

ESTUDO TÉCNICO
N.º 07/2013

Monitoramento e avaliação de programas sociais no Brasil e América Latina:
apontamentos conceituais, considerações metodológicas e reflexões sobre as práticas

MDS

SAGI

MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E COMBATE À FOME

SECRETARIA DE AVALIAÇÃO E GESTÃO DA INFORMAÇÃO

ESPLANADA DOS MINISTÉRIOS BLOCO A SALA 307

CEP 70054-960 BRASÍLIA DF TEL. 61 2030 1501

Estudo Técnico

No. 07/2013

Monitoramento e avaliação de programas sociais no Brasil: apontamentos conceituais, considerações metodológicas e reflexões sobre as práticas

Responsável técnico pelo estudo

Paulo de Martino Jannuzzi

Estudos Técnicos SAGI é uma publicação da Secretaria de Avaliação e Gestão da Informação (SAGI) criada para sistematizar notas técnicas, estudos exploratórios, produtos e manuais técnicos, relatórios de consultoria e reflexões analíticas produzidas na secretaria, que tratam de temas de interesse específico do Ministério de Desenvolvimento Social e Combate à Fome (MDS) para subsidiar, direta ou indiretamente, o ciclo de diagnóstico, formulação, monitoramento e avaliação das suas políticas, programas e ações.

O principal público a que se destinam os Estudos são os técnicos e gestores das políticas e programas do MDS na esfera federal, estadual e municipal. Nesta perspectiva, são textos técnico-científicos aplicados com escopo e dimensão adequados à sua apropriação ao Ciclo de Políticas, caracterizando-se pela objetividade, foco específico e tempestividade de sua produção.

Futuramente, podem vir a se transformar em artigos para publicação: Cadernos de Estudos, Revista Brasileira de Monitoramento e Avaliação (RBMA) ou outra revista técnica-científica, para alcançar públicos mais abrangentes.

Palavras-chave: *Avaliação; Monitoramento; Políticas Públicas*

Unidade Responsável

Secretaria de Avaliação e Gestão da Informação

Esplanada dos Ministérios | Bloco A | Sala 307

CEP: 70.054-906 Brasília | DF

Fone: 61 2030-1501 | Fax: 2030-1529

www.mds.gov.br/sagi

Secretário de Avaliação e Gestão da Informação

Paulo de Martino Jannuzzi

Secretária Adjunta

Paula Montagner

Apresentação

O presente Estudo Técnico apresenta uma discussão conceitual e metodológica no campo de Monitoramento e Avaliação, além de reflexão sobre as experiências práticas no Brasil e outros países da Iberoamérica. Traz subsídios para especificação e estruturação de painéis de indicadores de monitoramento e pesquisas de avaliação de programas sociais¹.

1. Introdução

O processo de estruturação de Sistemas de Monitoramento e Avaliação (SM&A) de Políticas, programas ou mais amplamente, da ação governamental é um empreendimento complexo, como bem e extensamente revelam a literatura internacional sobre o tema e as experiências vivenciadas – e muitas vezes sofridas – por que passam os gestores e técnicos públicos que procuram colocá-los em pé.

Afinal, um SM&A pode ser entendido, em um sentido lato, como o conjunto de atividades – articuladas, sistemáticas e formalizadas – de produção, registro, acompanhamento e análise crítica de informações geradas na gestão de políticas públicas, de seus programas, produtos e serviços, por parte das organizações, agentes e público-alvo envolvidos, com a finalidade de subsidiar a tomada de decisão com relação aos esforços necessários para a melhoria da ação pública. Se trata, por tanto, de um conjunto de atividades inerentes ao ciclo de gestão e produção das Políticas Públicas, direcionadas à sistematização de informação acerca dos aspectos considerados críticos para o êxito dos programas.

Ademais de prover informação para análise das “entregas” de produtos e serviços e para a correção de eventuais falhas da ação governamental, os SM&A também se destinam a gerar conhecimentos a respeito dos impactos das políticas e programas, assim como de custos de produção dos mesmos. Sistemas de M&A seriam, portanto, instrumentos importantes para garantir maior transparência da

¹ Este texto é uma revisão ampliada de outro, intitulado “Estruturação de sistemas de monitoramento e especificação de pesquisas de avaliação, os problemas dos programas públicos no Brasil são”, desenvolvido para uso no Curso de Evaluación de Programas no âmbito da Escuela Iberoamericana de Administración y Políticas Públicas, no mês de agosto de 2010 em Brasília, promovido pelas Escola Nacional de Administração Pública.

forma de utilização dos recursos públicos e para subsidiar decisões sobre o mérito e relevância das políticas y programas na repartição orçamentária.

Os SM&A podem se diferenciar segundo vários critérios, como discutem MacDavid&Hawthorne (2006), Mackay (2007), Owen (2007), Cunill y Ospina (2008). Podem ser mais dirigidos à melhoria interna de programas, gestão e decisão orçamentárias ou prestação de contas da ação governamental ao parlamento ou à sociedade. O alcance conceitual e programático dos sistemas de M&A podem ser macro-contextual quando fazem referência a políticas ou estratégias de governo; meso-contextual quando dirigidas ao acompanhamento gerencial e analítico de programas e organizações; e mesmo micro-ambiental quando a ênfase da avaliação recai sobre indivíduos, processos o produtos muito específicos. Podem ainda abarcar a integralidade de ações de governo, algumas áreas setoriais ou um conjunto seletivo de projetos prioritários.

Dependendo de onde estão localizados (Ministério da Fazenda, Ministério de Planejamento, Ministérios setoriais, assessorias de altos dirigentes, Casa Civil, órgão de controle interno ou externo) e da cultura existente em M&A no governo, estes sistemas podem dispor de um conjunto mais amplo ou restrito de ferramentas e produtos: mapas estratégicos, balanços de gestão, sistemas informatizados, indicadores de desempenho, indicadores de políticas e programas, painéis de indicadores, painéis situacionais, pesquisas de satisfação, desenhos experimentais ou *quasi*-experimentais, grupos focais, entre outros.

Como se vê, a estruturação de sistemas de M&A envolvem decisões de natureza política, administrativa e técnica. Aportar reflexões sobre esse último aspecto é o que este texto se propõe. Mais especificamente, este texto busca trazer alguma contribuição de natureza conceitual e metodológica para ajudar na estruturação dos sistemas de monitoramento e na especificação de pesquisas de avaliação que realmente se prestem ao objetivo de permitir o seguimento contínuo e aperfeiçoamento efetivo da ação governamental e, mais propriamente, a gestão de programas sociais, a partir da perspectiva de um Ministério setorial (ou órgão equivalente, voltado a um campo específico da Política Social).

Inicia-se com uma discussão sobre os desafios da legitimação e estruturação que os SM&A vem passando no Brasil e outros países. Depois, faz-se uma exposição do ciclo de formulação e avaliação de políticas públicas e programas, para, em seguida, discutir as especificidades dos sistemas de indicadores de monitoramento e das pesquisas de avaliação; de como e devem se integrar.

Este texto parte da hipótese de que as dificuldades apontadas no monitoramento e avaliação de programas no Brasil – e as frustrações daí decorrentes – 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 dos instrumentos de monitoramento e avaliação das ações governamentais. Isso leva às conhecidas frustrações com resultados tão ansiosamente esperados e aos questionamentos sobre a utilidade dos estudos.

Explicitando melhor a motivação para elaboração desse texto – agora, sob o risco de simplificar demais a natureza e minimizar o alcance dos estudos avaliativos conduzidos por colegas pesquisadores nas universidades e centros de pesquisa no país –, a abordagem marcadamente disciplinar da pesquisa aplicada no campo, a sobrevalorização de algumas abordagens e modelos específicos de avaliação, 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 contribui para conformação de tal quadro. Este texto procura, assim, em uma perspectiva modesta em substância, mas comprometida no seu sentido público, 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 pela comunidade de pesquisadores acadêmicos e profissionais.

2. O monitoramento e a avaliação no Brasil e Iberoamérica: um campo em processo de legitimação e estruturação

A crescente atuação do Estado Brasileiro e de diversos países da Iberoamérica na promoção de políticas sociais, em especial a partir da Constituição Federal de 1988 (no caso do Brasil), 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 na área da saúde e na área da qualificação profissional, ou entre equipes de gestores do governo federal e técnicos de prefeituras de pequenos municípios no interior do país.

Há mesmo quem diga – nos círculos acadêmicos e nas diferentes esferas da administração pública – que não é por falta de estudos, diagnósticos e indicadores que os programas sociais não conseguem alcançar os resultados e impactos esperados no Brasil. Afinal, é fato que centros de pesquisa, universidades, empresas de consultoria e equipes de técnicos do setor público têm produzido, de forma sistemática ou por meio de projetos contratados, um amplo conjunto de trabalhos, artigos e estudos sobre diferentes aspectos da realidade social, referidos direta ou indiretamente ao contexto de operação dos vários programas públicos. Entretanto, também é verdade que muitos desses trabalhos, mesmo os contratados para subsidiar a formulação de programas específicos, acabam se revelando como diagnósticos descritivos bastante gerais, aplicáveis a diferentes programas sociais,

com contribuições certamente relevantes, mas desprovidas de informações mais particulares e “customizadas” para formatação de ações públicas mais dirigidas, no grau que se requer atualmente. Nos diagnósticos contratados pelo setor público, contribuem para isso, sem dúvida, as imprecisões ou lacunas dos termos de referência – que espelham, muitas vezes, a falta de clareza dos objetivos e natureza do programa público a ser implementado, fato decorrente, por sua vez, das deficiências de formação do gestor público – e a exiguidade de tempo e recursos para realização de levantamentos de campo específicos. Nesse contexto, acabam se justificando estudos diagnósticos mais gerais, baseados em trabalhos anteriores e em fontes de dados e pesquisas já disponíveis.

Também 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 inexistência de sistemas de informações de monitoramento, seja pelas decisões metodológicas quanto às técnicas, amostras e casos investigados, muitas vezes escolhidos mais em função dos prazos e 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 pesem 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 (EUA) nos anos 1970, a julgar pelo relatório do General Accounting Office que, mobilizado por demanda do Senado americano em 1974, constatou que as informações das pesquisas sociais e estudos avaliativos financiados com recursos públicos não contribuíam para o desenho das políticas públicas, pela dispersão e fragmentação de temas investigados, falta de coordenação e orientação do que era de interesse governamental investigar, dificuldade de acesso e compreensão dos relatórios de pesquisas(GAO 1977).

Mesmo mais recentemente têm-se constatado limitações e problemas nos estudos avaliativos de programas públicos nos EUA, às vezes com desdobramentos até piores que os descritos por Cotta. De fato, como colocam Worthern et al. (2004):

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 implementação de programas por parte de técnicos de alto escalão que, ao não incorporar 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 esquece-se 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 mesmo conjunto de serviços, de 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.

De fato, um dos achados sistemáticos das avaliações de programas realizadas pelo Tribunal de Contas da União é a inexistência de sistemas de informação para acompanhamento das atividades dos programas. É ilustrativa, nesse sentido, a constatação, reproduzida a seguir, acerca da avaliação do Programa Nacional Biblioteca na Escola (PNBE), operado pelo Fundo Nacional de Desenvolvimento Escolar (FNDE) do Ministério da Educação (MEC), com objetivo de distribuir acervo de livros para formação de bibliotecas nas escolas.

