público ao impacto junto aos públicos beneficiários. Muitos são os exemplos de pesquisas de avaliação que, mesmo motivadas por preocupação legítima e meritória com a eficácia e a efetividade do gasto público, revelam, infelizmente, resultados triviais, metodologicamente questionáveis ou com baixa apropriação para reformulação dos programas.

Essas dificuldades no monitoramento e na avaliação de programas no Brasil – e as frustrações decorrentes distos – são consequência, em boa medida, de desenvolvê-los seguindo modelos e prescrições muito particulares e padronizados, sem a devida “customização” que deveriam ter em função do estágio de maturidade dos programas – ou do momento do “ciclo de vida” em que eles se encontram –, ou, ainda, de conduzi-los sem uma análise prévia do grau de avaliabilidade dos programas e sem uma compreensão de que essas atividades se integram a processos mais amplos, e igualmente importantes, da gestão de programas.

Forçando um pouco o argumento para torná-lo mais claro – ainda que em prejuízo do reconhecimento do esforço meritório e do trabalho competente observado em diversos setores da administração pública brasileira –, o fato é que as lacunas de formação e o pouco domínio de conceitos e técnicas no campo de monitoramento e avaliação, na comunidade de gestores, acabam trazendo problemas na especificação destas atividades no governo. Isso leva às conhecidas frustrações com resultados tão ansiosamente esperados e aos questionamentos sobre a utilidade dos estudos.

Explicitando melhor a motivação para elaboração deste texto – sob o risco de simplificar demais a natureza e minimizar o alcance dos estudos avaliativos conduzidos por pesquisadores nas universidades e nos centros de pesquisa no país –, a abordagem marcadamente disciplinar da pesquisa aplicada no campo, a sobrevalorização de algumas técnicas e modelos específicos de avaliação e o desconhecimento do contexto de operação da ação pública e da forma com que os resultados dos estudos podem ser usados mais efetivamente pelos gestores também contribuem para conformação de tal quadro.

Este trabalho procura, assim, contribuir para melhoria na especificação da demanda de instrumentos de monitoramento e pesquisa de avaliação por parte da comunidade de gestores públicos e na estruturação dos serviços a serem oferecidos pelo universo de pesquisadores acadêmicos e profissionais. Mais precisamente, este texto procura trazer alguma contribuição de natureza conceitual e metodológica para auxiliar na estruturação de sistemas de monitoramento e na especificação de pesquisas de avaliação, que realmente se prestem ao objetivo de permitir o acompanhamento contínuo e o aprimoramento efetivo da ação governamental e, portanto, a gestão de programas.

Inicia-se com uma problematização atual acerca dos estudos avaliativos no Brasil, passando-se a uma breve exposição do ciclo de formulação e avaliação de políticas públicas e programas, para, ao fim, apresentar aspectos conceituais no campo e discutir como as pesquisas de avaliação – delineadas segundo as diferentes técnicas quantitativas ou qualitativas da pesquisa social – podem e devem ser definidas a partir do estágio de maturidade e das perguntas efetivamente relevantes para os gestores e os técnicos.

2 MONITORAMENTO E AVALIAÇÃO: UM CAMPO DE ESTUDOS EM FORMAÇÃO NO BRASIL

A crescente atuação do Estado brasileiro na promoção de políticas sociais – em especial, a partir da CF/88 – tem sido acompanhada, com alguma defasagem, do aprimoramento da capacidade de formulação e avaliação de programas públicos no país. De modo geral, os estudos de avaliação de demandas sociais, os diagnósticos propositivos para intervenções específicas, os sistemas de indicadores de monitoramento e as pesquisas de avaliação de programas vêm se tornando mais consistentes e presentes no cotidiano dos gestores públicos, ainda que em um ritmo menos veloz que o necessário e de forma bastante desigual pelas esferas de governo e regiões do país. Em que pese os avanços na capacidade de gestão de programas, ainda são expressivos os contrastes que se podem observar, por exemplo, no acompanhamento dos programas nas áreas da saúde e da qualificação profissional, ou entre equipes de gestores do governo federal e técnicos de prefeituras de pequenos municípios no interior do país.

É preciso reconhecer as limitações das pesquisas de avaliação de programas públicos como fontes de informação para readequação dos programas existentes ou proposição de novos. Nas avaliações de impacto, as constatações empíricas acerca da efetividade dos programas são, não raras vezes, ambíguas ou pouco consistentes. Nos casos em que se adotam delineamentos metodológicos *quasi-experimentais*, as condições para garantir a validade interna dos estudos acabam por inviabilizar a apropriação e a generalização dos resultados para avaliações mais abrangente de programas (WORTHERN *et al.*, 2004).

Vários estudos avaliativos focados no processo de implementação dos programas acabam tendo baixo poder de generalização devido às lacunas de conhecimento sobre as atividades desenvolvidas e o papel dos agentes institucionais envolvidos nos programas, seja pelas limitações ou pela inexistência de sistemas de informações de monitoramento, seja pelas decisões metodológicas quanto a técnicas, amostras e casos investigados, muitas vezes escolhidos mais em função dos prazos e dos recursos disponíveis do que pelas perguntas as quais se quer responder. Avaliações voltadas à análise da consistência teórica e do desenho lógico de programas – que poderiam gerar conhecimento mais generalizável na área – são menos frequentes ou, pelo menos, menos conhecidas do que outros tipos de estudos avaliativos no país.

Enfim, em que pese os avanços na área, é preciso reconhecer que a insatisfação com os estudos avaliativos ainda é mais frequente do que o desejável. Como já revelava Cotta há mais de dez anos:

As metodologias de avaliação de programas sociais têm sido objeto de severas críticas. Basicamente, afirma-se que, na prática, as avaliações não subsidiam o processo decisório

porque seus resultados são inconclusivos, inoportunos e irrelevantes. Inconclusivos em função das próprias limitações deste tipo de estudo, inoportunos devido à morosidade do processo avaliativo e irrelevantes porque não respondem às demandas informacionais de todos os agentes sociais afetos à intervenção (COTTA, 1998, p. 118).

Situação semelhante parece ter passado os Estados Unidos nos anos 1970, a julgar pelo relatório do General Accounting Office (GAO), que, mobilizado por demanda do Senado americano em 1974, constatou que as informações das pesquisas sociais e dos estudos avaliativos financiados com recursos públicos não contribuíam para o desenho das políticas públicas, pela dispersão e fragmentação de temas investigados, pela falta de coordenação e orientação do que era de interesse governamental investigar, pela dificuldade de acesso e pela compreensão dos relatórios de pesquisas (GAO, 1977).

Mesmo mais recentemente, têm-se constatado limitações e problemas nos estudos avaliativos de programas públicos nos Estados Unidos, às vezes com desdobramentos até piores que os descritos por Cotta. De fato, como colocam Worthern *et al.*:

De vez em quando, uma “avaliação” mal concebida ou mal executada produz informações que, no melhor dos casos, seriam enganosas e, no pior, absolutamente falsas. Embora essas ocorrências sejam raras, podem causar problemas graves. Como geralmente tem ar de respeitabilidade, essas avaliações não costumam ser questionadas, e o resultado é que decisões importantes sobre programas e serviços essenciais baseiam-se inadvertidamente em informações falaciosas (WORTHERN *et al.*, 2004, p. 44).

