

Do Quê ao Pra Quê da Meta-Avaliação em Saúde

Zulmira Maria de Araujo Hartz
André-Pierre Contandriopoulos

Focalizaremos aqui os objetivos e os princípios éticos e profissionais que justificam a necessidade da meta-avaliação dos Estudos de Linha de Base do Projeto de Expansão e Consolidação do Saúde da Família – ELBs/Proesf (Brasil, 2004), particularmente importante nos países em que a política nacional de avaliação ainda está em estágio incipiente, como é o caso do Brasil. Depois, nos concentramos nas bases teóricas das diretrizes metodológicas do estudo, contextualizadas e orientadas para a tomada de decisão resultante da aprendizagem organizacional – uma visão de aconselhamento científico que foi priorizada em nossa atuação de consultores, fortalecendo a credibilidade e a institucionalização de práticas democráticas da avaliação em saúde.

DO 'QUÊ' E DO 'PORQUÊ'

Uma boa avaliação requer que seu próprio esforço seja avaliado. Muitas coisas podem sair, e freqüentemente saem, indevidamente na sua realização. Portanto é necessário checar problemas como vieses, erros técnicos, dificuldades administrativas e uso inadequado (...) Para facilitar a comunicação será usado o rótulo de Scriven, 'meta-avaliação', para se referir à avaliação da avaliação, e 'avaliação primária' será usada para aquelas que são objeto da meta-avaliação. (Stufflebeam, 1974: 1)

Se a demanda dos governos para a avaliação de políticas e/ou programas em certos países como Estados Unidos, Canadá e Inglaterra vem de várias décadas, a propensão a adotar uma orçamentação por programas e uma gestão orientada por resultados, como prática corrente – também no Brasil –, influencia mundialmente um crescimento significativo do interesse pela avaliação, ainda que apresentando uma diversidade de abordagens no seu controle de qualidade. Assim, tão sintomática quanto o interesse crescente direcionado à avaliação de programas é a formulação, pelas sociedades nacionais e governos, de dispositivos legais, diretrizes políticas, princípios, regras de deontologia ou parâmetros específicos que buscam claramente balizar a prática da avaliação. Para Richard Marceau (2005) parece tratar-se de uma resposta à demanda amplamente difundida da prática da avaliação, assegurando-a da essencial 'credibilidade' ao seu reconhecimento, noção esta que retomaremos posteriormente dada sua relevância no âmbito das políticas públicas (Mark, 2001; Monnier, 1992; Baron & Monnier, 2003; Contandriopoulos, 2006).

O principal objetivo de uma meta-avaliação é verificar até que ponto os estudos analisados respondem aos pré-requisitos de uma avaliação adequada (mérito ou qualidade) e à necessidade de informações dos interessados (valor ou utilidade), especificada no contrato da 'avaliação primária' ou do 'meta-avaliando' (Stufflebeam 1974, 2001; Coosky, 1999). Apesar de os 'padrões' avaliativos constituírem a base da meta-avaliação, incluímos aqui as noções de 'princípios, diretrizes ou normas', subentendidos nos 'parâmetros' da língua portuguesa, considerando que desempenham papéis coadjuvantes ou similares na institucionalização da avaliação em alguns países, sobretudo na comunidade europeia (Widmer, 2004).

Aqui, usaremos estas palavras em sentido intercambiável, prática observada em diversas publicações antigas e recentes (Stufflebeam, 1974; Hudson, Mayne & Thomlison, 1992; Picciotto, 2005; IEG/WB, 2007), mas lembramos que podem inspirar evocações ou significados distintos na literatura específica: 'princípio' como uma proposição-guia, geralmente um fundamento ético-profissional ou convicção que legitima determinadas ações; 'diretrizes' freqüentemente enunciadas como regra geral ou recomendações e 'padrão' como um termo preferido para 'normas obrigatórias', ou algo que serve como exemplo reconhecido, com base no qual ações são julgadas ou monitoradas, conformando mecanismos de certificação (Picciotto, 2005; Bustelo, 2006).