O trabalho constatou que o FNDE tem mostrado grande eficácia operacional na distribuição dos acervos do Programa para as escolas beneficiadas. Foram atendidas 20 mil escolas em 1998 e 36 mil em 1999, com previsão de atendimento de 139 mil escolas em 2002.

Todavia, também se verificou que o PNBE não tem atividades de monitoramento e avaliação bem estruturadas. O MEC não dispõe de informações que permitam conhecer o nível de utilização dos acervos, bem como os problemas que podem estar afetando a efetividade do uso dos livros no desenvolvimento escolar dos alunos nas escolas beneficiadas. Tal

fato é mais preocupante justamente no momento em que aumenta a escala de atendimento do Programa. (TCU, 2002, p.10).

Para isso, contribui o estágio ainda incipiente do conhecimento do campo aplicado de monitoramento e avaliação de programas no Brasil, seja dentro da estrutura pública, seja nas universidades e nos centros de pesquisa. Enquanto nos EUA 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, com profissionais com formação multidisciplinar e com 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 naquele 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). A avaliação de programas, que se concentrava, inicialmente, nas áreas de educação (sobretudo na investigação de programas de alfabetização), de saúde pública (nas ações de combate à mortalidade por doenças infecto-parasitárias) e de qualificação profissional, passa, no pós-guerra, a abranger áreas como prevenção da violência juvenil, planejamento familiar, nutrição, programas habitacionais, 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. Afinal, 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 os despesas militares da Guerra do Vietnã, no governo Nixon, e para 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 criação de uma comunidade profissional de avaliadores².

A avaliação de programas deixaria de ser preocupação apenas de cientistas sociais, em temas de pesquisas acadêmicas nas universidades, ganhando projeção e interesse na comunidade de gestores públicos como instrumento para aprimoramento dos serviços públicos. Essa mudança qualitativa da produção e apropriação de conhecimentos na área – que parece se processar neste momento no Brasil – é muito bem registrada pela seguinte passagem de Rossi et al. (2004):

In its early years, evaluation was shaped mainly by the the interests of social researchers. In later stages, however, the consumers of evaluation research exercised a significant influence on the field. Evaluation is now sustained primarily by funding from policymakers, program planners, and administrators who use the findings and by the interests of general public and the clients of the programs evaluated. Evaluation results may not make front-page headlines, but they are often matters of intense concern to informed citizens, program sponsors, and decision makers, and those whose lives are affected, directly or indirectly, by the programs at issue. (ROSSI et al., 2004,p.9).

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

² Vide, por exemplo, as publicações, eventos e sites das associações profissionais ou acadêmicas de avaliação nesses países, entre os quais: American Evaluation Association (www.eval.org); Canadian Evaluation Society (www.evaluationcanada.ca); European Evaluation Society (www.europeanevaluation.org); e Societé Française d’Evaluation (www.afe.asso.fr).

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 os programas procuram atuar³. 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 conhecimento mais circunstanciado dos problemas, as soluções idealizadas para mitigá-los e para conferir legitimidade técnica-científica na avaliação da ação governamental. Mas, como já alertado em um dos primeiros 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 questões que são demandadas ao avaliador responder, pela menor ou maior disponibilidade de dados, 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 com conhecimento técnico e metodológico apropriado à temática em questão, mas também com experiência empírica anterior e, sobretudo, com desprendimento para inovar e improvisar nas diversas circunstâncias e dificuldades que insistem em aparecer no cotidiano prático da área. Não é incomum, pois, que

³ De fato, no Brasil, os trabalhos de avaliação de programas públicos e seus autores encontram-se dispersos pelas principais associações científicas disciplinares, tais como: Anpad (administração e administração pública); Anpec (economia); Anpocs (ciências sociais); Anped (educação); Abep (estudos populacionais); Abet (estudos do trabalho); Abrasco (saúde coletiva). Vale destacar alguns dos principais periódicos em que se pode encontrar estudos avaliativos: Revista do Serviço Público; Revista Brasileira de Ciências Sociais; Revista de Administração Pública; Revista São Paulo em Perspectiva; Textos de Discussão (Ipea); Planejamento e Políticas Públicas; Pesquisa de Planejamento Econômico; Texto de Discussão Ence, entre outros (alguns desses periódicos estão disponíveis no portal www.scielo.br). 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 (www.abave.org.br); a Rede Brasileira de Monitoramento e Avaliação (<http://redebrasileirademea.ning.com>); o portal Políticas Públicas em Foco (www.boletim-fundap.cebrap.org.br), mantido pela Fundap e Cebrap; o sítio do Tribunal de Contas da União (www.tcu.gov.br) e da Controladoria-geral da União (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 de pós-graduação em administração pública, economia social, saúde coletiva, direito etc.

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.

Essas considerações iniciais – um tanto extensas, mas necessárias para contextualização e justificativa da discussão aqui pretendida – ilustram as dificuldades ou deficiências dos sistemas de monitoramento e os frequentes problemas de especificação de pesquisas e avaliação para permitir o acompanhamento das ações e programas governamentais no Brasil e possivelmente outros países da Iberoamérica. Muitos são os exemplos de pesquisas de avaliação que, mesmo motivadas por preocupação legítima e meritória com a eficácia e efetividade do gasto público, revelam, infelizmente, resultados triviais, metodologicamente questionáveis ou com baixa apropriação para reformulação dos programas. Equacionar tal situação requer, ademais de decisões políticas e administrativas, aprofundar o conhecimento sobre aspectos técnicos dos sistemas de M&A. É o que se apresenta nas seções seguintes.

3. O ciclo de Políticas e Programas Públicos

Antes de passar a uma explicação mais detalhada sobre o ciclo de Políticas e suas etapas, vale formalizar dois termos citados correntemente no texto – políticas públicas e programas. Isso é importante para circunstanciar os limites do campo de diálogo estabelecido neste texto, já que avaliação de políticas públicas e avaliação de programas públicos são termos muito imbricados, mas referem-se a contextos muito diferentes (em amplitude) da análise da intervenção estatal.

Evitando entrar em uma discussão muito extensa sobre os diferentes significados do termo – sistematizados em Villanueva (2006) – política pública é, na definição de Nascimento (1991), o conjunto de decisões tomadas por aqueles que detêm competência legal para deliberar em nome da coletividade – as instituições de Estado – visando à solução de um problema ou ao redirecionamento de uma

tendência, com a intenção de orientar sua evolução para um fim estabelecido como o desejável. Como esclarece Saravia (2006), em

[...] uma perspectiva mais operacional, poderíamos dizer que ela [a política pública] é um sistema de decisões públicas que visa a ações ou omissões, preventivas ou corretivas, destinadas a manter ou modificar a realidade de um ou vários setores da vida social, por meio da definição de objetivos e estratégias de atuação e da alocação dos recursos necessários para atingir os objetivos estabelecidos. (SARAVIA, 2006, p.29).

As decisões que conformam tal política pública – distributiva, redistributiva ou regulatória, universal ou focalizada – assumem diferentes formatos: podem ser enunciados de diretrizes estratégicas de governo, leis e decretos normativos, que especificam de forma mais clara e operacional tais diretrizes, criação de organizações ou programas que vão conferir operacionalidade às diretrizes, leis e normas. (VILLANUEVA, 2006).

O programa público é, pois, um dos instrumentos operacionais das políticas públicas. Mais especificamente, trata-se de um conjunto sistêmico de ações programadas e articuladas entre si, com objetivo de atender uma demanda pública específica, encampada na agenda de prioridades de políticas públicas do Estado ou governo. Assim, um programa social é um conjunto de atividades direcionadas para solucionar um problema vivenciado pela sociedade, no seu todo ou por grupos. (ROSSI et al., 2004).

Avaliação de políticas públicas, ou melhor, análise de políticas públicas, termo preferido por Owen (2007), é, pois, uma atividade muito mais ampla que a avaliação de programas. Trata do contexto político-social de surgimento da política, dos atores participantes. Volta-se mais ao esclarecimento de seu processo de construção do que da recomendação prática de como aprimorá-la, constituindo, na realidade, um campo de investigação mais propriamente acadêmico que a perspectiva técnica-profissional em que se realiza a avaliação de programas. Vale observar que há Programas, inclusive, com escopo temático e escala de operação muito mais abrangente que o usual, revelando-se mais como um guarda-chuva de outros programas mais específicos. Outros programas podem ser projetos de curto alcance, ou atividades que visam à produção de um produto ou serviço finalístico ou da área-

meio. A discussão aqui empreendida é suficientemente ampla para contemplar essas modalidades.

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 no Diagrama 1⁴. Em que pese as críticas de longa data quanto à forma simplificada com que esse diagrama apresenta o processo político e sua própria veracidade empírica, a separação em etapas se presta aos objetivos de evidenciar, ao longo do processo, ênfases diferenciadas no planejamento, operação ou avaliação dos programas. Justifica-se ainda para fins didáticos e para orientar o recorte analítico na pesquisa acadêmica na área (NASCIMENTO, 1991).

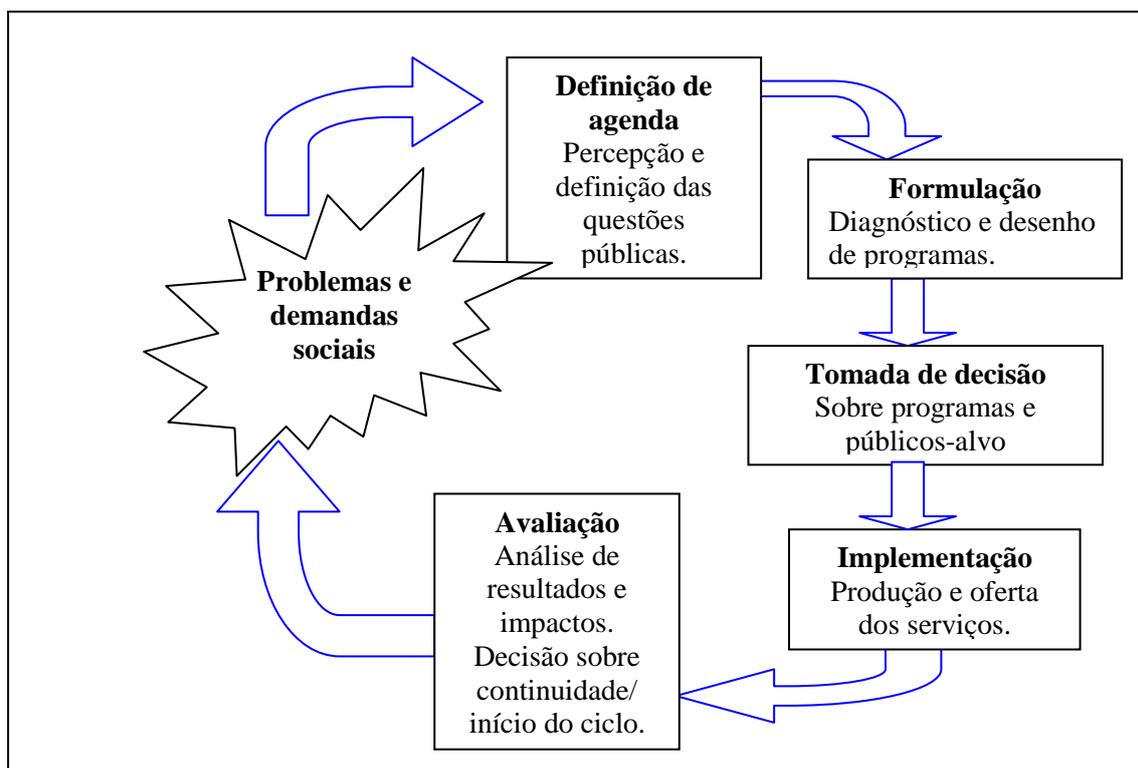
Nesse modelo, a primeira etapa – Definição da Agenda Política (Agenda-Setting) – corresponde aos múltiplos caminhos e processos que culminam com o reconhecimento de uma questão social como problema público e da necessidade da ação governamental para sua solução; isto é, a legitimação da questão social na pauta pública ou agenda das políticas públicas do país, em determinado momento. A etapa seguinte – Formulação de Políticas e Programas (Policy Formulation) – refere-se aos processos e 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 implementação da ação governamental, na alocação de recursos e desenvolvimento dos processos previstos nas alternativas e programas escolhidos anteriormente. Por fim, é preciso analisar se os esforços empreendidos estão atuando no sentido esperado de solucionar o problema original – etapa de Avaliação das Políticas e Programas (Policy Evaluation). É necessário avaliar se é preciso realizar mudanças

⁴ Como apresentado nos vários textos reunidos na valiosa coletânea organizada por Saravia e Ferrarezi (2006) em políticas públicas, publicada pela ENAP e disponibilizada em seu sítio (www.enap.gov.br), o Ciclo de Políticas Públicas pode ser descrito com um número maior ou menor de etapas. De modo geral, distinguem-se, pelo menos, três macroetapas: formulação (no qual estaria a formação da agenda), a implementação e a avaliação. Essa publicação, organizada segundo as etapas do ciclo, traz artigos clássicos que aprofundam a vasta discussão envolvida em cada uma delas.