Conspiram para conformação de tais problemas, entre tantos fatores, a crença desmesurada na capacidade de antecipação e na implementação de programas por parte de técnicos de alto escalão que, ao não incorporarem a contribuição de agentes envolvidos no trabalho, acabam por desenhar processos e rotinas que desconsideram as distintas realidades de operação dos programas. Esse tecnocratismo ingênuo invariavelmente se esquece das dificuldades de articulação e colaboração dos três níveis de governo ou de pastas sociais em uma mesma esfera e das diferenças de capacidade de gestão e controle social país afora. Tão ou mais grave, também desconsidera a heterogeneidade socioeconômica da população demandante na formulação dos programas, considerando-a como único público-alvo, para o qual se deve disponibilizar um conjunto de serviços, das regiões Norte a Sul, do centro à periferia das cidades. Assim, com problemas nas fases iniciais do ciclo de gestão de programas – no reconhecimento das questões sociais e no desenho das intervenções idealizadas para mitigá-los –, não se poderiam esperar resultados, de fato, impactantes.

Desconhecimento sobre o estágio de avaliabilidade dos programas é outro fator a minar a credibilidade dos instrumentos de monitoramento e avaliação.

A prematuridade na encomenda de estudos avaliativos de resultados e impactos, quando se sabe que o programa ainda se encontra em fase de implantação ou com problemas de gestão, e a antecipação de avaliações externas, com natureza mais de auditoria em detrimento de avaliações de caráter mais formativo, conduzidas internamente e voltadas ao aprimoramento incremental do programa, são problemas mais frequentes do que se poderia esperar. São aportados volumes significativos de recursos em pesquisas de avaliação, procurando garantir representatividade amostral de resultados em nível nacional, esquecendo-se de estruturar painéis ou sistemas de indicadores de monitoramento, que – com todas as limitações de cobertura, mas com a tempestividade devida – permitiriam identificar boa parte dos problemas em tempo de serem corrigidos.

Para isso, contribui o estágio ainda incipiente do conhecimento do campo aplicado de monitoramento e avaliação de programas no Brasil, seja na estrutura pública, seja nas universidades e nos centros de pesquisa. Enquanto, nos Estados Unidos e em países europeus desenvolvidos, a cultura de avaliação de programas já se encontra em estágio maduro, depois de mais de três ou quatro décadas de desenvolvimento, com marcos conceituais abrangentes e testados, profissionais com formação multidisciplinar e instituições especializadas, no Brasil, a institucionalização do campo é bem mais recente.

Como comentam Worthern *et al.* (2004), a cultura de monitoramento e avaliação de programas nos Estados Unidos – assim como os instrumentos de planejamento e programação orçamentária – começou a se fortalecer nos anos 1960 no governo Lyndon Johnson, quando da estruturação de vários programas sociais em nível federal, no contexto do que se denominou *Great Society*, e da necessidade, portanto, de garantir melhor gestão dos programas públicos nesse país. Rossi *et al.* (2004) identificam esforços sistemáticos em avaliação de programas sociais norte-americanos, já a partir da década de 1930, como desdobramento do papel crescente do Estado americano no financiamento de serviços sociais – respondendo aos efeitos da crise de 1929. Esta avaliação, que se concentrava, inicialmente, nas áreas de educação – sobretudo na investigação de programas de alfabetização –, saúde pública – nas ações de combate à mortalidade por doenças infecto-parasitárias – e qualificação profissional, passa, no pós-guerra, a abranger áreas como prevenção da violência juvenil, planejamento familiar, nutrição, programas habitacionais e desenvolvimento rural.

No entanto, nos anos 1960, o campo da avaliação de programas passa por um *boom*, com a ampliação dos programas sociais, como já mencionado, e também, como lembram Rossi *et al.* (2004), com o avanço e a disseminação das técnicas aplicadas da pesquisa social, desenvolvidas nas universidades americanas.

Ao contrário do que se poderia imaginar, nas décadas seguintes, a ênfase na desregulamentação e descentralização da prestação de serviços públicos e no questionamento sobre a pertinência e o tamanho dos programas sociais nos governos republicanos conservadores, que assumiram o poder mais tarde, acabou por valorizar a prática da avaliação, ao disseminar a cultura de monitoramento e avaliação para governos estaduais, municipais e organizações não governamentais (ONGs). Afinal, até mesmo um “estado regulador” requer instrumentos de medição de resultados para acompanhar a prestação dos serviços financiados com recursos públicos, até para justificar, com alguma base técnica e transparência, a descontinuidade de programas sociais “ineficazes” e redirecionar os gastos para outras áreas – como para as despesas militares da Guerra do Vietnã, no governo Nixon, e para a defesa militar, no governo Reagan. Respondendo a essas demandas, comentam os autores que, nos anos 1970, são estruturados cursos de pós-graduação, de natureza disciplinar e, mais tarde, multidisciplinar, com forte ênfase em técnicas de pesquisa – quantitativas e qualitativas –, abrindo a possibilidade para a criação de uma comunidade profissional de avaliadores.¹

No caso brasileiro, atualmente, parte significativa da avaliação de programas públicos tem sido realizada por equipes de centros de pesquisa e universidades, com boa experiência em análise de macropolíticas, conjuntura social ou projetos de pesquisa acadêmicos, mas, em geral, sem experiência na implementação real de programas e sem reunir o conhecimento multidisciplinar requerido para abordagem dos problemas complexos em que estes procuram atuar.² Não há dúvidas de que, em geral, essas avaliações de programas contam com equipes competentes no emprego das boas práticas da pesquisa social e no uso rigoroso de técnicas quantitativas e qualitativas de coleta e análise de dados. Como nas pesquisas acadêmicas, é fundamental garantir o rigor metodológico nos estudos avaliativos, para permitir o conhecimento mais circunstanciado dos problemas e as soluções idealizadas para mitigá-los e conferir legitimidade técnica-científica na avaliação da ação governamental. Mas, como já alertado em um dos primeiros

1. Ver, por exemplo, as publicações, os eventos e os *sites* das associações profissionais ou acadêmicas de avaliação nesses países, entre os quais: American Evaluation Association – disponível em: <www.eval.org>; Canadian Evaluation Society – disponível em: <www.evaluationcanada.ca>; European Evaluation Society – disponível em: <www.europeanevaluation.org>; e Société Française d’Evaluation – disponível em: <www.afe.asso.fr>.

2. De fato, no Brasil, os trabalhos de avaliação de programas públicos e seus autores encontram-se dispersos pelas principais associações científicas disciplinares. A partir dos anos 2000, foram criados espaços mais multidisciplinares para discussão e apresentação de estudos de avaliação de políticas públicas; entre eles, a Associação Brasileira de Avaliação Educacional – disponível em: <www.abave.org.br>; a Rede Brasileira de Monitoramento e Avaliação – disponível em: <<http://redebrasileirademea.ning.com>>; o portal Políticas Públicas em Foco – disponível em: <www.boletim-fundap.cebrap.org.br> –, mantido pela Fundação do Desenvolvimento Administrativo (FUNDAP) e pelo Centro Brasileiro de Análise e Planejamento (CEBRAP); e os *sites* do Tribunal de Contas da União (TCU) – disponível em: <www.tcu.gov.br> – e da Controladoria-Geral da União (CGU) – disponível em: <www.cgu.gov.br>, além dos fóruns criados pelos programas de pós-graduação lato e stricto sensu – acadêmicos e profissionais – e das linhas de pesquisa nesse sentido em programas deste tipo em administração pública, economia social, saúde coletiva, direito etc.

manuais abrangentes de avaliação de programas no início dos anos 1970, programas sociais são “objetos de pesquisa” inerentemente difíceis de serem tratados, pelo ambiente muitas vezes inóspito e pouco colaborativo, pelo contexto político menos ou mais favorável, pela natureza fugidia das perguntas e das questões que são demandadas ao avaliador responder, pela menor ou maior disponibilidade de dados e pelo tempo e recursos disponíveis para conduzir análises mais consistentes (WEISS, 1972). Em manuais mais recentes, como os já citados de Worthern *et al.* (2004) e Rossi *et al.* (2004), recomenda-se que os estudos avaliativos sejam realizados por pesquisadores não apenas com conhecimento técnico e metodológico apropriado à temática em questão, mas também com experiência empírica anterior e, sobretudo, desprendimento para inovar e improvisar nas diversas circunstâncias e dificuldades que insistem em aparecer no cotidiano prático da área. Não é incomum, pois, que estudos avaliativos realizados por “puristas metodológicos” cheguem a resultados já conhecidos do gestor de programas, ou concluam – com ingênua assertividade – acerca da necessidade de descontinuidade do programa avaliado, pela suposta baixa efetividade apurada, segundo suas escolhas metodológicas sobre o que investigar e como fazê-lo.