A gênese de padrões nacionais de avaliação, da mesma maneira que para a auditoria ou contabilidade, faz parte do contexto social da profissão. Historicamente, é no setor educacional dos Estados Unidos, nas déca-

das de 1970 e 1980, que têm origem os padrões de avaliação dos programas públicos (Joint Committee on Standards for Educational Evaluation – JCEE). Muitas associações profissionais americanas e canadenses aderem a esses parâmetros que, entre outros, inspiraram fortemente as sociedades de avaliação suíça, africana e alemã. Eles são igualmente respaldados pelos organismos internacionais voltados para o financiamento e avaliação dos programas de ajuda ao desenvolvimento (IEG/WB, 2007) em que se enquadra o Proesf. Quatro grandes princípios ou atributos norteiam seus enunciados: ‘utilidade, exequibilidade ou factibilidade, propriedade e precisão ou acurácia’. Ainda que o JCEE constitua referência primordial em matéria de parâmetros, a Associação Americana de Avaliação (AEA) ampliou seu escopo validando, com seus membros, os princípios e diretrizes para os avaliadores: 1) a investigação sistemática e rigorosa; 2) a competência e, entre outras, a competência cultural; 3) a integridade e a honestidade; 4) o respeito às pessoas e 5) a responsabilidade em relação ao interesse público e coletivo, além do interesse das partes em questão. Foi a primeira vez que se registrou esta última consideração como uma diretriz formal (Marceau, 2005).

A Sociedade Canadense de Avaliação (CES) elaborou, também, diretrizes em matéria de ética dos avaliadores, as quais, juntamente com os padrões de avaliação originais, foram incorporadas, ampliadas e operacionalizadas pelo Centro de Excelência em Avaliação do Secretariado do Tesouro Federal (similar ao nosso Tribunal de Contas da União), disponíveis em relatórios recentes de avaliação dos programas governamentais (CEE/TBS, 2004a, 2004b).

A Sociedade Francesa de Avaliação empreendeu um processo regular da revisão de padrões internacionais visando a construir o mais amplo consenso possível entre seus membros e, em 2003, aprovou sua Carta de Avaliação das Políticas e Programas Públicos, que consiste de um preâmbulo e seis princípios, ou diretrizes: pluralismo, independência, competência, respeito à integridade dos indivíduos, transparência e responsabilidade, cada um brevemente descrito no documento que pode ser consultado no seu *site* (SFE, 2003).

Em contraste com a Suíça, a Alemanha e a França, o Reino Unido tem uma tradição mais longa de avaliação e pertence ao grupo europeu de países que cedo atingiram um nível relativamente alto de institucionalização – inclusive com a criação da United Kingdom Evaluation Society (UKES), em 1994 –, mas somente em 2003 a UKES apresentou um ‘guia de boas práticas’ (incluindo auto-avaliação programática) cuja peculiaridade está na dis-

tingão de normas dirigidas aos avaliadores, comissionários e participantes, mas de cujos resultados ainda não se conhece a aplicação (Widmer, 2004).

Robert Picciotto (2005), ao se referir às semelhanças entre as normas nacionais de avaliação de sete países (Austrália/Nova Zelândia; Canadá, França, Alemanha, Suíça, Estados Unidos e Inglaterra), alerta para o fato de que, se por um lado todas se referem a uma dupla audiência (organizações que comissionam ou comanditam, preparam, conduzem e usam as avaliações, bem como indivíduos que a elas interessem ou que sejam por elas afetados), é igualmente patente a assimetria entre o elevado peso atribuído às responsabilidades dos avaliadores e o atribuído às obrigações que deveriam assumir comanditários ou gestores de programas, explicitadas apenas pela Inglaterra.

Para Rakesh Mohan e Kathleen Sullivan (2006), a escolha dos padrões depende de exigências profissionais e governamentais, da cultura organizacional ou do próprio avaliador (auditores tenderão a usar seus próprios padrões), mas todos têm em comum as seguintes características: envolvimento dos atores interessados; manutenção de independência e explicitação dos conflitos de interesse; credibilidade e suficiência das evidências apresentadas para dar suporte aos resultados, conclusões e recomendações; proteção de informações confidenciais; uso de processos que garantam qualidade; apresentação de relatório equilibrado e manutenção dos registros de demonstração do uso apropriado da metodologia.