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.

Diagrama 1: O Ciclo de Políticas e Programas Públicos



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, à lista de preocupações da sociedade ou 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 a outros 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, 1997). 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 “digere” inovações a seu próprio tempo e estilo⁵.

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 & Ramesh (2003). Vai sendo formada por meio de diferentes mecanismos de pressão externa de grupos organizados em partidos políticos, sindicatos, associações patronais, 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 os 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. Outro ator político de grande

⁵ A resistência à mudança e à incorporação de inovações não é problema só no setor público, mas também nas empresas privadas. Os mecanismos de motivação, pressão e coerção nessas últimas acabam mobilizando mudanças, às vezes a custos pessoais e institucionais elevados. No setor público, os instrumentos para mobilização envolvem, em tese, mais coordenação de esforços e convencimento, até porque as resistências podem ser bastante legítimas. Ou seja, nem toda mudança é necessariamente uma inovação “positiva”. Para o bem ou para o mal, são conhecidos os casos em que a resistência da burocracia pública acaba revertendo ou minando propostas consideradas muito ousadas.

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 desses, 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, 1997). Enfim, as organizações políticas – partidos, sindicatos, grupos de interesse – são fundamentais para que as demandas entrem na agenda política do governo e para que, lá presentes, possam se transformar em ações e programas concretos.

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 concretude que conferem aos problemas sociais deles decorrentes.

Na etapa da formulação da política, os problemas, propostas e demandas explicitados na agenda transformam-se em leis, programas e propostas de ações. Isto é, a formulação de política envolve a busca de possíveis soluções para as questões priorizadas na agenda. Nesse momento, os elementos operacionais da política em questão precisam ser explicitados: diretrizes estratégicas, propostas de leis, decretos normativos, 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, de 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 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 junto com as pressões para introduzir a questão na agenda formal de governo. Se não, colocada a questão, há sempre atores com maior ou menor influência na formatação dos possíveis programas e ações (JANN; WEGRICH, 2007).

Levantadas as diferentes alternativas para uma dada questão social, é preciso escolher qual deverá ser adotada. Enquanto que na definição da agenda política e formulação de programas a participação de agentes não pertencentes ao governo ou Estado desempenha papel central, a tomada de decisão envolve os gestores e técnicos mais diretamente relacionados à política ou ao programa. Esse é 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. Como bem colocam Howlett & Ramesh (2003):

These actors can and do, of course, engage in various kinds of lobbying activities aimed at persuading, encouraging, and sometimes even coercing authoritative office-holders to adopt options of which they approve. However, unlike office-holders, those other actors have, at best, a voice in the decision-making process, but they do not have a vote per se. (HOWLETT & RAMESH, 2003, p.163).

Chega então o momento da implementação dos programas, projetos e ações pontuais ou mais abrangentes. A fase de Implementação 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. Aqui a política adquire uma concretude objetiva maior e, por isso, começa a revelar mais explicitamente as dificuldades de sua introdução e 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 ações realizadas, 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 programas futuros. Como bem coloca Frey (1997), a 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, 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 programas, sendo “[...] más habitual que cambien o se combinem com outras. Se há llegado a decir que lãs 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. Como observam Jann & Wegrich (2007):

Policy process rarely features clear-cut beginnings and endings. At the same time, policies have always been constantly reviewed, controlled, modified and sometimes even terminated; policies are perpetually reformulated, implemented, evaluated and adapted.... Moreover, policies do not develop in a vacuum, but are adopted in a crowded policy space that leaves little space for innovation[...] (JANN; WEGRICH, 2007, p.44-45)

Contudo, na visão de Howlett & 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. Esse modelo seria também suficientemente geral para ser aplicado no entendimento do processo na maioria dos âmbitos e contextos de

formulação de políticas e programas– em nível federal, estadual, local e mesmo setorial.

Jann & Wegrich (2007) acrescentam ainda que o sucesso e a “resiliência” desse modelo clássico às críticas de pesquisadores acadêmicos é que ele parece como um modelo 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, com controle efetivo das atividades na implementação dos programas públicos e objetividade na avaliação desses.

4. Os sistemas de indicadores de monitoramento no Ciclo de Políticas e Programas

A preocupação com a construção de indicadores de monitoramento da ação governamental é tão ou mais antiga que a própria avaliação de programas públicos, se forem consideradas as medidas de performance mais geral do setor público baseadas na entrega de obras e produtos e na computação de indicadores de custos de serviços, a partir da execução orçamentária. Como registram McDavid e Hawthorn (2006) em livro que procura oferecer uma visão integrada entre avaliação de programas e medição do desempenho governamental:

While we have tended to situate the beginnings of performance measures in the United States, in the 1960s, with the development of performance management systems such as programmed planned budgeting systems (PPBS)[...] there is good evidence that performance measurement and reporting was well-developed in some American local governments early in the 20th century. (MCDAVID; HAWTHORN, 2006, p.283).

Os autores citam as iniciativas de alguns governos locais nos EUA, antes da 2ª Guerra Mundial, na elaboração de relatórios com computação regular de medidas de eficiência dos serviços públicos então oferecidos, como a construção e manutenção das ruas e rodovias, distribuição de água e recolhimento de lixo. Os produtos e serviços disponibilizados pelo setor público municipal eram bastante simples e tangíveis, assim como a contabilização dos custos para sua produção.

Mas é a introdução das técnicas de programação orçamentária por programas no Governo Kennedy que marcaria um momento de maior

aprimoramento na medição do desempenho – expresso por indicadores de eficiência e efetividade – na esfera federal americana. Ao propor a estruturação do orçamento segundo conjunto de atividades e programas afins – agora bem mais complexos que os oferecidos pelos municípios na primeira metade do século XX – e não segundo as agências e organizações que as executavam, esperava-se conseguir apurar resultados e custos mais consistentes dos serviços públicos disponibilizados e, portanto, medidas de eficiência mais precisas. (MCDAVID; HAWTHORN, 2006).

Em que pesem o fracasso na implementação do orçamento-programa – na vinculação entre recursos-produtos ao nível de programas, e o contexto de crise fiscal do estado, comentam os autores que a orientação voltada à medição dos resultados ganhou força nos anos 1970 e na década seguinte, Os governos conservadores desse período – em especial na Inglaterra, com Thatcher, e nos EUA, com Reagan – imprimem reformas no setor público, que vieram a ser conhecidas como o movimento da Nova Administração Pública, preconizando a adoção de instrumentos de controle e gestão de resultados do setor privado.

Sem entrar na discussão acerca dos excessos, limitações e aspectos meritórios de tal movimento – debate esse já empreendido no Brasil por vários autores, com muito mais competência, autoridade e espaço, do que o disponível para este texto –, o legado de preocupação com a transparência e responsabilização pública criou oportunidades para aprimoramento da gestão de programas, seja para a estruturação de sistemas de indicadores de monitoramento, seja para a realização das pesquisas de avaliação. Neste tópico do texto, discute-se o primeiro desses instrumentos de gestão: os sistemas de monitoramento.

O modelo idealizado de “ciclo de vida” de políticas e programas apresentado anteriormente é uma referência conceitual interessante para ilustrar como os sistemas de indicadores de monitoramento podem ser estruturados e como as pesquisas de avaliação podem ser especificadas de forma a potencializar seu emprego na gestão dos programas. A proposta básica desta seção é mostrar a importância de se dispor de indicadores relevantes e periodicamente atualizados para acompanhar as atividades e a produção de serviços dos programas, que permitam corrigir desvios, reprogramar atividades ou mesmo especificar pesquisas

de avaliação para entender por que determinados processos não estão se encaminhando no sentido idealizado originalmente.

Como discutido em texto anterior⁶, as atividades de cada etapa do ciclo apoiam-se em um conjunto específico de indicadores. Na definição da agenda, os indicadores são recursos valiosos para dimensionar os problemas sociais, servindo como instrumentos de advocacy e pressão de demandas sociais não satisfeitas. Indicadores produzidos pelas instituições oficiais de estatísticas – sobretudo os computados a partir de censos demográficos e pesquisas amostrais regulares – prestam-se bem a esse papel, pela legitimidade que gozam perante diferentes públicos. No Brasil, o relatório “Síntese de Indicadores Sociais”, publicado anualmente pelo IBGE, é uma referência importante nesse sentido, provocando grande interesse na mídia quando de sua divulgação⁷. O Índice de Desenvolvimento Humano (IDH) computado para países, municípios, mulheres e negros é outro bom exemplo de como uma medida simples pode mobilizar os veículos de comunicação e recolocar anualmente o debate sobre desigualdade social e os impactos das políticas públicas na mídia e nas esferas de governo⁸.

Na elaboração dos diagnósticos para formulação de programas, os indicadores são imprescindíveis para qualificar os públicos-alvo, localizá-los e retratá-los de modo tão amplo e detalhado quanto possível. É preciso adequar as intervenções às características e necessidades dos demandantes dos programas. De modo geral, na elaboração de diagnósticos socioeconômicos propositivos para programas públicos empregam-se indicadores de várias áreas temáticas analíticas ou de atuação governamental. Para desenvolver programas no campo da educação básica, por exemplo, é preciso conhecer não apenas as condições de acesso à escola, a infraestrutura escolar, o desempenho dos alunos, mas também aspectos relacionados às condições de vida dos estudantes, como moradia, nível de pobreza e

⁶ Jannuzzi (2005), disponível para *download* na Revista do Serviço Público, v.56, n.2, em www.enap.gov.br. Outra referência para aprofundamento nesse campo é Jannuzzi (2004).

⁷ Essa publicação do IBGE, assim como diversas outras da instituição, está disponível para *download* em www.ibge.gov.br.

⁸ Uma discussão dos limites e potencialidades do IDH pode ser vista em Guimarães e Jannuzzi (2005).

rendimento familiar, acesso a serviços de saúde, escolaridade dos pais, fatores esses que certamente podem afetar ou potencializar as ações programáticas específicas⁹.