3 O CICLO DE POLÍTICAS PÚBLICAS E PROGRAMAS

Nos manuais clássicos de Ciência Política, o processo de formulação de políticas públicas tem sido apresentado recorrentemente pelo ciclo de etapas sucessivas (*Policy Cycle*), com mais ou menos estágios, como ilustrado na figura 1. Em que pesem as críticas de longa data quanto à forma simplificada com que este apresenta o processo político e sua veracidade empírica, a separação em etapas se presta aos objetivos de evidenciar, ao longo do processo, ênfases diferenciadas no planejamento, na operação ou na avaliação dos programas. Justifica-se ainda para fins didáticos e para orientar o recorte analítico na pesquisa acadêmica na área (NASCIMENTO, 1991).

Nesse modelo, a primeira etapa (definição da agenda política – *Agenda-Setting*) corresponde aos múltiplos caminhos e processos que culminam com o reconhecimento de uma questão social como problema público e da necessidade da ação governamental para sua solução; isto é, a legitimação desta questão na pauta pública ou na agenda das políticas públicas do país, em determinado momento. A etapa seguinte (formulação de políticas e programas – *Policy Formulation*) refere-se aos processos e às atividades relacionados à construção de possíveis soluções, encaminhamentos e programas para lidar com a questão recém-legitimada na agenda. É preciso, então, em sequência – na tomada de decisão técnica-política (*Decision Making*) –, escolher o rumo a seguir, de ação efetiva ou não, decidindo-se por uma ou algumas das alternativas formuladas. A quarta etapa (implementação de políticas e programas – *Policy Implementation*) corresponde

aos esforços de efetivação da ação governamental na alocação de recursos e no desenvolvimento dos processos previstos nas alternativas e nos programas escolhidos anteriormente. Por fim, é preciso analisar se os esforços empreendidos estão atuando no sentido esperado de solucionar o problema original (avaliação das políticas e dos programas – *Policy Evaluation*). É necessário verificar se é preciso realizar mudanças nos programas implementados para garantir sua efetividade; descontinuá-los, se o problema deixou de compor a agenda; ou então adaptá-los a uma nova realidade, reiniciando o ciclo.

É oportuno registrar que a avaliação, enquanto *etapa* do ciclo, realiza-se após a implementação. Trata-se de um momento de natureza mais reflexiva para continuidade ou não do programa. Distingue-se, portanto, das *atividades* de monitoramento e avaliação, que se realizam mediante os sistemas de indicadores e as pesquisas de avaliação, instrumentos investigativos que podem ser empregados a qualquer momento do ciclo, como discutido mais adiante. Denominar essa etapa decisiva do ciclo como de avaliação somativa talvez ajudasse a evitar o duplo sentido que o termo assume na área.

A agenda política corresponde ao conjunto de assuntos e problemas que os gestores públicos e a comunidade política entendem como mais relevantes em dado momento e, não necessariamente, a lista de preocupações da sociedade ou os destaques da imprensa (JANN; WEGRICH, 2007). Afinal, o reconhecimento de uma questão social como problema de governo ou Estado não é um processo simples e imediato, que responde automaticamente às estatísticas disponíveis, por mais reveladoras que sejam da gravidade da questão, quando comparada a outros países ou momentos do passado. Não é a vontade de um técnico do setor público, um pesquisador acadêmico ou governante eleito, com conhecimento empírico consistente da realidade ou visão ousada, que garante imediatamente sua incorporação na agenda formal de governo. Como coloca Parada (2006, p. 73), “No toda idea entra a la agenda. No todos temas de la agenda se convierten em programas.”

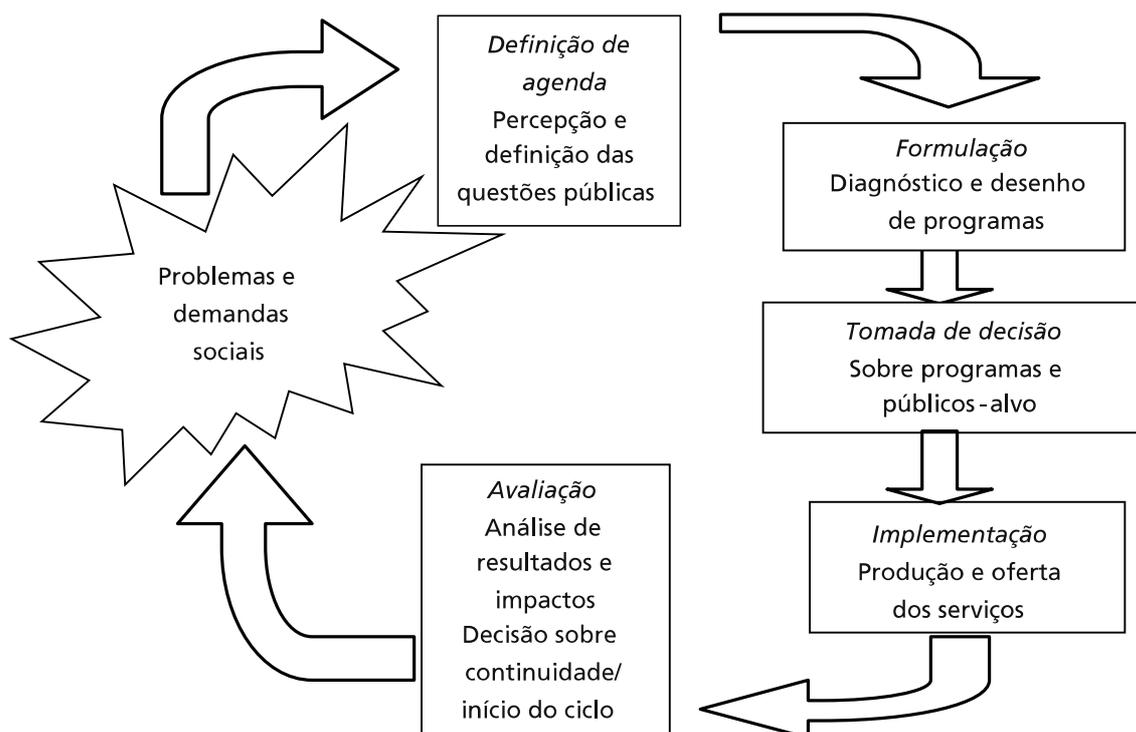
Quando existe a convicção de que um problema social precisa ser dominado política e administrativamente, é que ele se transforma em uma questão pública (FREY, 2002). Se parece haver solução técnica viável e factível para determinada questão social, essa entra mais facilmente na agenda. Afinal, a estrutura do setor público, pelos mecanismos institucionais existentes e operantes, é um ambiente que “digerer” inovações a seu tempo e estilo.

Há várias interpretações diferentes de como a agenda é construída. Em todas as perspectivas, ela é entendida como um processo coletivo e conflituoso de definição, como assinalam Howlett e Ramesh (2003). Vai sendo formada por meio de diferentes mecanismos de pressão externa de grupos organizados em partidos políticos, sindicatos, associações patronais, na imprensa e outras instituições, pelas

temáticas por eles entendidas como relevantes; por iniciativa interna do governo e mobilização social em função de compromissos assumidos nas eleições; pela atuação da burocracia pública no encaminhamento das demandas setoriais etc.