Se no conjunto dos autores e países mencionados as distinções certamente resultam de processos culturais e participativos, a similaridade dos princípios acordados sugere que a área atingiu um alto grau de profissionalismo. A comparação de padrões na União Européia, por exemplo, mostra que as semelhanças são bem maiores do que as pequenas diferenças de orientação (em nenhum caso os documentos se mostram inconsistentes ou contraditórios), e incentivar o uso mais amplo dessas ferramentas na prática da avaliação será um ponto crucial para seu aperfeiçoamento e desenvolvimento futuro (Widmer, 2004).

No Brasil, a complexidade da avaliação da gestão descentralizada em saúde, orientada pela atenção básica, tem-se constituído em objeto privilegiado da literatura nacional em avaliação, explorando diversos recortes e abordagens (Vieira da Silva *et al.*, 2006), mas o mesmo avanço não se observa no âmbito da meta-avaliação (Hartz, 2006). A múltipla inserção institucional da avaliação dos ELBs – em se tratando de uma investigação comissionada a grupos de pesquisa de instituições acadêmicas, selecionados por edital internacional elaborado em comum acordo entre os financi-

adores internos e externos (Ministério da Saúde e Banco Mundial) – reforça o interesse em se assegurar a qualidade e utilização do produto apresentado com a decisão de se realizar uma meta-avaliação, ‘instituída’ pelo gestor nacional do Proesf.¹

DO ‘COMO’ E DO ‘PRA QUÊ’

Em uma interpretação do conceito de meta-avaliação orientada para a aprendizagem, ela deve ser parte de um diálogo aberto de julgamento e recomendações coletivo, pelas diversas partes envolvidas no processo (Uusikylä & Virtanen, 2000: 52).

A meta-avaliação pode ser definida ‘operacionalmente’ como um processo de descrição, julgamento e síntese de um estudo ou qualquer procedimento de avaliação, fundamentada em padrões e critérios propostos e validados pelas associações profissionais e outros organismos auditores. Ela consiste, portanto, na verificação da coerência teórica e prática da pesquisa com os padrões internacionais e/ou governamentais do controle de qualidade das avaliações de políticas públicas (Stufflebeam, 2001, 2000). Esta aparente simplicidade da definição mascara a complexidade dos estudos de meta-avaliação como ação reflexiva e de investigação de suas práticas acadêmicas e funcionais.

Todos os autores concordam que, como para a avaliação de qualquer intervenção, a meta-avaliação também deve ser introduzida preferencialmente com a pesquisa avaliativa ainda em curso (meta-avaliação formativa), de modo que possa contribuir para o aperfeiçoamento do próprio estudo, e não apenas como meta-avaliação somativa, ao final da investigação, objeto da análise, quando poderá agregar credibilidade, mas não maior validade e utilidade aos resultados apresentados. Ambos, avaliadores e clientes, beneficiam-se, assim, desta dupla apreciação, favorecendo ainda o controle de vieses.

Uma das questões a serem colocadas é que pouco se tem avançado em relação à preocupação de Daniel Stufflebeam (1974), quando afirmava que o ‘estado da arte’ da meta-avaliação é limitado e que a existência de padrões de qualidade não supre a deficiência de textos que tratem do equacionamento de problemas conceituais e de sua própria ‘mecânica’. O que se observa em alguns artigos publicados na seção “Evaluating Evaluations” do *American Journal of Evaluation*, criada em 1999, é um certo reducionismo: limitam-se à aplicação dos critérios do JCEE previamente comentados (Coosky, 1999; Grasso, 1999; Sanders, 1999; Scott-Little *et al.*, 2002). Whorthen, Sanders e Fitzpatrick também chamam a atenção para o fato de que, apesar da grande publicidade e da ampla aceitação desses padrões do Joint Committee, pouco cresceu o número de avaliações submetidas a um