Indicadores construídos a partir dos censos demográficos são particularmente úteis nessa fase, pela amplitude do escopo temático investigado e pela possibilidade de desagregação territorial ou por grupos sociodemográficos específicos. O Censo Demográfico 2010 potencializou ainda mais essas duas características – escopo e desagregabilidade –, constituindo-se em marco importante para reavaliação das demandas sociais da população brasileira, nesse momento de ampliação da estrutura de proteção social no país.

Na seleção de alternativas programáticas idealizadas para atender à questão pública colocada na agenda e eleger prioridades de intervenção, é preciso dispor de indicadores que operacionalizem os critérios técnicos e políticos definidos. Indicadores sintéticos, como o já citado Índice de Desenvolvimento Humano, o Índice de Desenvolvimento da Educação Básica, o Índice da Pegada Humana (Footprint Index) ou as classificações geradas por técnicas de análise multicritério podem ser úteis para apoiar decisões nessa fase do ciclo, como discutido em Scandar et al. (2006) e Jannuzzi et al. (2009)¹⁰.

Para acompanhar a implementação dos programas e, posteriormente, para gestão, quando eles entrarem em plena operação, são necessários indicadores que permitam acompanhar regularmente as ações programadas, do provisionamento e alocação do gasto à produção dos serviços e, idealmente, aos eventuais resultados e impactos esperados quando da idealização dos programas. São os indicadores que estruturam processos formais de monitoramento, entendido, nas palavras de Coutinho (2001, p.11) como “um processo sistemático e contínuo que, produzindo informações sintéticas e em tempo eficaz, permite rápida avaliação situacional e a intervenção oportuna que confirma ou corrige as ações monitoradas”.

⁹ Um bom exemplo de indicadores para diagnósticos de programas setoriais é demonstrado na publicação “Construindo o diagnóstico municipal”, disponível em www.cepam.sp.gov.br. Outro exemplo de diagnóstico apoiado em indicadores multitemáticos é o Diagnóstico para Programa de Qualificação Profissional do Estado de São Paulo, disponível em www.emprego.sp.gov.br. Os Relatórios de Informação do MDS também são outros bons exemplos (www.mds.gov.br/sagi).

¹⁰ No sítio www.anipes.org.br está disponível um aplicativo (Pradin) que permite entender o potencial da Análise Multicritério na tomada de decisão em políticas públicas.

Dispostos em painéis ou em sistemas informatizados, os indicadores de monitoramento devem ser atualizados com regularidade e tempestividade adequada à tomada de decisão. Idealmente, esses indicadores devem ser específicos e sensíveis às ações programadas, a fim de se tornarem, de fato, úteis para inferir se o programa está sendo implantado conforme planejado ou para permitir as correções de rumo (aspectos que só poderão ser comprovados mais precisamente por meio de pesquisas de avaliação específicas, durante ou após a implantação).

Para isso, esses painéis ou sistemas de indicadores de monitoramento devem se valer dos dados continuamente atualizados nos cadastros e registros administrados pelos gestores e operadores envolvidos nos programas. Devem estar interligados aos sistemas informatizados de gestão do programa, no qual são registrados atendimentos prestados, informações dos agentes que o operam, características dos beneficiários, processos intermediários, que produzirão os efeitos idealizados pelo programa. Diferentemente do que se passa nas pesquisas de avaliação, na estruturação de sistemas de monitoramento não se prevê levantamentos primários de dados. Pode ser necessário criar rotinas de coleta de dados que operem fora dos círculos normais de produção de serviços dos programas, mas é preciso fazer esforços para aproveitar as informações geradas no âmbito de operação cotidiana dos programas. A criação de mais uma rotina para registro de informação, pelo beneficiário do programa ou pelo agente envolvido na implementação, pode implicar atrasos indesejados e, pior, declarações mal preenchidas que acabam não se prestando à sua finalidade original.

Vale observar que, em geral, as informações compiladas e enviadas para atualização dos registros de acompanhamento do Plano Plurianual, no Ministério do Planejamento, Orçamento e Gestão, são somente a ponta do iceberg de dados armazenados em planilhas eletrônicas e gavetas nos escritórios da administração pública. Aliás, é curioso que, nas médias e altas gerências do setor público – prefeituras, secretarias de estados ou ministérios –, reclame-se de falta de informação para gestão em meio a tanta informação continuamente produzida nas unidades de prestação de serviços públicos (escolas, postos de saúde, delegacias, centros de referência da assistência social, postos de intermediação de mão de obra, agências do INSS etc). Para superar esse paradoxo da “escassez na abundância”, é

preciso aprimorar os processos de gestão da informação nos escritórios em que se planejam e coordenam as políticas e programas públicos. Os fluxos de informação, os processos de tratamento, validação, classificação e, naturalmente, de armazenamento precisam ser pensados de forma articulada, valendo-se de aplicativos e ferramentas de integração de dados (e não dos custosos – em planejamento, tempo é dinheiro – sistemas informatizados de 4ª, 5ª ou enésima geração propostos por consultorias experientes em automação de processos na iniciativa privada, mas pouco conhecedoras da complexidade operacional da gestão de programas públicos).

É possível que algumas informações produzidas pelas agências estatísticas – no caso brasileiro, o IBGE, o Datasus, o Inep, a Secretaria de Avaliação do Ministério de Desenvolvimento Social, entre outros¹¹ – possam ser úteis para a construção de indicadores de monitoramento de programas, sobretudo de natureza universal ou com público-alvo numeroso. Em geral, essas fontes proveem informações para avaliações de políticas ou para um conjunto de políticas e macroações governamentais – mais adequadas para composição de mapas estratégicos da ação governamental – e não para os propósitos de monitoramento de programas específicos, em função da escala territorial de desagregação dos dados ou do tempo que levam para serem produzidas. Esse quadro pode mudar para melhor com as transformações anunciadas nas pesquisas amostrais do IBGE para a década de 2010, com maior integração conceitual, maior possibilidade de incorporação de temas suplementares, ampliação e adensamento da amostra pelo interior do país. Não só se poderá dispor de um escopo mais amplo de estatísticas e indicadores sociais divulgados mais regularmente ao longo do ano, como também referidos para domínios territoriais e grupos sociodemográficos bem mais diversos.

Aos indicadores mensais de emprego, hoje restritos às seis principais regiões metropolitanas, somar-se-ão muitos outros indicadores – educacionais, habitacionais, de rendimentos etc. – divulgados em base trimestral (semestral ou anual) para todos os estados brasileiros. Mediante o emprego de métodos

¹¹ As pesquisas e dados destas instituições podem ser consultadas, respectivamente, em www.ibge.gov.br, www.datasus.gov.br, www.inep.gov.br. No portal www.anipes.org.br, podem ser acessados sítios de órgãos subnacionais de estatística no Brasil.

estatísticos avançados, usando informação combinada dessas pesquisas com outras fontes de dados, pesquisadores de centros de análise de políticas públicas e universidades poderão estimar indicadores sociais com bom grau de precisão para domínios territoriais ainda mais específicos¹².

De qualquer forma, as informações estatísticas mais gerais são úteis para que se possam disponibilizar alguns indicadores de contexto socioeconômico no âmbito do sistema de monitoramento. Afinal, todo sistema aberto, como são os programas públicos, está sujeito aos efeitos de fatores externos, que podem potencializar ou atenuar resultados. Sistemas de indicadores de monitoramento de programas de qualificação profissional, por exemplo, devem dispor de indicadores de mercado de trabalho e de produção econômica como informações de contexto, pelos impactos que uma conjuntura econômica menos ou mais favorável podem ocasionar na operação do programa.

Como alternativa ou complemento às pesquisas estatísticas oficiais, é possível construir indicadores de contexto ou mesmo de monitoramento de programas, a partir dos registros administrativos de programas de grande cobertura populacional como o Cadastro Único de Programas Sociais do Ministério de Desenvolvimento Social e Combate a Fome; a Relação Anual de Informações Sociais e o Cadastro Geral de Empregados e Desempregados do Ministério do Trabalho; e o Cadastro Nacional de Informações Sociais do INSS/Ministério da Previdência. Naturalmente, os registros de provimento e execução orçamentária do Sistema Integrado de Administração Financeira do Governo Federal (Siafi) também são importantes fontes de informação para construção de indicadores de monitoramento (sobretudo porque permitem computar indicadores de regularidade de repasse de recursos, dimensão crucial para programas que envolvem diversas ações intermediárias para viabilizar a consecução das atividades mais finalísticas). Vale registrar que os ministérios responsáveis pela gestão desses cadastros têm feito

¹² Vide, nesse sentido, o que as agências estatísticas americanas produzem regularmente, acessando o portal www.fedstats.gov. O *Bureau of Labor Statistics* produz, por exemplo, estimativas mensais de taxa de desemprego para 372 áreas metropolitanas e taxas anuais para os mais de 3 mil condados e 50 maiores cidades americanas. O *Bureau of Census*, por sua vez, traz anualmente estimativas populacionais para mais de 20 mil localidades e indicadores sociais referidos a mais de 10 temáticas – de ocupação, moradia à educação – para municípios acima de 65 mil habitantes.

esforços importantes com o objetivo de disponibilizar as informações neles depositadas para pesquisadores e público em geral¹³.

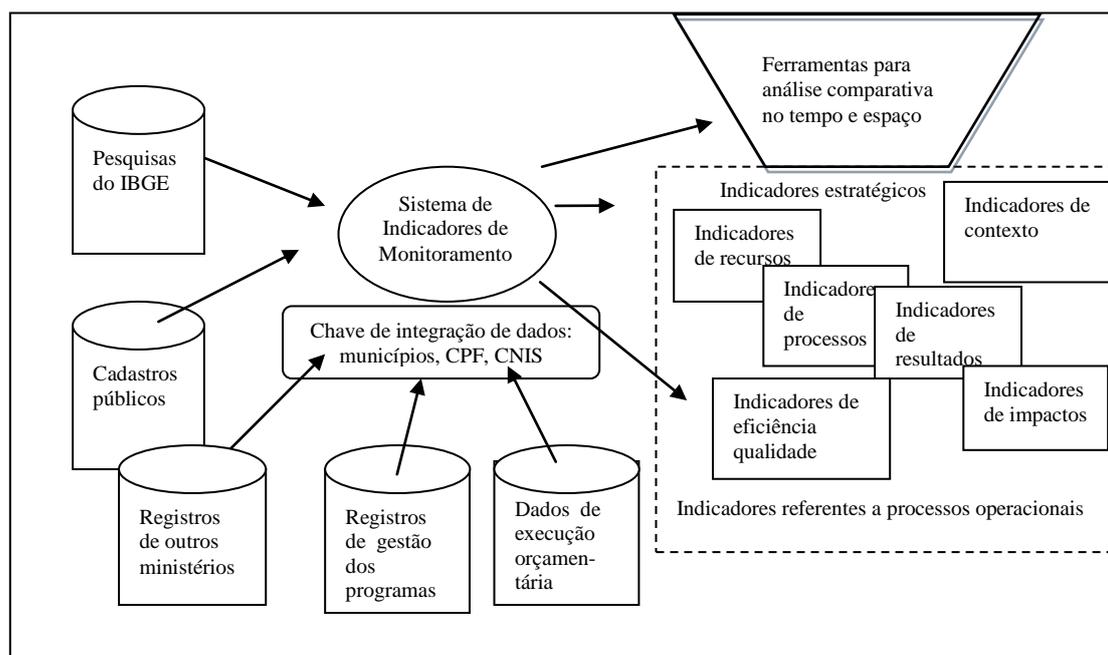
Um bom sistema de indicadores de monitoramento não é necessariamente composto de grande quantidade de informação, mas sim um sistema em que a informação foi selecionada de diferentes fontes e está organizada de forma sintetizada e mais adequada ao uso analítico pelos diferentes gestores (Diagrama 2). É preciso encontrar um ponto de equilíbrio entre o “caos informacional”, potencialmente gerado pela estruturação de sistemas de monitoramento construídos de baixo para cima (em que participam inicialmente técnicos e gestores da base e depois de níveis táticos e mais estratégicos), e a pobreza analítica das propostas desenvolvidas de cima para baixo. Um sistema de indicadores de monitoramento não é um sistema de gestão operacional do programa, que provê acesso aos incontáveis registros diários e individuais de operação de convênios, prestação de serviços, recursos transferidos, projetos e atividades concluídas. Um sistema de monitoramento vale-se do(s) sistema(s) de gestão dos programas para buscar informações, integrá-las segundo unidades de referência comum (município, escola etc.), sintetizá-las em indicadores e conferir-lhes significado analítico. Ao apresentar informações sintetizadas na forma de indicadores, que podem ser analisados no tempo, por regiões e públicos-alvo, ou que podem ser comparados com metas esperadas, os sistemas de monitoramento permitem ao gestor avaliar se os diversos processos e inúmeras atividades sob sua coordenação estão se “somando” no sentido preconizado. Um sistema de monitoramento não é, pois, um conjunto exaustivo de medidas desarticuladas, mas uma seleção de indicadores de processos e ações mais importantes.