Os atores políticos são elementos catalisadores desse processo. Eles são diversos e possuem características distintas: políticos e burocratas; empresários, trabalhadores/sindicatos, servidores públicos e meios de comunicação (RUA, 1998). Grandes empresários, individualmente ou por meio de seus *lobbies* (grupos de pressão), encaminham suas demandas e pressionam os atores públicos, nas decisões governamentais, em favor de seus interesses. Outros agentes políticos de grande importância são os agentes internacionais, tais como o Fundo Monetário Internacional (FMI), o Banco Mundial, a Organização das Nações Unidas (ONU), a Organização Mundial do Comércio (OMC) etc; ou, ainda, os países com os quais o Brasil mantém relações políticas. Esses atores podem afetar não apenas a economia, como também a política interna do país (RUA, 1998). Além destes, os meios de comunicação são outros agentes importantes no processo – de modo especial, a televisão, com difusão massificada em todas as camadas da população –, pois dispõem de recursos para influenciar a opinião pública na formação das demandas (FREY, 2002). Enfim, as organizações políticas (partidos, sindicatos e grupos de interesse) são fundamentais para que as demandas entrem na agenda política do governo e, lá presentes, possam se transformar em ações e programas concretos.

FIGURA1
O ciclo de políticas e programas públicos



A amplitude da agenda política estaria condicionada também ao nível de desenvolvimento econômico e tecnológico do país, já que tais fatores, ao viabilizar ganhos crescentes de produtividade na economia, permitiriam o atendimento de demandas de grupos de interesse organizados, pela ampliação da oferta, cobertura e diversidade dos serviços sociais (CAREY, 2006). O ritmo de constituição de tal agenda de bem-estar social seria potencializado, em contextos de maior liberdade e participação política, com maior atuação de partidos políticos, sindicatos, imprensa e outros grupos de interesse. Os níveis de urbanização e envelhecimento populacional também seriam outros condicionantes importantes para entender a formação da agenda política, pela visibilidade e pela concretude que conferem aos problemas sociais deles decorrentes.

Na etapa da formulação da política, os problemas, as propostas e as demandas explicitados na agenda transformam-se em leis, programas e proposições de ações. Isto é, a formulação de política envolve a busca de possíveis soluções para as questões prioritizadas na agenda. Nesse momento, os elementos operacionais da política em questão precisam ser explicitados: diretrizes estratégicas, propostas de leis, decretos normativos, bem como programas e projetos. Em uma visão simplificada – e um tanto romântica e ingênua, como diria Charles Lindblom – do processo, tais programas e propostas seriam elaborados pelo emprego de técnicas gerais de planejamento de projetos; conhecimentos setoriais específicos; análise de viabilidade política, custo-benefício ou custo-efetividade das soluções; revisão crítica de experiências anteriores; e boas práticas identificadas. Essas tarefas são conduzidas por técnicos do setor público, com maior ou menor participação de consultores externos, seguindo orientações emanadas de gestores públicos no topo da hierarquia.

Em função da natureza do problema e da orientação geral de governo, os programas propostos se classificariam em ações de cunho distributivo, redistributivo ou regulamentatório (HOWLETT; RAMESH, 2003). Na realidade, essa fase não está tão desvinculada assim do cotidiano da política ou tão isolada das influências das propostas dos grupos de interesse, como se supõe. Muitas vezes, as propostas de encaminhamento de soluções chegam com as pressões para introduzir a questão na agenda formal de governo. Se não, colocado o problema, há sempre atores com maior ou menor influência na formatação dos possíveis programas e ações (JANN; WEGRICH, 2007).

Levantadas as diferentes alternativas para dada questão social, é preciso escolher qual deverá ser adotada. Enquanto na definição da agenda política e na formulação de programas a participação de agentes não pertencentes ao governo ou ao Estado desempenha papel central, a tomada de decisão envolve os gestores e os técnicos mais diretamente relacionados à política ou ao programa. Esses são o ônus e o bônus do exercício da atividade pública. Isso não significa que outros

agentes e atores não estejam ativos no processo, de modo a influenciar a decisão em um ou outro sentido, introduzindo – legitimamente em muitos casos – considerações de ordem política em rotinas de cunho mais eminentemente técnico.

Chega então o momento da implementação dos programas, dos projetos e das ações pontuais ou mais abrangentes. Esta fase é o momento em que a política pública começa a ser executada, em que os recursos alocados começam a se transformar em ações capazes de agir e mudar certa realidade. Neste ponto, a política adquire uma concretude objetiva maior e, por isso, começa a revelar mais explicitamente as dificuldades de sua introdução e sua operação pelos agentes encarregados de “por a mão na massa” (NASCIMENTO, 1991). Trata-se da gestão das ações para fazer uma política sair do papel e funcionar efetivamente, compreendendo o conjunto de realizações, por grupos ou atores de natureza pública ou privada, para a consecução de objetivos definidos. Envolve programação de atividades, provisionamento de recursos financeiros, alocação de recursos humanos, mobilização de agentes, interlocução com atores estratégicos, manejo dos mecanismos que assegurem a governabilidade das atividades e correção de cursos, caso se identifiquem obstáculos ou surpresas não antecipadas.

É na avaliação – ou melhor, na avaliação somativa – que são confrontados os resultados esperados com os resultados alcançados. É o momento em que se dá um retorno à etapa inicial de formulação, com correções no plano de ação, caso seja uma política de ordem contínua; ou quando se decide pelo encerramento da política. É também uma etapa de verificação dos impactos desejados e indesejados. Trata-se de verificar os efeitos obtidos para deduzir consequências indesejadas nas ações e nos programas futuros. Como bem coloca Frey (2002) avaliação é imprescindível para o desenvolvimento e a adaptação contínua das formas e dos instrumentos de ação pública, o que pode ser denominado como a fase de “aprendizagem política”.

O ciclo de políticas públicas tal como descrito é, certamente, um modelo idealizado, com aderência bastante limitada ao cotidiano verificado no Brasil ou em outros países. Como bem observa Saravia (2006, p. 29), “O processo de política pública não possui uma racionalidade manifesta. Não é uma ordenação tranquila na qual cada ator social conhece e desempenha o papel esperado.” Pelo que sugere Lindblom (2006), atualização da agenda, formulação de políticas e programas, bem como implementação e avaliação seriam estágios permanentes e concomitantes do processo político. A atualização da agenda política e as inovações dos programas seriam, de fato, muito mais incrementais e contínuas do que gostariam os gestores públicos. Tampouco haveria, segundo Parada (2006, p. 72), um momento de avaliação cabal das políticas e dos programas, sendo “(...) más habitual que cambien o se combinem con outras. Se há llegado a decir que las políticas publicas son imortables.”

Boa parte da crítica feita a esse modelo se refere ao fato de ele sugerir que a administração pública, seus gestores, os atores políticos e os técnicos atuem de forma bastante sistemática e cooperativa, como se estivessem todos envolvidos na resolução de um problema consensualmente percebido, empregando métodos racionais e objetivos na busca da solução, de acordo com uma sequência linear de etapas bem-delineadas. Como diria Lindblom (2006), mais uma vez, o imperativo da racionalidade técnica no processo, a análise exaustiva dos problemas, a busca de soluções ótimas e a crença no poder revelador e conciliador do discurso técnico-científico – tão presente nos órgãos de planejamento público – não seriam empiricamente constatáveis, nem factíveis e muito menos desejáveis.

Na realidade, a formulação de políticas configura-se como um processo que envolve a interação de muitos agentes, com diferentes interesses. É marcada por apoios entusiasmados de alguns, resistências legítimas ou não de outros. Está repleta de avanços e retrocessos, com desdobramentos não necessariamente sequenciais e não plenamente antecipáveis.