¹ No sentido de *instituted evaluation*, noção emprestada a Demarteau (2002) indicando avaliação ‘socialmente organizada’ e envolvendo obrigatoriamente três componentes: metodologia e instrumentos bem definidos; pessoas socialmente autorizadas para sua condução e uso dos resultados formalizados.

exame mais minucioso, em relação ao período anterior à sua publicação, e até “uma inspeção informal revela que só uma pequena proporção dos estudos avaliatórios chega a ser avaliada, mesmo da maneira mais superficial” (Whorthen, Sanders & Fitzpatrick, 2000: 612).

Assim, é interessante notar mesmo um certo retrocesso, se compararmos tal quadro à pluralidade dos artigos divulgados em periódicos até 1994, quando foi publicada a última versão dos padrões do JCEE, até com uma expressiva presença dos estudos de uso da avaliação (Anexo 1). É fundamental destacar que Patricia Rogers, ao sistematizar a compilação bibliográfica original, observava que a meta-avaliação, como qualquer pesquisa avaliativa, pode privilegiar diferentes perspectivas ou paradigmas (técnica, gerencial, crítica) e a mesma variedade de abordagens da avaliação de programas. Na listagem em pauta, o artigo de Thomas D. Cook e Charles L. Gruder (1978) continua incontornável para os que fazem meta-avaliação somativa, e o de Alan Winberg (1991) é pioneiro na análise das unidades internas de avaliação, ou seja, a que tem como foco a ‘função da avaliação’. Este último é visto atualmente como um dos objetos prioritários da meta-avaliação, dados os avanços observados no processo de sua institucionalização e a necessidade de entendê-la como ferramenta democrática de governança (Uusikylä & Virtanen, 2000; Bustelo, 2006; Beywl, 2006).

Em uma perspectiva institucional, uma análise crítica de avaliações deve servir à aprendizagem política e organizacional referindo-se, respectivamente, aos mecanismos de disseminação de conhecimento nas diversas arenas da política e ao contínuo e reflexivo questionamento avaliativo nas organizações, entendido como a única maneira de aumentar a sua utilização. Para ser útil, a avaliação deve ter foco no contexto (i.e. *on the texts*), nos ‘textos’ produzidos pelos diversos atores e seus significados.

Os itens discutidos pelos avaliadores (o ‘avaliando’, sua eficácia, teoria do programa etc.) são apenas textos. O programa e sua lógica (...) técnicas de todos os tipos, as descobertas da avaliação (sejam tabelas numéricas ou conjuntos de recomendações) estão condicionados pela maneira como foram compreendidos e interpretados por cada ator (incluindo o avaliador), de acordo com suas restrições formais, morais e normativas. (Bezzi, 2006: 58)

A contextualização necessita ser particularmente valorizada quando se trata dos atores institucionais de entes federados diversos com demandas contraditórias da (meta)avaliação² do nacional ao local, ressaltando a necessidade de formas mais participativas de avaliação:

o governo central estará interessado em dados agregados para descobrir tendências e resultados globais, enquanto os municípios precisarão de

² A notação ‘(meta)avaliação’ reforça a idéia de meta-avaliação e meta-avaliando na mesma afirmação.

informações desagregadas na produção de impactos específicos. Isso tem duas implicações na relação entre avaliação e aprendizado (...) nem todos os atores entenderão determinada avaliação como igualmente relevante e utilizável. Avaliações que são mais válidas do ponto de vista de governos descentralizados podem ser vistas como sendo de pouco valor em nível central. (Van der Meer & Edelenbos, 2006: 209)