Um sistema que não provê acesso orientado às centenas de indicadores disponíveis talvez não se preste ao propósito de monitoramento (ainda que possa ser útil como base de dados para estudos avaliativos a posteriori). Também não se presta ao monitoramento um sistema em que a informação não está organizada segundo o nível de relevância operacional-estratégica do gestor usuário. Ao gerente

¹³ Vide, nesse sentido, os aplicativos da Senarc/MDS (www.mds.gov.br), os sistemas de consulta on-line da Rais e do Caged no MTE (www.mte.gov.br), e os dados da Previdência na ferramenta Infologo em www.dataprev.gov.br.

de processos operacionais básicos, deve estar disponível a informação essencial para o bom desempenho das atividades de seus coordenados. Ao gestor mais estratégico, devem estar disponíveis indicadores que lhe permitam acompanhar os macroprocessos segundo o modelo lógico do programa.

Diagrama 2: Integração de informações de diferentes fontes na estruturação de um Sistema de Indicadores de Monitoramento

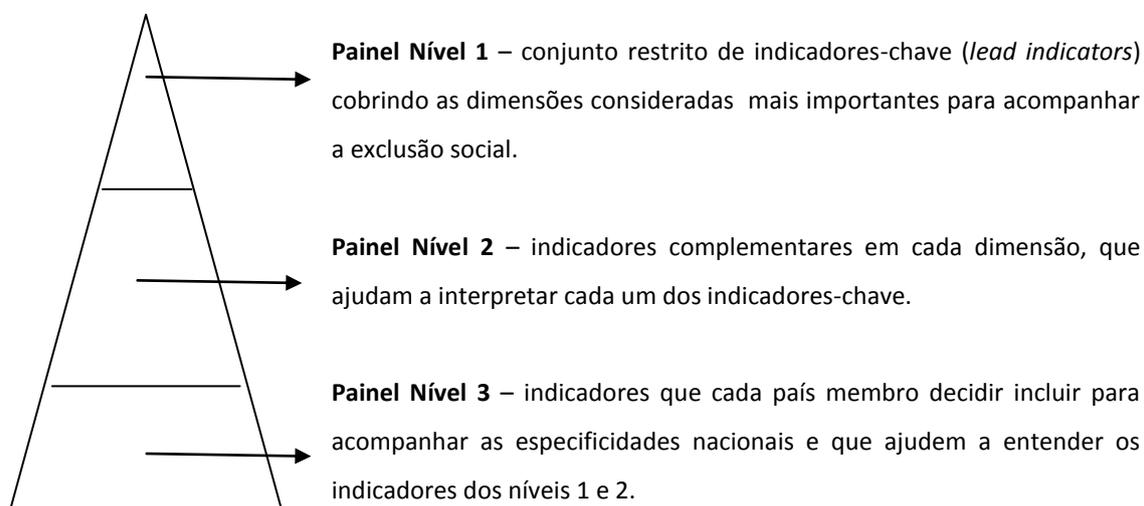


Para um e para outro gestor, os indicadores devem ser os pertinentes à sua esfera de decisão, ajustados à referência temporal e territorial que lhes compete e interessa. As novas ferramentas de integração de dados permitem construir painéis de indicadores de forma “customizada”, possibilitando, inclusive, acesso a informação mais detalhada se assim o gestor o desejar. Podem-se construir painéis em camadas “explicativas”, isto é, organizando indicadores segundo uma estrutura nodal, em que um primeiro conjunto reduzido de indicadores estratégicos seja acompanhado de um segundo conjunto mais amplo de indicadores mais específicos, que ajudem a entender o comportamento e a evolução dos primeiros, e assim por diante.

Na realidade, trata-se de um sistema de monitoramento que reúne informações sintéticas – para análise de tendências gerais das atividades estratégicas – e informações analíticas – para entendimento mais aprofundado das tendências

observadas. A proposta de acompanhamento das metas de inclusão social nos países da Comunidade Europeia segue essa lógica de estruturação, dispondo os indicadores em três painéis articulados (ATKINSON et al., 2005), como se vê no Diagrama 3:

Diagrama 3: Organização de Painéis de Indicadores em SM&A



Tal proposta de organização vem acompanhada de algumas premissas básicas para escolha de indicadores que parece oportuno resgatar neste texto, pois podem ser úteis em outras aplicações (Quadro 1). Vale registrar que tal escolha deveria se orientar também pela análise da aderência dos indicadores às propriedades de relevância social, validade de constructo, confiabilidade, periodicidade, sensibilidade às mudanças, especificidades das ações programadas, como discutido em Jannuzzi (2005).

Quadro 1: Premissas para escolha de indicadores de monitoramento da inclusão/exclusão social na Comunidade Europeia

- O conjunto de indicadores não pode se pretender exaustivo e deve ser equilibrado entre as dimensões da exclusão social (saúde, educação, moradia etc.). Um conjunto muito amplo de indicadores leva à perda de objetividade, perda de transparência e credibilidade.
- Os indicadores devem ter uma interpretação normativa claramente definida (Para monitorar a exclusão social a taxa de desemprego cumpre tal requisito; já um indicador de produtividade do trabalho não).
- Os indicadores devem ser mutuamente consistentes, isto é, não devem sugerir tendências inconsistentes (indicadores de desigualdade como o Índice de Gini e a Proporção de Massa Salarial Apropriada podem ter comportamentos diferentes ao longo do tempo, já que medem aspectos distributivos diferentes).
- Os indicadores devem ser inteligíveis e acessíveis a toda a sociedade. São preferíveis medidas simples, de fácil entendimento. Deve-se resistir às simplificações indevidas (indicadores sintéticos).

Se o programa foi especificado segundo as boas práticas e técnicas de planejamento de projetos, deve haver um desenho lógico de encadeamento de atividades e etapas. Tal sistema deve conseguir oferecer evidências acerca da execução do gasto, da produção, eficiência e qualidade dos serviços, do consumo e usufruto por parte do público-alvo e, se possível, antecipar dimensões impactadas pelo programa. Isto é, um bom sistema de monitoramento deve prover indicadores de insumo, processo, resultado e possíveis impactos do programa. Exemplificando, tal sistema deve permitir monitorar, simultaneamente: o dispêndio realizado por algum tipo de unidade operacional prestadora de serviços ou sub-projeto; o uso dos recursos humanos, financeiros e físicos; a geração de produtos e a percepção dos efeitos gerados pelos programas. Em Resende e Jannuzzi (2008), é apresentado o exemplo de Painel de Indicadores de Monitoramento do Plano de Desenvolvimento da Educação (PDE) estruturado na lógica insumo-processo-resultado-impacto. Na dimensão insumo, deu-se prioridade aos indicadores voltados para o financiamento da educação. Para a dimensão processo, foram definidos indicadores que pudessem ser produzidos com certa regularidade e estivessem relacionados ao processo de ensino-aprendizagem, como percentual de docentes com nível superior, percentual de alunos atendidos por turno integral, entre outros. Como medida de resultados, os indicadores de proficiência da Prova Brasil, defasagem idade-série e taxa de

abandono. Como apontamentos de impactos potenciais, o ingresso de jovens no ensino superior¹⁴.

A compilação de indicadores e de programas (Dici-VIP), organizada pela Secretaria de Avaliação e Gestão da Informação do Ministério de Desenvolvimento Social e Combate à Fome, é um produto útil para servir como referência à estruturação de sistemas de monitoramento de programas nessa perspectiva processual, não apenas no âmbito do ministério, mas de outros setores da administração pública¹⁵.

O documento “Guia metodológico para construção de indicadores do PPA” (MPO, 2009), elaborado pelo Ministério do Planejamento, Orçamento e Gestão, é outra referência importante para especificação de sistemas de indicadores de monitoramento de programas, pela extensa revisão bibliográfica empreendida na sua elaboração; pela preocupação em incorporar aspectos de propostas desenvolvidas em outros estados e contextos do setor público brasileiro; pela experiência de mais de 10 anos na elaboração de planos plurianuais no governo federal; e pela preocupação com clareza na exposição de conceitos, etapas e fontes de dados.

Em geral, há maior disponibilidade de indicadores de processos, que espelham os esforços e produtos gerados nas ações programadas, para os quais há registros formais e maior controle operacional pelos gestores. Ainda assim, em alguns casos, é possível dispor, no painel de monitoramento, de indicadores de resultados e impactos junto ao público-alvo dos programas, buscando informações nos registros e cadastros públicos citados anteriormente. No caso do Sistema de Indicadores de Monitoramento de Programas de Qualificação Profissional, ilustrado há pouco, seria possível dispor de alguma medida de impacto do programa, com a integração de informações cadastrais dos trabalhadores qualificados com os dados da Rais ou mesmo do CadÚnico, mediante o uso do Número de Identificação do

¹⁴ No referido texto, empregou-se o aplicativo Monit – disponível em www.anipes.org.br – para compor um painel situacional com gráficos dos indicadores escolhidos para duas unidades territoriais de interesse de monitoramento. Uma breve apresentação do aplicativo é encontrada em Jannuzi e Miranda (2008), disponível no Boletim de Estatísticas Públicas n.4, no mesmo site.

¹⁵ Consulte o “Dicionário de variáveis, indicadores e programas sociais” no endereço <http://aplicacoes.mds.gov.br/sagi/dicivip>. Esse dicionário também está disponível para *download* em www.mds.gov.br/sagi.

Trabalhador ou do CPF (de fato, procedimentos análogos são usados correntemente pelo Ministério do Desenvolvimento Social para fins de fiscalização na concessão de benefícios do Programa Bolsa Família).