Contudo, na visão de Howlett e Ramesh (2003), esse modelo de representação tem a grande virtude de facilitar o entendimento do processo complexo de interação de diversos agentes, nos múltiplos estágios por que passa a formulação de políticas públicas, oferecendo um marco metodológico geral para análises isoladas de cada etapa do processo ou das relações de cada uma com as demais, à frente ou à jusante. Tal modelo seria também suficientemente geral para ser aplicado no entendimento do processo na maioria dos âmbitos e dos contextos de formulação de políticas e programas – em nível federal, estadual, local e, até mesmo, setorial.

Jann e Wegrich (2007) acrescentam ainda que o sucesso e a “resiliência” desse modelo clássico às críticas de pesquisadores acadêmicos é que ele parece como um do tipo prescritivo ideal a ser adotado, em que gestores eleitos governam com corpos burocráticos com boa capacidade de diagnóstico de problemas, com posturas ativas e inovadoras na formulação de programas, operando com racionalidade técnica na tomada de decisões, controle efetivo das atividades na implementação dos programas públicos e objetividade na avaliação destes.

4 AS PESQUISAS DE AVALIAÇÃO NO CICLO DE POLÍTICAS E PROGRAMAS

O fato de se denominar de avaliação a fase posterior à implementação de programas no ciclo de políticas e programas cria uma distinção artificial entre as atividades de monitoramento, exercidas com base nos sistemas de indicadores descritos no tópico anterior, e as de realização de pesquisas de avaliação, que podem ser realizadas a qualquer momento do ciclo.

Na realidade, monitoramento e avaliação de programas são termos cunhados para designar procedimentos técnicos formais de acompanhamento de programas,

focados na análise da eficiência, eficácia e efetividade, visando ao aprimoramento da ação pública (COHEN; FRANCO, 1999). Monitoramento e avaliação são processos analíticos organicamente articulados, que se complementam no tempo, com o propósito de subsidiar o gestor público de informações mais sintéticas e tempestivas sobre a operação do programa – resumidas em painéis ou sistemas de indicadores de monitoramento – e mais analíticas sobre o funcionamento deste, levantadas nas pesquisas de avaliação.

As tipologias clássicas usadas para classificar essas pesquisas acabam também favorecendo alguma confusão, como considerar monitoramento como avaliação de processo. Esta última é aquela realizada quando se requer avaliar as atividades de implementação dos programas, a extensão de cobertura do público-alvo atendido, a qualidade dos serviços oferecidos, ou ainda entender as dificuldades que estão “emperrando” a efetiva operação do programa, em contraposição à pesquisa de avaliação de resultados e impactos, realizada idealmente depois de resolvidos os problemas de gestão, quando se requer, já na fase de avaliação do ciclo, uma resposta cabal às perguntas: O problema social que originou o programa foi equacionado? Qual a contribuição desse programa para isso? Este deve continuar, ser expandido ou descontinuado? (CARVALHO, 2003).

Uma tipologia de pesquisas avaliativas particularmente interessante, face sua aderência ao ciclo de políticas e programas já exposto, é proposta por Rossi *et al.* (2004), como ilustrado na figura 2. Para esses autores, os estudos avaliativos deveriam compor um programa de investigação completo, seguindo a lógica implícita do “ciclo de vida” dos programas. Antes de mais nada, seria preciso analisar a pertinência do programa na agenda de políticas públicas (*Needs Assessment*), investigando se responde a uma demanda social efetivamente existente.

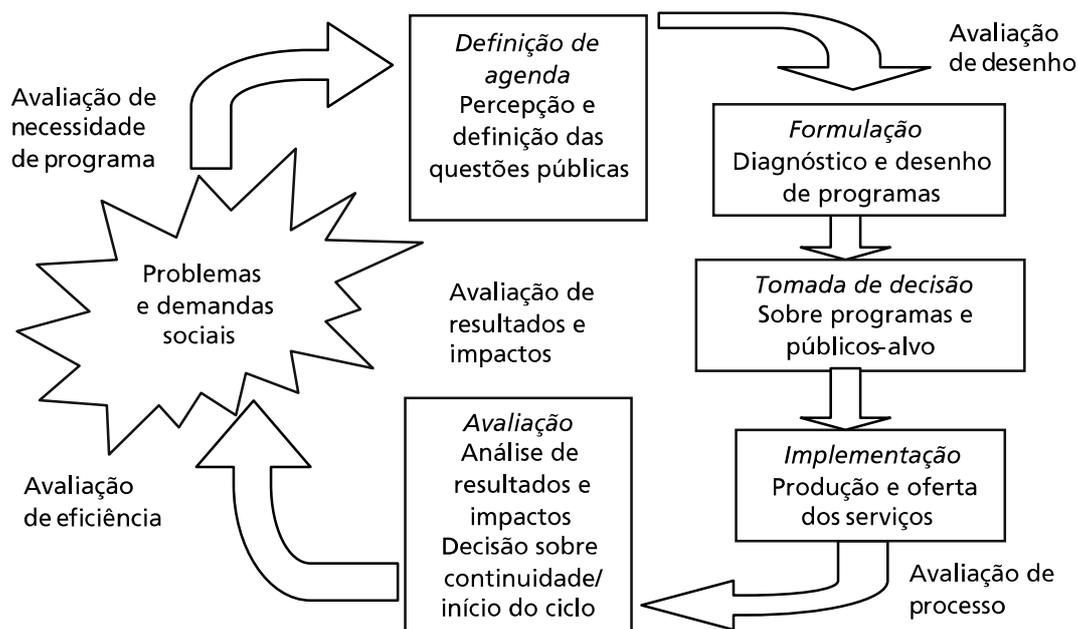
Se o programa existe para responder a uma demanda pública de fato, e não a uma necessidade de autopreservação de uma estrutura organizacional criada em algum momento do passado, deveria se investigar a teoria ou o modelo de intervenção em que o programa se assenta (*Design Assessment*). Afinal, este é o melhor desenho lógico de intervenção, no quadro de relações institucionais prevalecente entre as esferas de governo? E frente às distintas capacidades de gestão e controle social pelo país afora? Respondidas essas questões, faria sentido, então, investigar o processo de implementação do programa (*Program Process Assessment*), e se constatada a inexistência de graves problemas quanto à entrega dos serviços ao público-alvo, faria sentido avaliar seus impactos e seus resultados (*Impact Assessment*). Ao fim, tendo-se constatado que o programa é socialmente justificável e conta com um desenho de implementação adequado, procedimentos operacionais bem-estabelecidos e resultados efetivamente comprovados, passar-se-ia para a avaliação da eficiência ou do custo-efetividade

(*Efficiency Assessment*), para responder se o custo da operação do programa legitima-se pelos efeitos obtidos e se é replicável para outras escalas de operação.

A proposta de Rossi *et al.* (2004) é certamente um avanço conceitual na forma de se entender os diferentes focos dos estudos avaliativos, para além daquelas dicotomias “avaliação de processos/avaliação de impactos”, “avaliação *ex-ante*/avaliação *ex-post*” etc. Ao propor ênfases avaliativas para cada etapa do “ciclo de vida” do programa, os autores oferecem um plano de avaliação abrangente, orientando a especificação das pesquisas neste sentido segundo uma lógica de abordagem que tenderia a garantir, *a priori*, resultados relevantes e úteis ao fim do processo. Tal proposta ajudaria a definir melhor a questão crucial para sucesso de uma pesquisa de avaliação de programas: que aspectos do programa devem ser avaliados e quais devem ser abordados posteriormente, em função do estágio de maturidade do programa, para maximizar as chances de se obter informações úteis e relevantes para seu aprimoramento no momento.

FIGURA 2

Tipologia de avaliação de Rossi *et al.* e o ciclo de políticas e programas



Fonte: Jann e Wegrich (2007).
Elaboração do autor.