Para que os atores aprendam com a avaliação, a análise deve fazer sentido na perspectiva dos atores envolvidos. Interpretações e interações de atores desempenham um papel não somente na produção de resultados, mas também na geração de procedimentos de aprendizado como consequência da avaliação. Estes pressupostos tornam imperativa a indicação de abordagens orientadas por um modelo pluralista (Monnier, 1992), inserida na interface da quarta e quinta gerações de avaliação, como pode ser visualizado no Quadro 1, adaptado de Gaëlle Baron e Eric Monnier (2003). Os autores apóiam-se em uma longa experiência da avaliação de políticas públicas em diversos países, fundamentando seus argumentos de que a credibilidade dos estudos avaliativos e suas chances de utilização decorrem dessa 'co-produção' dos participantes, a qual permite que assimilem todo o processo analítico e seus resultados. A credibilidade de uma avaliação depende, portanto, da capacidade de produzir informação coerente para todos os sujeitos envolvidos na intervenção (Contandriopoulos, 2006). É necessário que as instâncias de decisão, às quais a avaliação se dirige, possam participar e utilizar as informações geradas nos seus modelos de ação, o que justifica a indicação de uma abordagem contextualizada e co-produzida. Como salienta Cecília Minayo, avaliados e avaliadores são "como aprendizes que, estudando e refinando conceitos e técnicas, crescem junto com o processo de análise. Nessa concepção, a avaliação é muito menos classificatória e mais promocional" (Minayo, Assis & Souza, 2005: 46).

Quadro 1 – Tipologia e níveis de participação em (meta)avaliação

Geração	Tipo de avaliação	Decisão do conteúdo	Pilotagem	Julgamento e recomendações
1ª a 3ª	Gerencial	Administração responsável pela gestão do programa		Equipe de avaliação
4ª (a)	Co-produzida	Autoridades públicas	Gestonários + Cooperadores (Equipe de avaliação)	
4ª (b)	Pluralista	Autoridades públicas	Autoridades + Sociedade civil + Operadores – Avaliadores	
5ª	Emancipadora	Sociedade civil ³		

Fonte: Baron & Monnier, 2003.

³ A quinta geração combina as duas anteriores, mas ela implica a vontade explícita de aumentar o poder dos participantes graças ao processo de avaliação; esta abordagem, como as demais, compromete-se com a melhoria das políticas públicas, mas também a ajudar os grupos sociais a ela relacionados a melhor compreender os próprios problemas e as possibilidades de modificá-los, incorporando multimeios nessa 'emancipação'.

Essas mesmas condições aplicam-se na meta-avaliação dos ELBs/Proesf, cooperada pelos grupos de pesquisa, configurando um programa nacional multicêntrico e multi-institucional de avaliação da Atenção Básica à Saúde. Portanto, é preciso entender o papel da meta-avaliação além da mera aplicação dos padrões e critérios, compartilhando o aprendizado individual e coletivo ao se privilegiar estratégias coerentes com esta intencionalidade.

Considerando-se a inexistência de normas nacionais (associativas ou governamentais), os princípios e referências previamente mencionados – utilidade, exeqüibilidade, propriedade, precisão ou acurácia – balizam, sem engessar, os padrões de qualidade desta meta-avaliação. Seu 'caráter formativo' é incontornável, dada a incompletude própria aos estudos de 'Linha de Base', que presume necessários ajustes na avaliação somativa posterior, e o interesse formal em aprofundar e qualificar o trabalho realizado para maior e melhor utilização pelas organizações envolvidas por meio dos estudos de casos exemplares das diversas instituições de pesquisa envolvidas.