Com o avanço da informatização no setor público brasileiro, é possível também obter informações referidas a unidades de prestação de serviços muito específicas (em tese, escolas, hospitais, postos policiais etc.), com boa periodicidade de atualização (frequência escolar mensal, atendimentos ambulatoriais semanais etc.), o que permitiria a construção de indicadores de monitoramento relacionados ao contexto de vivência dos beneficiários de programas e ao tempo adequado de tomada de decisão. De fato, projetos sociais de alcance local têm recorrido à busca periódica de informações nesses postos de prestação de serviços públicos como estratégia de monitorar resultados e impactos de sua ação. Centros de promoção de cursos de artesanato, atividades culturais e esportivas voltados à reintegração social de jovens em comunidades muito violentas, projetos que vieram a surgir com frequência nos últimos anos, pela ação direta de prefeituras ou organizações filantrópicas, podem ter seus resultados e impactos inferidos pelo eventual aumento das taxas de frequência à escola, diminuição dos atendimentos ambulatoriais decorrentes de ferimentos ou das ocorrências policiais envolvendo jovens, entre outras informações coletadas localmente.

Naturalmente, nos dois casos aqui exemplificados – programa de Qualificação Profissional e projeto social de reintegração social de jovens – os indicadores citados podem estar sendo afetados por outros fatores – conjuntura mais favorável do mercado de trabalho, no primeiro caso; policiamento mais ostensivo, no segundo – e não propriamente pela excelência do programa ou projeto. Não seriam, pois, exatamente indicadores de impacto, mas talvez indicações potenciais de impacto, que para efetiva atribuição ou vinculação causal com o programa ou projeto requereriam uma pesquisa de avaliação específica. Diferentemente dessas últimas, em que a investigação da atribuição de um efeito a um programa pode ser questão a avaliar para inferência a posteriori; em sistemas de monitoramento, assume-se a priori, acreditando-se na validade do desenho de implementação, a vinculação de ações (e de seus indicadores) com os resultados (na

forma como podem ser medidos pelas informações geradas pelo programa e outras fontes secundárias de baixo custo de aquisição). (MCDAVID; HAWTHORN, 2006).

Indicadores de eficiência, produtividade na prestação dos serviços, de cobertura de público-alvo, de qualidade de produtos ou satisfação dos beneficiários também deveriam constar da matriz de indicadores de monitoramento. Em programas de qualificação profissional, por exemplo, com base nos registros de operação deles próprios, é possível computar indicadores como custo de formação por aluno, número médio de alunos formados por entidade credenciada, taxa de frequência ou abandono de qualificandos.

Há cinco sistemas de macroacompanhamento de programas que podem ilustrar alguns dos aspectos estruturantes aqui apresentados: o Painel de Indicadores do Brasil Sem Miséria – MONIB-, o Siga Brasil, do Senado; a Matriz de Informação Social (MIS), da Secretaria de Avaliação e Gestão da Informação do Ministério do Desenvolvimento Social e Combate à Fome (Sagi/MDS); o SIM-Trabalho, da Secretaria de Emprego e Relações do Trabalho de São Paulo; e a Agenda 2012, da Prefeitura de São Paulo.

Talvez a experiência mais completa em termos das propostas aqui apresentadas seja aquela em estruturação na Secretaria de Avaliação do MDS, com o Painel de Indicadores do Brasil Sem Miséria – MONIB – e o Painel de Conjuntura, que organizam painéis de indicadores em diferentes níveis e temas, e diferentes formas de apresentação e funcionalidades.

As outras quatro relacionadas não são propriamente sistemas de indicadores de monitoramento para gestores de programas, mas sim ferramentas para consulta e acompanhamento de programas públicos por parte do cidadão, já que estão franqueados aos usuários da Web. O Siga Brasil¹⁶ permite acesso a dados e indicadores históricos sobre a execução orçamentária de programas do governo federal para estados e municípios. A MIS¹⁷ reúne um conjunto amplo de dados de contexto e indicadores de programas federais, que interessam a gestores e pesquisadores da área de assistência social. O SIM-Trabalho¹⁸ é um sistema voltado

¹⁶ Acesse por meio de um buscador na Web, procurando por Siga Brasil + Senado.

¹⁷ Acesse por meio do endereço <http://aplicacoes.mds.gov.br/sagi>.

¹⁸ Acesse por meio do sítio www.emprego.sp.gov.br.

a elaboração de diagnósticos para especificação de planos locais de qualificação profissional permitindo também acesso a informações sobre a execução de programas de intermediação de mão de obra e microcrédito. A Agenda 2012¹⁹ é um portal de comunicação da Prefeitura de São Paulo que apresenta a evolução das metas de gestão, nas várias secretarias, por distritos e subprefeituras, com atualização semestral.

Como já assinalado, um bom sistema de monitoramento deve permitir a análise comparativa dos indicadores ao longo do tempo e para distintas unidades de prestação dos serviços ou atendimento do programa, assim como em relação a padrões normativos de referência. A análise da evolução dos indicadores, de seus avanços, retrocessos e permanências permite identificar falhas operativas do programa, ainda que, em geral, não possibilite a explicação dessas. Esta aí uma diferença básica entre os sistemas de monitoramento e as pesquisas de avaliação: os primeiros, ao assumir que existe uma lógica de encadeamento de atividades, delineiam comportamentos esperados aos indicadores; nas pesquisas de avaliação, alertados pelos desvios não esperados na evolução dos indicadores, buscam-se as explicações valendo-se dos métodos e técnicas da pesquisa social. (MCDAVID; HAWTHORN, 2006).

No exemplo apresentado anteriormente, se um gestor precisar entender porque a taxa de abandono de alunos em cursos de qualificação em dada localidade ou em uma instituição está aumentando, talvez ele consiga alguma interpretação em seu próprio sistema de monitoramento – se este for um sistema que dispõe de informações analíticas, além de sintéticas. Pode ser que haja indicadores de contexto que mostrem forte aumento das vagas no mercado de trabalho regional (daí a evasão). A extração de informações do sistema para um pacote estatístico ou aplicativo de mineração de dados pode permitir ao gestor refinar sua hipóteses, se ele dispuser de conhecimento ou consultoria especializada no uso das ferramentas²⁰.

¹⁹ Acesse por meio do sítio www.agenda2012.com.br. A Agenda 2012 foi criada para responder às exigências da Emenda 30 à Lei Orgânica do Município, aprovada pela Câmara Municipal, ao final de 2008. Essa iniciativa vai ter, certamente, repercussões em outras localidades do país.

²⁰ Um pacote estatístico gratuito de ampla difusão é o Epi-info, disponível em www.lampada.uerj.br/, onde se pode encontrar também material de consulta e treinamento. O Weka, disponível em www.cs.waikato.ac.nz/ml/weka, é um aplicativo gratuito para mineração de dados, isto é, dispõe de rotinas para análise descritiva e exploratória de dados.

Ou talvez ele não tenha qualquer pista adicional e busque uma explicação com o dirigente da instituição ou gestor responsável pela área. É o que Worthern et al. (2004, p.38) classificam como avaliação informal, que ocorre “sempre que uma pessoa opta por uma entre várias alternativas existentes, sem antes ter coletado evidência formal do mérito relativo dessas alternativas”. Embora não sejam pautadas em procedimentos sistemáticos, tais avaliações nem sempre “ocorrem no vácuo”.

A experiência, o instinto, a generalização e o raciocínio podem, todos eles, influenciar o resultado das avaliações informais, e qualquer desses fatores, ou todos eles, pode ser a base de bons julgamentos. (WORTHERN et al., 2004, p.38).

Se o seu conjunto de indicadores de monitoramento não lhe permite chegar a explicações satisfatórias; se suas estratégias de visitação ou contato informal não lhe agregam evidências convincentes; ou se o processo de evasão ganha dimensões preocupantes em algumas áreas e não em outras; talvez seja o momento de o gestor pensar em encomendar uma pesquisa de avaliação.

5. As pesquisas de avaliação no Ciclo de Políticas e Programas

O fato de se denominar avaliação a fase posterior à implementação de programas no Ciclo de Políticas e Programas cria uma distinção artificial entre as atividades de monitoramento, exercida 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, 2000). 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 informações mais analíticas sobre o funcionamento desse, levantadas nas pesquisas de avaliação.

Como bem conceitua documento do Tribunal de Contas da União:

O monitoramento e a avaliação dos programas de governo são ferramentas essenciais para a boa prática gerencial. A avaliação é um procedimento que deve ocorrer em todas as etapas permitindo ao gestor federal o acompanhamento das ações e sua revisão e redirecionamento quando necessário. Enquanto o monitoramento é uma atividade gerencial interna, que se realiza durante o período de execução e operação, a avaliação pode ser realizada antes ou durante a implementação, como ao concluir uma etapa ou projeto como um todo, ou mesmo algum tempo depois, devendo se preocupar com o impacto provocado pela intervenção pública em seus beneficiários. (TCU, 2006, p. 75).

As tipologias clássicas usadas para classificar as pesquisas de avaliação acabam também favorecendo alguma confusão, como por exemplo 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? O programa deve continuar, ser expandido ou descontinuado? (Carvalho, 2003).

Cohen e Franco (1994) as classificam, respectivamente, de avaliações formativas e avaliações somativas. Estas últimas seriam avaliações do tipo ex-post, distintas das avaliações ex-ante, que se destinam a estudos de necessidade e viabilidade de criação de um programa. Worthern et al. (2004) resgatam outra classificação clássica, diferenciando as pesquisas de avaliação segundo seus protagonistas. Tem, então, a avaliação interna, conduzida pela equipe do programa; a avaliação externa, realizada por equipe de consultores externos contratados; a avaliação mista, que prevê a interação dos dois grupos; e a avaliação participativa, construída juntamente com os beneficiários dos programas (que também contribuem para o planejamento).

Cada modalidade tem vantagens e limitações. Se a avaliação externa parece permitir um olhar externo não enviesado e comprometido com a continuidade do programa, conferindo maior isenção à análise (supondo que os consultores são absolutamente profissionais, não preocupados com as chances de contratação em uma a próxima demanda de avaliação); a avaliação interna, a mista e a participativa parecem permitir aportes mais circunstanciados na análise dos processos e resultados, além de facilitar a incorporação de melhorias na gestão do programa – objetivo último da avaliação. Avaliações internas podem ser rápidas e baratas, ao contrário das outras modalidades. Avaliações participativas podem viabilizar aprimoramentos mais efetivos, pelo envolvimento dos beneficiários no sucesso do programa.

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 no Diagrama 3. 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 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 impactos e resultados do programa (Impact Assessment). Ao final, tendo-se constatado que o programa é socialmente justificável, conta com um desenho de implementação adequado, com procedimentos operacionais bem estabelecidos e com resultados efetivamente comprovados, passar-se-ia para a Avaliação da Eficiência ou 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 de avaliação segundo uma lógica de abordagem que tenderia a garantir, a priori, resultados relevantes e úteis ao final 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.

When developing the questions around which the plan for an evaluation will resolve, therefore, it is best for the evaluator to start at the bottom of the evaluation hierarchy [Avaliação da necessidade do programa] and consider first what is known and needs to be known about the most fundamental issues. When the assumptions that can be safely made are identified and the questions that must be answered are defined, then it is appropriate to move to the next level of hierarchy.