Afinal, não caberia especificar uma pesquisa de avaliação de resultados e impactos se há dúvidas com relação ao estágio de maturidade dos processos e das ações previstas na implementação do programa, ou ainda se existem suspeitas de que o desenho desta padece de supostos não sustentáveis em sua estrutura lógica de intervenção. Não caberia proceder-se a investigações complexas de análise de custo-efetividade se não há certeza de que os resultados e os impactos gerados pelo programa são significativos. Enfim, evidências de baixa realização do programa

poderiam ser consequências menos de falhas no desenho ou da forma como ele foi implementado, e sim da própria inadequação de existência deste, como parece ser o caso dos programas voltados a oferecer experiência de primeiro emprego para jovens (MADEIRA, 2004).

A hierarquia de focos avaliativos de Rossi *et al.* (2004) não implica que o gestor só possa dispor de informações acerca da eficiência, da qualidade dos serviços e da cobertura de atendimento do público-alvo nas fases mais adiantadas do “ciclo de vida” do programa. Pressupõe-se que há um sistema de indicadores de monitoramento já estruturado que reúne informações indicativas nesse sentido. Pode até ser necessário encomendar pesquisas de avaliação de resultados e impactos ou estudos mais aprofundados de análise da eficiência em estágios ainda incipientes da implementação do programa, para responder a demandas externas de avaliação. O que os autores defendem é que se organize um plano consistente de análise em todas as etapas do “ciclo de vida” do programa e que se resista à tentação de “colocar a carroça antes dos bois”, encomendando-se pesquisas de avaliação de forma precoce e desarticulada. É preciso ser diligente com o risco de produção de resultados irrelevantes ou, pior, com a produção de resultados precipitados em pretensa legitimidade científica.

5 OS DIFERENTES DELINEAMENTOS METODOLÓGICOS DAS PESQUISAS DE AVALIAÇÃO

Cada um desses tipos de estudos avaliativos requer uma abordagem metodológica e/ou disciplinar diferente. Estudos de viabilidade ou necessidade de programas podem se valer de análises documentais comparativas, de recurso às técnicas de avaliação econômica ou, até mesmo, às análises institucionais típicos da Ciência Política. Estudo de desenhos lógicos de implementação envolve revisão de documentos usados quando da formulação do programa e emprego das técnicas de planejamento de projetos e de comparação com outros modelos de intervenção social semelhantes em distintos contextos e momentos.

A avaliação de processos pode requerer emprego de técnicas variadas de coleta de informações usadas na pesquisa social. A avaliação de resultados e impactos, por sua vez, pode requisitar a adoção de entrevistas com agentes institucionais, grupos de discussão com beneficiários dos programas, pesquisas amostrais, delineamentos *quasi-experimentais* ou estudos comparativos de casos.

A especificação do delineamento metodológico adequado aos objetivos da pesquisa de avaliação pretendida é, pois, aspecto crucial. A avaliação de programas é um empreendimento técnico-científico de uso de métodos da pesquisa social para investigar a situação, os problemas e os diferentes aspectos da gestão de um programa público, ao longo do seu “ciclo de vida”, da sua concepção ao

usufruto dos produtos e dos serviços por ele disponibilizado, considerando o contexto organizacional e político em que se insere, com a finalidade última de informar as necessidades de aprimoramento de suas ações, de modo a contribuir, com outros programas, na melhoria das condições sociais da população. Como empreendimento técnico-científico, tal como na condução de pesquisas acadêmicas, a escolha do método depende dos objetivos específicos almejados pelo avaliador. A busca de dados secundários, a observação e suas variantes – visitação, observação participante etc. –, as entrevistas – semiestruturadas ou estruturadas –, os grupos de discussão, as pesquisas de campo com questionários – com amostras probabilísticas ou intencionais –, e os desenhos *quasi-experimentais* e não experimentais são algumas das estratégias metodológicas mais empregadas nas pesquisas de avaliação de programas no Brasil.

Cada técnica tem suas características vantagens e limitações. Os métodos quantitativos, organizados sob a égide do modelo hipotético-dedutivo, como os levantamentos amostrais e os experimentos, compreendem técnicas bastante estruturadas – destinadas à investigação de problemas específicos –, voltadas ao dimensionamento de quantidades ou da intensidade de relações entre variáveis, supondo distanciamento do pesquisador em relação ao objeto investigado. Os métodos qualitativos, baseados no princípio da produção do conhecimento pela lógica indutiva, do particular para o geral, como a observação, as entrevistas e os grupos de discussão, compreendem técnicas pouco ou semiestruturadas, para investigar exploratoriamente problemas complexos, pressupondo proximidade do pesquisador ao objeto avaliado.

Foge ao escopo deste trabalho detalhar prescrições metodológicas acerca do emprego de uma ou outra técnica, dada a complexidade de cada uma e a boa disponibilidade de manuais brasileiros de métodos e processos de pesquisa social, como os de Laville e Dionne (1999), Cano (2002), Richardson *et al.* (2002) e Babbie (1999), entre outros. Vale incluir nessa lista, como importante material de referência para delineamento metodológico de pesquisas de avaliação, o documento organizado pela Secretaria de Avaliação e Gestão da Informação do Ministério de Desenvolvimento Social e Combate à Fome (Sagi/MDS)³ (PAES-SOUSA; VAITSMAN, 2007). São apresentados 18 estudos, na sua maioria pesquisas de avaliação de programas do ministério nas áreas de assistência social, segurança alimentar e nutricional e renda da cidadania. Na exposição do estudo avaliativo de cada programa, são apresentados as instituições e os pesquisadores participantes, o período de realização, os objetivos da avaliação, os aspectos metodológicos acerca das técnicas de coleta de dados empregadas, além das características da amostra e dos sujeitos entrevistados. As bases de dados de algumas dessas pesquisas de ava-

3. Disponível em: <www.mds.gov.br>, no link Institucional – Sagi.

liação foram disponibilizadas no Consórcio de Informações Sociais da Associação Nacional de Pós-graduação e Pesquisa em Ciências Sociais (ANPOCS), permitindo a pesquisadores e avaliadores buscar respostas preliminares para questões de outros programas similares.⁴

Esse material evidencia, na prática, como a natureza do programa, o estágio em que ele se encontra, os objetivos de avaliação e a disponibilidade de tempo e recursos acabam orientando as decisões metodológicas acerca de instrumentos e técnicas de pesquisa, o tamanho e a intencionalidade ou não das amostras das pesquisas de avaliação encomendadas pela Sagi. Ilustra-se como pesquisas de avaliação de processo conduzidas para vários programas procuram garantir levantamento de informações por meio de entrevistas semiestruturadas com agentes públicos em diferentes posições da produção dos serviços (gestores municipais e técnicos atendentes do público), em localidades intencionalmente selecionadas – o segundo tempo de implantação do programa, por exemplo. Percebe-se, contudo, predominância de uso de questionários bastante estruturados, até mesmo em estudos avaliativos de natureza mais exploratória. Constata-se também que há pouco emprego de grupos de discussão e não se utilizam ferramentas de análise estruturada de textos, inclusive nos estudos avaliativos que compreenderam mais de uma centena de entrevistas. Worthen *et al.* (2004) reputam essas técnicas como avanços metodológicos significativos nas abordagens qualitativas de avaliação.

Vale observar que as pesquisas de avaliação não implicam necessariamente o levantamento de dados por meio de algumas das técnicas anteriormente relacionadas. Os dados disponíveis nos registros operacionais do programa, o sistema de indicadores de monitoramento, as pesquisas do IBGE e as informações de outros órgãos podem permitir fazer análises preliminares a baixo custo e tempo. O problema é que, em geral, não têm a especificidade necessária para responder questões mais particulares do gestor.