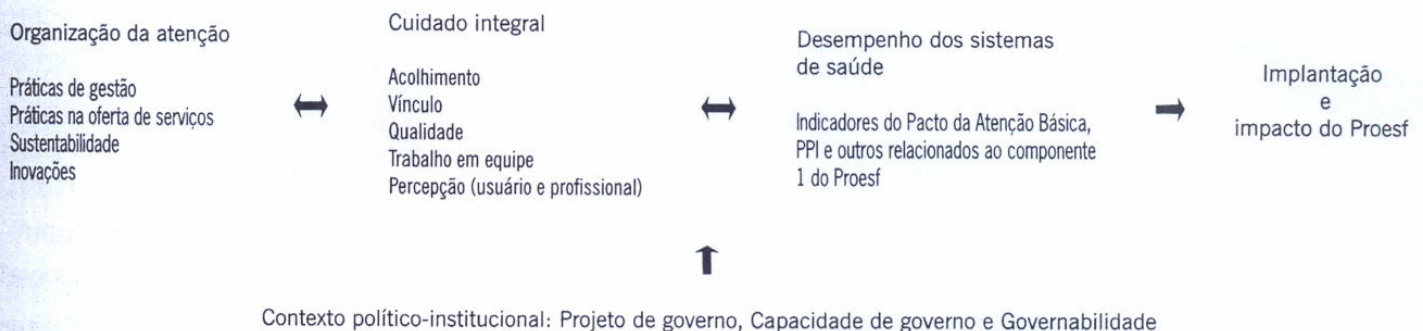
Se os padrões de avaliação originais alcançaram um elevado consenso no conjunto das Américas para os diferentes setores de políticas sociais, também no campo da saúde e das intervenções comunitárias (CDC, 1999; Hartz, 2006) eles não devem ser considerados um livro de receitas no seu enunciado ou aplicação, cabendo negociações e escolhas ao serem adaptados por cada 'meta-avaliador' (Whorthen, Sanders & Fitzpatrick, 2000). Em uma perspectiva internacional, a noção da necessidade de 'padrões abertos' está sendo desenvolvida em face de dificuldades inerentes à transferência de parâmetros entre diferentes culturas e contextos (Love & Russon, 2004), mas também à complexidade das políticas e programas avaliados.

Conseqüentemente, a inclusão de padrões relacionados ao adequado tratamento da 'especificidade das intervenções' nas pesquisas avaliativas (*specificity standards*), como proposto no estudo exploratório de meta-avaliação das intervenções comunitárias de promoção da saúde nas Américas (Hartz *et al.*, 2007), parece uma pista promissora a ser explorada, também, por outras outras meta-avaliações setoriais. Essa abordagem inspirou-se em estudos de meta-avaliação de natureza formativa, propostos pelas sociedades profissionais interessadas em aumentar a qualidade e a credibilidade dos estudos de avaliação e que focalizam, em formato agregado e com anonimato assegurado, um conjunto de relatórios e/ou publicações de um determinado campo de ação (Stufflebeam, 2001). Primando pela coerência e consistência entre princípios/padrões profissionais e governamentais na avaliação das políticas públicas, julgou-se que ela deveria ser (re)qualificada, associando-se os critérios canadenses utilizados na

revisão de relatórios governamentais, com vista à melhoria da qualidade da avaliação das políticas públicas no governo federal (CEE/TBS, 2004).

Outro registro, levando-se em consideração os objetivos específicos dos ELB (Quadro 2), é a necessidade de se apreciar, por meio de estudos de casos exemplares, as respostas dadas a estas demandas. Para tal entendeu-se a sistematização e co-validação dos resultados apresentados, e reconstruídos na matriz multidimensional do modelo teórico (Figura 1), como uma passagem obrigatória de comparabilidade e aprofundamento das lições aprendidas sobre os níveis de qualificação da Linha de Base municipal. Este 'marco analítico', inserido no Termo de Referência (TdR) contratual, constitui claramente um avanço conceitual e um desafio para os pesquisadores e (meta)avaliadores por incorporar em seu modelo, como dimensões inter-relacionadas à implantação e ao impacto do programa, a organização e a integralidade do cuidado, o desempenho dos sistemas locais e o contexto político. A construção das matrizes de aprendizagem com os resultados observados, inovações conceituais e operacionais efetuadas pelas equipes da avaliação primária têm fundamental importância para iluminar com maior intensidade e nitidez as contribuições das avaliações realizadas e as recomendações de maior relevância para a gestão e outras avaliações futuras, voltadas para o impacto do Proesf.

Figura 1 – Marco analítico dos ELB



Fonte: Brasil. MS, 2004.