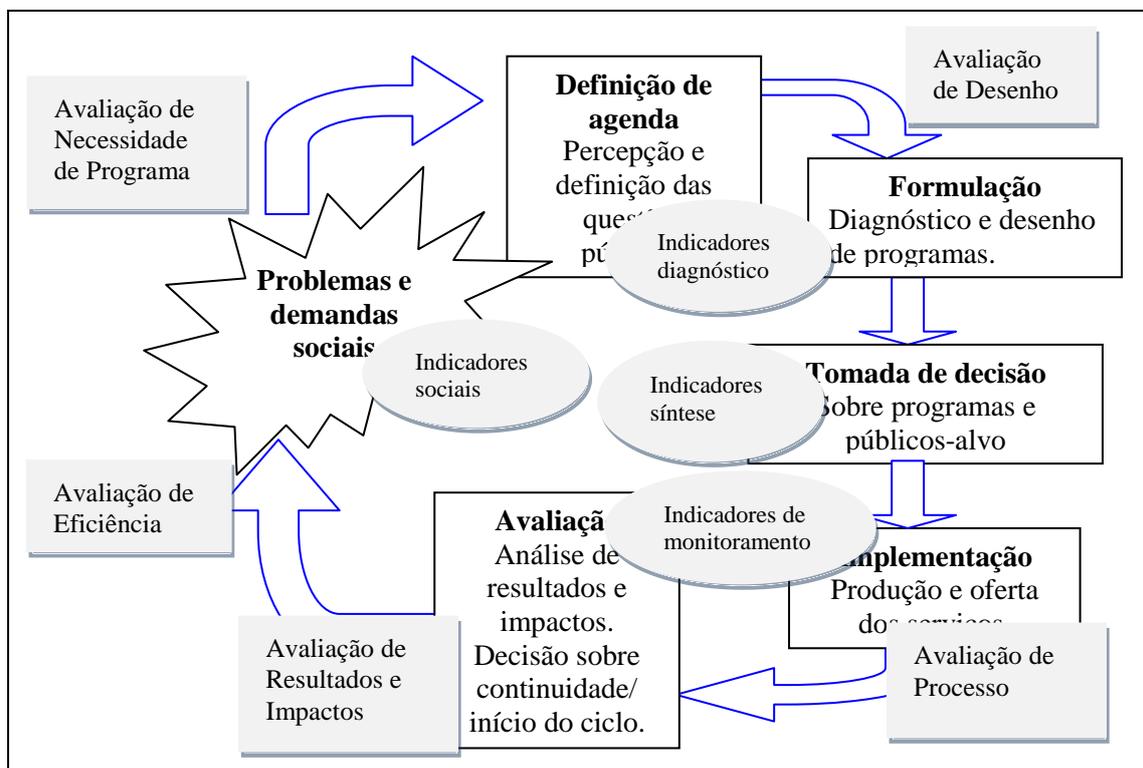
By keeping in mind the logical interdependencies between the levels in the evaluation hierarchy and the corresponding evaluation building blocks [técnicas e instrumentos apropriados da pesquisa social], the evaluator can focus the evaluation on the questions most appropriate to the program situation. (ROSSI et al., 2004, p.81).

Afinal, não caberia especificar uma pesquisa de avaliação de resultados e impactos se há dúvidas com relação ao estágio de maturidade dos processos e ações previstas na implementação do programa ou ainda se existem suspeitas de que o desenho de implementação 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 impactos gerados pelo programa são significativos. Enfim, evidências de baixa efetividade do programa poderiam ser consequências menos de falhas no desenho ou da forma como ele foi implementado e sim da própria inadequação da existência do

programa, 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. (2006) não implica que o gestor só possa dispor de informações acerca da eficiência, da qualidade dos serviços, 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 avaliação 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 avaliação em todas as etapas do “ciclo de vida” do programa, e que se resista à tentação de “colocar a carroça antes dos bois”, encomendando 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.

Diagrama 3: Indicadores e as Pesquisas de Avaliação no Ciclo de Políticas e Programas



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 análise econômica ou mesmo às análises institucionais típicos da Ciência Política. Análise de desenhos lógicos de implementação envolve revisão de documentos usados quando da formulação do programa, emprego das técnicas de planejamento de projetos e de comparação com outros modelos de intervenção social semelhantes em outros contextos e momentos.

A avaliação de processos pode requerer emprego de técnicas variadas de coleta de informações usados na pesquisa social. A avaliação de resultados e impactos, por sua vez, pode requerer emprego 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, problemas e diferentes aspectos da gestão de um programa público, ao longo do seu “ciclo de vida”, da sua concepção ao usufruto dos produtos e serviços por ele disponibilizado, considerando o contexto organizacional e político em que ele se insere, com a finalidade última de informar as necessidades de aprimoramento de suas ações, de modo a contribuir, juntamente 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 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 pela boa disponibilidade de manuais brasileiros de métodos e técnicas 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 Fome²¹ (MDS, 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 pesquisadores participantes, o período de realização, os objetivos da avaliação, os aspectos metodológicos acerca das técnicas de coleta de dados empregadas, as características da amostra e sujeitos entrevistados. As bases de dados dessas pesquisas de avaliação foram disponibilizadas no Portal da Secretaria e 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 similares²².

Esse material evidencia, na prática, como a natureza do programa, o estágio em que ele se encontra, os objetivos de avaliação, a disponibilidade de tempo e de recursos acabam orientando as decisões metodológicas acerca de instrumentos e

²¹ Vide www.mds.gov.br, menu “Institucional – Sagi”.

²² Vide www.nadd.prp.usp.br/cis.

técnicas de pesquisa, o tamanho, 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 através de entrevistas semiestruturadas com agentes públicos em diferentes posições da produção dos serviços (gestores municipais, técnicos atendentes do público), em localidades intencionalmente selecionadas (segundo tempo de implantação do programa, por exemplo).

Grupos de discussão com equipes técnicas envolvidas nos programas e, sobretudo, com beneficiários desses permitem levantar rapidamente informações cruciais para aprimoramento de procedimentos e ações dos programas e também para captar dimensões de impacto não facilmente tangíveis. Exemplo ilustrativo da técnica é descrito por Carvalho (2003), na estratégia para avaliar os resultados e impactos do programa Jovem Cidadão em São Paulo. Esse programa visava proporcionar, aos estudantes de 16 a 21 anos da rede pública estadual de ensino médio, a primeira oportunidade de experiência profissional no mercado de trabalho.

Por se tratar de um público composto por jovens, a técnica utilizada foi a de grupos de discussão, aplicando-se “dinâmicas que estimulam e facilitam a manifestação dos participantes, como simulações e jogos, tornando a pesquisa quase lúdica. Tais dinâmicas ajudam a atenuar dificuldades como inibição, timidez, desconfiança de determinados públicos. Isso é particularmente útil no caso de jovens, segmento que geralmente se caracteriza por falas lacônicas, em código, em especial diante de adultos.”[...] CARVALHO, 2003, p.189)

A análise estruturada de textos é uma técnica usada para interpretar relatos de entrevistas ou de grupos de discussão de forma mais estruturada, tratando palavras, frases, parágrafos como entidades empíricas, permitindo contabilizar frequências, recorrências e relacionamentos entre ideias de um conjunto de textos e documentos. No campo dos estudos de avaliação de programas, em especial nas pesquisas de satisfação ou avaliação de impacto junto a beneficiários de programas sociais, a técnica pode ser utilizada como recurso analítico para sistematizar, de forma mais objetiva e padronizada, as manifestações, opiniões e críticas presentes nos discursos dos entrevistados, e como estratégia metodológica para garantir maior replicabilidade das avaliações em outros contextos territoriais e temporais. Worthen

et ál. (2004) consideram que estas técnicas podem produzir avanços significativos nos enfoques qualitativos de avaliação.

Não existe a priori um método universal, mais legítimo ou com maior “status científico” para toda e qualquer pesquisa de avaliação, como não existe um único método para as pesquisas acadêmicas. Como bem assinalado em um importante manual da Pesquisa Social Americana acerca da prática da pesquisa científica:

It seems to us futile to argue whether or not a certain design is “scientific”[...] It is not a case of scientific or not scientific, but rather one of good or less good design [...]

The proof of the hypotheses is never definitive[...]

There is no such thing as a single “correct” design. Different workers will come up with different designs favoring their own methodological and theoretical predispositions[...]

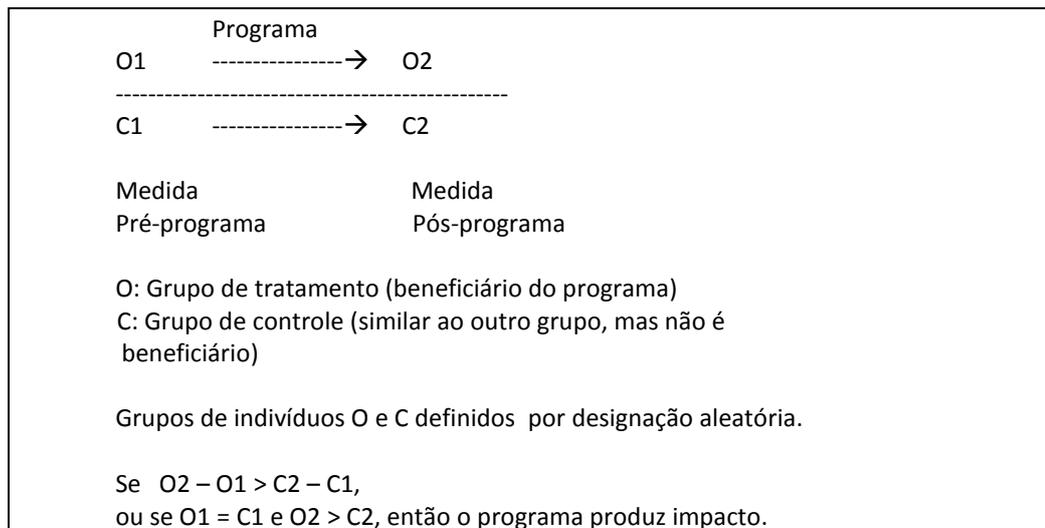
All research design represents a compromise dictated by the many practical considerations that go into social research. None of us operates except on limited time, money, and personnel budgets[...] A research design must be practical.

A research design is not a highly specific plan to be followed without deviation, but rather a series of guideposts to keep one headed in the right direction[...] (SUCHMAN in MILLER, 1991, p.98)

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 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 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. Como ilustrado no Quadro 3, o desenho visa avaliar a evolução dos dois grupos – supostamente idênticos no começo do experimento –, ao longo de tempo, e constatar se ao final há evidências de que o primeiro mostrou melhor performance que o segundo, em uma variável entendida como reveladora dos efeitos do programa.

Quadro 3 Desenho experimental clássico



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 dimensão impactada guarda, pelo desenho lógico do programa, vinculação estreita com as ações e atividades desse? 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 uma determinada variável, não forem elevados – algo que o incrementalismo de Lindblom sugeriria –, as amostras de beneficiários atendidos e do grupo 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, onde 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 naquele 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 –, seja na pesquisa acadêmica, seja na pesquisa de avaliação de programas, a incorporação de avaliadores provenientes das várias disciplinas das Ciências Sociais – antropólogos, sociólogos, comunicólogos etc. – e a formalização mais rigorosa de abordagens de investigação mais qualitativas, mais adequadas aos problemas complexos e pouco estruturados da realidade social, acabaram por consolidar a percepção, na comunidade de avaliadores nos EUA, de que os estudos avaliativos requerem certo ecletismo metodológico, integrando métodos quantitativos e qualitativos.