Outras fontes de informação secundária úteis são os relatórios de avaliação de programas do TCU e da CGU. Os relatórios de avaliação de programas deste tribunal,⁵ elaborados desde 1998, são organizados em grandes tópicos, iniciando-se com a exposição dos objetivos da avaliação do programa (seção *O que foi avaliado*), da justificativa para sua realização (*Por que foi avaliado*), dos aspectos metodológicos da pesquisa avaliativa, e explicitando-se os métodos e as técnicas de coleta e análise dos dados, as amostras e os sujeitos investigados (*Como se desenvolveu o trabalho*) e uma breve apresentação do programa (*Histórico do programa*). Os resultados da avaliação (*O que o TCU encontrou*) são apresentados em seções, organizadas segundo os objetivos da avaliação. Ao fim, sistematizam-se as sugestões de aprimoramento para o programa (*O que pode ser feito para melhorar o programa*).

4. Disponível em: <www.nadd.prp.usp.br/cis>.

5. Disponível em: <www.tcu.gov.br>, no link Avaliação de Programas de Governo.

Os relatórios de avaliação da execução dos programas federais, conduzidos pela CGU,⁶ na forma de sorteios sistemáticos de estados e municípios desde 2003, são também fontes de informação muito interessantes para avaliações preliminares de processo de implementação de programas públicos. Um exemplo nesse sentido é o de Vieira (2008), que empregou os relatórios deste órgão para identificar os problemas de gestão municipal do Programa de Assistência Farmacêutica, em uma amostra de 597 municípios. A autora constatou que 90,3% deles apresentaram problemas na gestão de recursos ou serviços, como falta ou deficiência de controle de estoque de medicamentos, compras irregulares e condições inadequadas de armazenamento, entre outros aspectos.

Não existe *a priori* um método universal, mais legítimo ou com maior “*status científico*” para toda e qualquer pesquisa de avaliação, como não existe único procedimento para as pesquisas acadêmicas.

É estranho, pois, que ainda persista, em certas comunidades, o mito de que os delineamentos experimentais ou *quasi-experimentais* constituem-se nos métodos mais adequados e legítimos cientificamente para avaliação de impactos.

Esses desenhos metodológicos de avaliação são inspirados no formato clássico do experimento nas Ciências Naturais. Tal modelo tem o objetivo de investigar a estrutura e a intensidade de causalidade entre uma variável-efeito dependente e seus fatores determinantes. Para isso, é preciso garantir o controle da situação experimental em laboratório e o emprego de grupos de tratamento (beneficiários de programas) e controle (não beneficiários) escolhidos de forma aleatória. Nos desenhos *quasi-experimentais*, uma das condições básicas que definem o experimento não é verificada; em geral, a designação aleatória em grupos experimental e de controle. Tais desenhos visam avaliar a evolução dos dois grupos – supostamente idênticos no começo do experimento –, ao longo de tempo, e constatar se ao fim há evidências de que o primeiro mostrou melhor *performance* que o segundo, em uma variável entendida como reveladora dos efeitos do programa.

Na realidade, tal delineamento de pesquisa – experimental ou *quasi-experimental* – é um dos métodos usados para avaliação de impacto, não necessariamente o melhor ou mais factível. Nesse tipo de desenho avaliativo, há problemas éticos (Como escolher e justificar quem vai ser beneficiário e quem vai ficar de fora do programa?); operacionais (Como evitar a evasão dos beneficiários? Como garantir que os efeitos medidos são apenas do programa em foco, em meio a um contexto crescente de intervenções sociais?); metodológicos (O indicador empregado é a melhor medida para captar a dimensão impactada? O que se supõe como impacto guarda, pelo desenho lógico do programa, vinculação

6. Disponível em: <www.cgu.gov.br>, no link Auditoria e Fiscalização.

estreita com as ações e as atividades deste? O impacto deve ser medido sobre os beneficiários ou junto à comunidade a que pertencem?); e epistemológicos (Como os esforços de garantia da validade interna do experimento conspiram contra a generalização dos resultados? Como garantir que um programa bem-avaliado em circunstâncias tão artificiais possa repetir o êxito em situações normais?) (WEISS, 1972; MOHR, 1995; CANO, 2002).

Ainda que todos esses problemas fossem contornáveis, restaria um de natureza prática: se os efeitos potenciais do programa, tal como medidos em determinada variável, não forem elevados – algo que o incrementalismo de Lindblom sugeriria –, as amostras de beneficiários atendidos e do grupo de controle teriam que ser consideravelmente grandes para que os testes estatísticos possam ser aceitos sem hesitação (ROSSI *et al.*, 2004).

A mitificação desse desenho na avaliação de programas se deve, em alguma medida, pela origem dos estudos avaliativos centrados na investigação de programas nas áreas de educação e saúde pública, como já mencionado, em que tais modelos podem se viabilizar mais concretamente – pelas condições de simulação de “laboratório” em salas de aula ou pela tradição dos ensaios de tratamento clínico de doenças. A hegemonia circunstancial dos modelos quantitativos importados da pesquisa em Ciências Naturais, no debate sobre a cientificidade dos métodos de pesquisa a serem ensinados e adotados na pesquisa social americana nos anos 1960 – momento de expansão dos estudos avaliativos nesse país, como já mencionado – também é outro fator explicativo. Nesse contexto, como sugerem Worthern *et al.* (2004), o livro *Experimental and Quasi-Experimental Designs for Research*, escrito pelos psicólogos Donald Campbell e Julian Stanley, publicado em 1966, acabou sendo adotado como o manual de referência metodológica da comunidade de avaliadores que ora se formava. Afinal, apesar das advertências sobre as dificuldades de replicação das condições de controle laboratorial no contexto de operação dos programas sociais, “A elegância e a precisão do método experimental levaram a maioria dos avaliadores de programa a vê-lo como ideal.” (WORTHERN *et al.*, 2004, p. 116).

As críticas que se seguiram nas décadas seguintes sobre aspectos éticos, factibilidade operacional e poder de generalização dos resultados de desenhos experimentais – e suas aproximações *quasi-experimentais* –, incorporação de avaliadores provenientes das várias disciplinas das Ciências Sociais – antropólogos, sociólogos, comunicólogos etc. – e formalização mais rigorosa de abordagens de investigação mais qualitativas e adequadas aos problemas complexos e pouco estruturados da realidade social acabaram por consolidar a percepção, na comunidade de avaliadores nos Estados Unidos, de que os estudos avaliativos requerem certo ecletismo metodológico, integrando métodos quantitativos e qualitativos.

Dada a complexidade operacional, os contextos de implementação, os desenhos institucionais e a diversidade de públicos-alvo dos programas públicos no Brasil, não se pode definir uma técnica de investigação como melhor em qualquer situação. Muitos programas operam em contextos complexos, pouco estruturados para abordagens quantitativas – e muito menos para desenhos *quasi-experimentais*. Nessas situações, abordagens metodológicas menos estruturadas podem levantar evidências mais relevantes e úteis para aprimoramento dos programas. É possível que permitam a estruturação de questões mais específicas a investigar, para aplicação posterior de desenhos quantitativos de pesquisa. Abordagens quantitativas e qualitativas não são mutuamente excludentes em um projeto de pesquisa ou avaliação. São complementares, compatíveis e conectáveis.

Pluralismo metodológico, enfoques avaliativos mistos, triangulação de abordagens investigativas⁷ e complementariedade de técnicas. São essas as perspectivas da pesquisa de avaliação de programas partilhadas pelos autores das duas principais referências bibliográficas citadas neste estudo (WORTHERN *et al.*, 2004; ROSSI *et al.*, 2004). Rigor metodológico, capacidade de improvisação e maleabilidade técnica frente à complexidade do objeto de estudo, estas são as prescrições generalizáveis para qualquer equipe – necessariamente multidisciplinar – que queira encarar responsabilmente a pesquisa aplicada na avaliação de programas.

6 CONSIDERAÇÕES FINAIS

A ampliação do gasto social no Brasil e a diversificação dos programas voltados a atender às diversas demandas públicas vêm pressionando o setor público a aprimorar suas práticas de gestão. Nesse sentido, a preocupação com o aprimoramento técnico na elaboração de diagnósticos e nas atividades de monitoramento e avaliação de programas vem crescendo.

Estruturar melhor os sistemas de indicadores de monitoramento e especificar pesquisas de avaliação mais consistentes são desafios que precisam ser rapidamente enfrentados nos três níveis de governo (federal, estadual e municipal), sob pena de estender, por mais tempo ainda, a superação das iniquidades sociais no país e com o risco de se perder a crença de que os programas públicos podem ser vetores impactantes da mudança social no Brasil.

Este texto procurou mostrar que tais instrumentos precisam ser especificados, respondendo às demandas de informação do gestor nas diferentes fases do “ciclo de vida” do programa, de acordo com o estágio de maturidade deste. Não existe receita única, pronta e acabada para desenhar esses instrumentos. Existem experiências, recomendações e boas práticas.

7. Triangular significa abordar o objeto de pesquisa com três – ou mais – técnicas diferentes de investigação, como linhas não paralelas na forma de um triângulo cercando o objeto de pesquisa ao centro.

Não se trata de avaliar projetos sociais de ONGs, para uso instrumental de um financiador externo decidir se mantém ou não o aporte de recursos, mas se trata de avaliar programas públicos que tem sua legitimidade e sua perenidade garantida por outros critérios que não apenas os identificados em uma pesquisa de avaliação de impactos. Frente à complexidade de problemas sociais e de públicos-alvo por atender no país, trata-se de buscar estratégias de pesquisa que permitam, de fato, levantar informações que sejam úteis – e tempestivas – para aprimoramento da ação governamental.

REFERÊNCIAS

- ALMEIDA, A. C. **Como são feitas as pesquisas eleitorais e de opinião**. Rio de Janeiro: FGV, 2002.
- BABBIE, E. **Metodos de pesquisa de Survey**. Belo Horizonte: Editora UFMG, 1999.
- BENNETT, C. A.; LUMSDAINE, A. A. **Evaluation and experiment: some critical issues in assessing social programs**. New York: Academic Press, 1975.
- CANO, I. **Avaliação de programas sociais**. Rio de Janeiro: FGV, 2002.
- CAREY, G. W. GOVERNMENT. *In*: ENCYCLOPEDIA AMERICANA. Danbury: Scholastic, 2006. p. 127-131.
- CARVALHO, S. N. Avaliação de programas sociais: balanço das experiências. **São Paulo em Perspectiva**, São Paulo, v. 17, n. 3-4, p. 185-197, 2003.
- COHEN, E.; FRANCO, R. **Avaliação de projetos sociais**. Petrópolis: Vozes, 1999.
- COTTA, T. C. Metodologias de avaliação de programas e projetos sociais: análise de resultados e impacto. **Revista do Serviço Público**, Brasília, v. 49, n. 2, p. 103-124, 1998.
- COUTINHO, R. G. **Subsídios para organizar avaliações da ação governamental**. Brasília: Ipea, 2001 (Textos para Discussão, n. 776).
- FREY, K. Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas no Brasil. **Planejamento e Políticas Públicas**, Brasília, Ipea, n. 21, 2002.
- GENERAL ACCOUNTING OFFICE (GAO). **Social research and development of limited use to national policymakers: Report to the congress**. Washington, 1977.
- GUIMARÃES, J. R. S.; JANNUZZI, P. M. IDH: indicadores sintéticos e suas aplicações em políticas públicas – uma análise crítica. **Revista Brasileira de Estudos Urbanos e Regionais**, Salvador, v. 7, n. 1, p. 73-89, 2005.

HOWLETT, M.; RAMESH, M. **Studying public policy: policy cycles and policy subsystems**. DonnMills: Oxford University Press, 2003.

JANN, W.; WEGRICH, K. Theories of the policy cycle. *In*: FISCHER, F. *et al.* (Ed.). **Handbook of public policy analysis: theory, politics and methods**. Boca Raton: CRC Press, 2007. p. 43-62.

JANNUZZI, P. M. **Indicadores sociais no Brasil: conceitos, fonte de dados e aplicações**. Campinas: Alínea, 2004.

_____. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil. **Revista do Serviço Público**, Brasília, v. 56, n. 2, p. 137-160, abr./jun. 2005.

LAVILLE, C.; DIONNE, J. **A construção do saber**. Porto Alegre: Artmed, Editora UFMG, 1999.

LINDBLOM, C. Todavía tratando de salir del paso. *In*: SARAVIA, E.; FERRAREZI, E. (Org.). **Políticas públicas: coletânea**. Brasília: ENAP, 2006. p. 99-122.

MADEIRA, F. R. A improvisação na concepção de programas sociais: muitas convicções, poucas constatações – o caso do primeiro emprego. **São Paulo em Perspectiva**, São Paulo, v. 18, n. 2, p. 78-94, 2004.

MILLER, D. C. **Handbook of research design and social measurement**. Nwbury Park: Sage, 1991.

MOHR, L. B. **Impact analysis for program evaluation**. Thousand Oaks: Sage, 1995.

NASCIMENTO, B. **A reforma universitária de 1968: origem, processo e resultados de uma política pública para o ensino superior**. 1991. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

OWEN, J. M. **Program evaluation: forms and approaches**. New York: Guilford Press, 2007.

PAES-SOUSA, R.; VAITSMAN, J. (Org.). **Metodologias e instrumentos de pesquisas de avaliação de programas do MDS: Bolsa Família, assistência social, segurança alimentar e nutricional**. Brasília: MDS/Sagi, 2007.

PARADA, E. L. Política y políticas públicas. *In*: SARAVIA, E.; FERRAREZI, E. (Org.). **Políticas públicas: coletânea**. Brasília: ENAP, 2006. p. 67-95.

RICHARDSON, R. J. *et al.* **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 2002.

ROSSI, P. H. *et al.* **Evaluation: a systematic approach**. Thousand Oaks: Sage, 2004.

RUA, M. G. Algumas considerações preliminares sobre política e políticas públicas. *In*: RUA, M. G.; CARVALHO, M. I. V. (Org.). **O estudo da política: tópicos selecionados**. Brasília: Paralelo 15, 1998.

_____. **A avaliação no ciclo da gestão pública**. Brasília: ENAP, 2004. Apostila do curso de Especialização em Políticas Públicas da Educação com Ênfase em Monitoramento e Avaliação – MPA. Mimeografado.

SARAVIA, E. Introdução à teoria da política pública. *In*: SARAVIA, E.; FERRAREZI, E. (Org.). **Políticas públicas: coletânea**. Brasília: ENAP, 2006. p. 21-42.

SARAVIA, E.; FERRAREZI, E. (Org.). **Políticas públicas: coletânea**. Brasília: ENAP, 2006.

TRIBUNAL DE CONTAS DA UNIÃO (TCU). **Avaliação do TCU sobre o Programa Nacional Biblioteca na Escola**. Brasília, 2002.

VIEIRA, F. S. Qualificação dos serviços farmacêuticos no Brasil: aspectos inconclusos da agenda do Sistema Único de Saúde. **Revista Panamericana de Salud Pública**, v. 24, n. 2, p. 91-100, 2008.

VILLANUEVA, L. F. A. Estudio introdutório. *In*: SARAVIA, E.; FERRAREZI, E. (Org.). **Políticas públicas: coletânea**. Brasília: ENAP, 2006. p. 43-65.

WEISS, C. **Evaluation research**. New Jersey: Prentice Hall, 1972.

WORTHERN, B. R. *et al.* **Avaliação de programas: concepções e práticas**. São Paulo: EDUSP, Gente, 2004.