Na verdade, a maioria dos avaliadores sérios começou a entender que as abordagens quantitativas e naturalistas têm padrões metodológicos e rigor

que são diferentes, e não ausentes. A maioria dos avaliadores passou a aceitar a existência das múltiplas realidades ou pelo menos das múltiplas percepções da realidade. Com essa consciência e uma legitimidade maior, a avaliação qualitativa surgiu como alternativa real – ou complemento – da abordagem quantitativa tradicional. (WORTHERN et al., 2004, p.117)

Ainda assim, continuam os autores, alertando que o debate não está encerrado, na medida em que:

Aqueles que preferem o uso exclusivo ou majoritário de métodos quantitativos estão, em sua maior parte, aborrecidos com a aceitação dos estudos qualitativos (apesar do fato de o trabalho quantitativo ainda manter sua posição como abordagem dominante da avaliação e pesquisa [...]) Esses críticos da avaliação qualitativa queixam-se com frequência da subjetividade de muitos dos métodos e técnicas qualitativos, mostrando a preocupação de que a avaliação tenha abandonado a objetividade em favor de uma subjetividade exercida inabilmente. WORTHERN et al., 2004, p.117)

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.

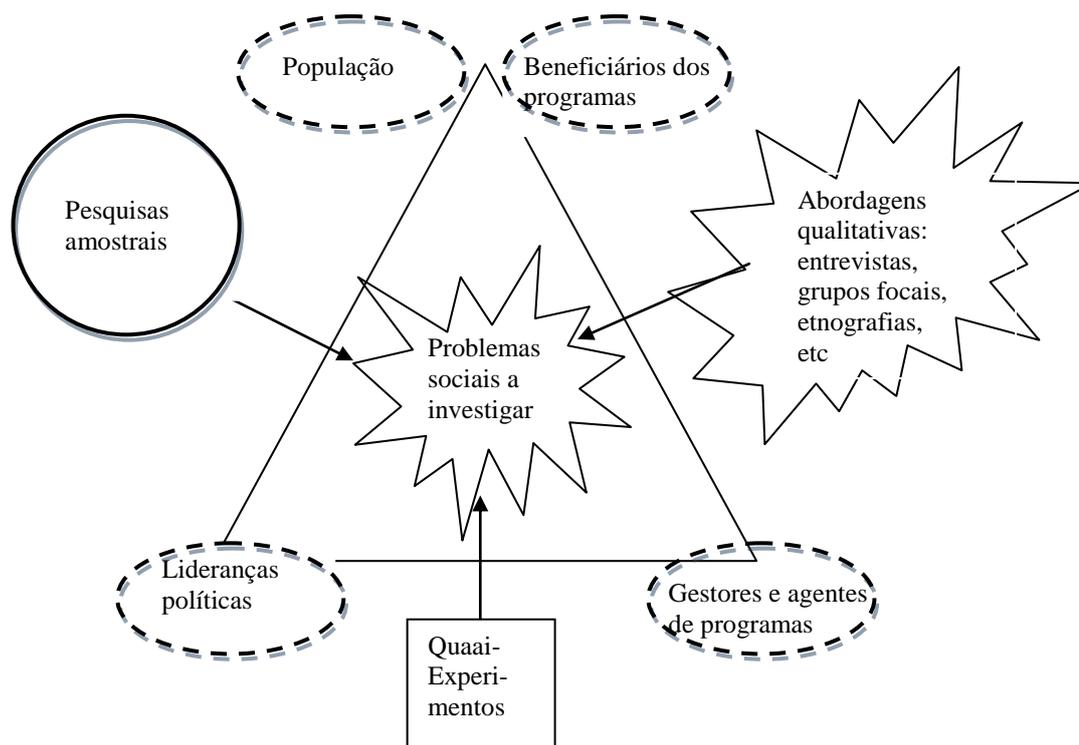
Desta forma, é a estratégia avaliativa que deve se ajustar à complexidade do programa ou problema a ser investigado, e não o contrário. É o problema a ser avaliado que deve orientar a escolha dos métodos, sujeitos, instituições a serem investigadas. Não é a suposta consistência técnica dos métodos de avaliação que deve condicionar a forma de como o programa deve ser implementado ou conduzido. O desenho ou estratégia de implementação do programa deve ser aquele que melhor se ajusta aos condicionantes, restrições e potencialidades existentes no contexto de operação do programa. Problemas sociais complexos requerem estratégias avaliativas igualmente complexas, exigindo muitas vezes a triangulação

de metodologias e sujeitos, ou melhor, a abordagem do objeto de estudo segundo técnicas diferentes de investigação e de sujeitos a serem entrevistados (Diagrama 5).

Pluralismo metodológico, enfoques avaliativos mistos, triangulação de abordagens investigativas²³, complementariedade de técnicas. São essas as perspectivas da pesquisa de avaliação de programas partilhadas pelos autores das duas principais referências bibliográficas aqui citadas – Worthern et al. (2004) e 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.

²³ 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.

Diagrama 5: “Problemas complexos, avaliações complexas” por meio de triangulação de técnicas e sujeitos



Deve-se observar ainda que na estratégia de avaliação não se requer necessariamente a obtenção de dados através de alguma técnica anteriormente relacionada. Os dados disponíveis nos registros operacionais dos programas, o sistema de indicadores de monitoramento, as pesquisas dos Institutos Nacionais de Estatística e dados de outros órgãos podem permitir realizar análises preliminares a baixo custo e tempo. O problema é que, em geral, essas fontes de informação não garantem a especificidade necessária para responder a perguntas mais concretas de interesse. Mas em várias situações é possível integrá-las em nível individual ou segundo algum agregado regional e analisá-las por meio de um modelo econométrico com certo poder avaliativo.

Outras fontes úteis de informação secundária são os relatórios de avaliação de órgãos de Controle Público, que no Brasil são o Tribunal de Contas da União (TCU) e a Controladoria Geral da União (CGU)²⁴. Os relatórios de avaliação de programas do TCU, elaborados desde 1998, são organizados em grandes tópicos, iniciando-se com a exposição dos objetivos da avaliação do programa (seção “O que foi

²⁴ Os relatórios destas instituições podem ser acessados em www.tcu.gov.br e www.cgu.gov.br.

avaliado”), a justificativa para a realização (“Por que foi avaliado”), os aspectos metodológicos da pesquisa de avaliação, explicando métodos e técnicas de coleta e análise de dados, as amostras e 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 investigação. Ao final, são sistematizados as sugestões de aperfeiçoamento para o programa (“O que pode ser feito para melhorar o programa”).

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

Os documentos, informes e relatórios de órgãos internacionais das Nações Unidas - – PNUD, Unicef, Unesco, Habitat, FAO, OIT, Cepal, etc. – assim como Banco Mundial, OCDE, Banco Interamericano de Desenvolvimento são outras fontes de informação para avaliação de Políticas, programas e ações governamentais em escala nacional e internacional. Estes relatórios, usualmente anuais, também trazem dados estatísticos, análises temáticas e experiências de diferentes países. Se pode fazer comparações de indicadores sociais, considerando os gastos orçamentários e a dinâmica econômica de cada país. Se pode conhecer experiências e programas de outros países, as melhores práticas, os desafios e dificuldades comuns. Há que se considerar, como em toda avaliação ou estudo, que a visão de mundo e os enfoques metodológicos dos autores e da instituição executora determinam os resultados que são registrados nos relatórios.

Quadro 3: Alguns dos principais relatórios sociais internacionais

Instituição	Informe Social
CEPAL	Panorama Social de América Latina y Caribe
PNUD	Relatório de Desenvolvimento Humano
UNESCO	Global Education Digest
UNICEF	Estado Mundial da Infancia
FAO	Estado mundial da agricultura e alimentação
OIT	Tendencias mundiales del empleo
Banco Mundial	Relatório de Desenvolvimento Mundial
OCDE	OECD Factbook
Bco Interamericano	Development Effectiveness Overview

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 sob 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 desse. Não existe uma receita única, pronta e acabada para desenhar esses instrumentos. Existem experiências, recomendações e boas práticas. Que elas sirvam para enfrentar os desafios da estruturação dos nossos sistemas de M&A.

Referências Bibliográficas

ALMEIDA, A. C. *Como são feitas as pesquisas eleitorais e de opinião*. Rio de Janeiro: FGV, 2002.

- ATKINSON, T. *et al. Social Indicators: the EU and Social Inclusion*. Oxford: Oxford Univ. Press, 2005.
- BABBIE, E. *Metodos de pesquisa de Survey*. Belo Horizonte: Ed.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. *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, Ernesto; FRANCO, Rolando. *Avaliação de projetos sociais*. Petrópolis, RJ: 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.
E contribuição para o debate. *São Paulo em Perspectiva*, São Paulo, v. 17, n. 3-4, p. 185-197, 2003.
- COUTINHO, R. G. Subsídios para organizar avaliações da Ação Governamental. *Textos para discussão*, Brasília, n. 776: IPEA, 2001.
- FREY, Klaus. Políticas Públicas: Um debate conceitual e reflexões referentes à prática da análise de políticas no Brasil. *Planejamento e Políticas Públicas*, 21. Brasília: IPEA, 2002.
- GENERAL ACCOUNTING OFFICE. *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. Est. 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. FISCHER, F. *et al. 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.

JANNUZZI, P. M., MIRANDA, W. L. Pradin e Monit: ferramentas para tomada de decisão no ciclo de políticas públicas. *Boletim de Estatísticas Públicas*, Salvador, n.4, p.62-73, 2008.

JANNUZZI, P. M., MIRANDA, W. L., SILVA, D. S. G. da. Análise Multicritério e Tomada de Decisão em Políticas Públicas: Aspectos Metodológicos, Aplicativo Operacional e Aplicações. *Revista Informática Pública*, Belo Horizonte, ano II, p. 69-87, 2009.

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

LINDBLOM, C. Todavía tratando de salir del paso. SARAVIA, E.; FERRAREZI, E.(orgs.) *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.

MCDAVID, J. HAWTHORN.L. ***Program Evaluation and Performance Measurement: An Introduction to Practice***. Thousand Oaks: Sage, 2006.

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

MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E COMBATE À FOME. *Metodologias e instrumentos de pesquisas de avaliação de programas do MDS: Bolsa Família, Assistência Social, Segurança Alimentar e Nutricional*. Rômulo Paes-Sousa; Jeni Vaitsman (orgs.). Brasília, DF: MDS; SAGI, 2007.

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO. *Guia metodológico para construção de indicadores para o PPA*. Brasília, 2009.

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) – UFRJ, Rio de Janeiro.

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

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

REZENDE, L. M. e JANNUZZI, P. M. Monitoramento e Avaliação do PDE: IDEB e painel de indicadores. *Revista do Serviço Público*, Brasília, v. 59, n. 2, p. 121-150, 2008.

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, Maria das Graças e CARVALHO, M. I. V. (orgs). *O estudo da política: tópicos selecionados*. Brasília: Paralelo 15, 1998.

RUA, M.G. *A avaliação no ciclo da gestão pública*. Apostila do curso de Especialização em políticas públicas da Educação com Ênfase em monitoramento e Avaliação – MPA. ENAP. MIMEO, 2004.

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

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

SCANDAR NETO, W. *et al*. Sistemas de indicadores ou indicadores sintéticos: do que precisam os gestores de programas sociais. In: HENRIQUES, R. *et al* (Org). *Educação na diversidade : como indicar as diferenças?* (Coleção Educação para Todos, Série Avaliação). Brasília: MEC/Unesco, n. 8, v. 25, p. 16-34. 2006.

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

VIEIRA, Fabiola Sulpino. 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. SARAVIA, E.; FERRAREZI, E.(orgs.) *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.