Quadro 2 – Objetivos específicos dos ELB/Proesf

- Desenvolver metodologias de análise sobre as práticas de saúde da família, integrando indicadores epidemiológicos, assistenciais de saúde, que considerem as especificidades regionais, sociais, institucionais e políticas do país, tendo em vista o aperfeiçoamento e a consolidação das práticas atuais de avaliação da Atenção Básica/SUS.
- Coletar e sistematizar dados primários e secundários para qualificar, na implantação do Proesf (Linha de Base), os processos implicados na produção dos efeitos da intervenção nas diferentes dimensões (político-institucional, organizacional da assistência, cuidado integral e desempenho dos sistemas de saúde) e do estado de saúde das populações-alvo.
- Compartilhar lições aprendidas com as metodologias utilizadas e os resultados obtidos durante a pesquisa, mediante publicações, seminários e outros meios de divulgação dirigidos a distintos públicos-alvo.
- Apoiar a utilização e institucionalização dos resultados da avaliação nos municípios selecionados, integrando ações envolvendo Ministério da Saúde, secretarias estaduais e secretarias municipais de Saúde, pólos de educação permanente em saúde com Instituições de Ensino Superior (IESs) no monitoramento do Proesf.
- Viabilizar a realização de um estudo de seguimento em amostra estratificada representativa dos municípios incluídos na Linha de Base com uma comparação *ex-post* de grupos experimentais e de controle, nas áreas com e sem PSF, considerando-se, pelo menos, os indicadores para análise do 'caso-controle'.
- Subsidiar o processo decisório de progressão dos municípios ('gatilhos') da fase 1 à fase 2 do Proesf.

Fonte: Brasil. MS, 2004.

CONCLUSÕES: DAS EXPECTATIVAS E LIMITES DA META-AVALIAÇÃO

Se tivermos um pouco de sorte, já convencemos o leitor de que o conceito de meta-avaliação é útil e há instrumentos apropriados para usar com esta finalidade (...) Embora a raridade de boas meta-avaliações seja compreensível, não é facilmente perdoável, pois permite que práticas avaliáveis de má qualidade passem despercebidas e, pior ainda, que sejam repetidas. (Whorthen, Sanders & Fitzpatrick, 2000: 612)

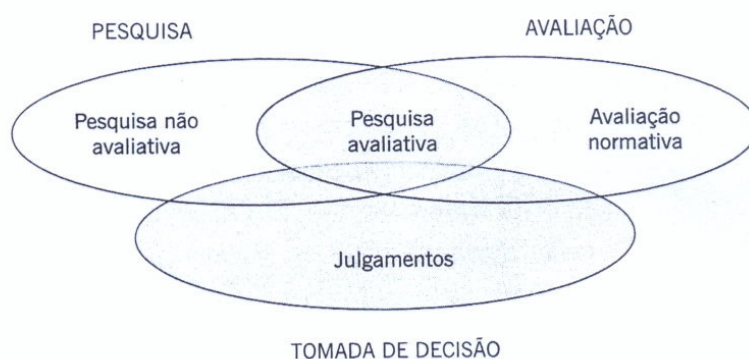
Os ELB/Proesf são tratados como estratégia fundamental para a institucionalização do monitoramento e da avaliação da atenção básica em âmbito nacional, enfatizando-se formalmente a articulação e a indispensável colaboração entre avaliadores e gestores, a academia e os serviços de saúde, no apoio à tomada de decisão, por meio da participação, difusão e uso das investigações. Este presumido consenso quanto à idéia de que avaliações sistemáticas melhoram as decisões, a fim de readaptar os conhecimentos em função do saber fornecido pela avaliação, traduz-se na necessidade de institucionalizar a avaliação em todos os níveis do sistema de saúde, de criar uma verdadeira cultura de avaliação. Mas não basta institucionalizar a avaliação, é preciso questionar continuamente a capacidade da avaliação em produzir as informações e julgamentos necessários para ajudar as instâncias decisórias a melhorar o desempenho de suas ações (Contandriopoulos, 2006).

Para que a avaliação desempenhe um papel nas decisões que estruturam o sistema de saúde, é necessário que um grande número de atores esteja convencido da importância dos resultados obtidos. Sem uma mobilização desta ordem, a avaliação não será verdadeiramente veículo de mudanças. Não basta concordar, em termos gerais, sobre a necessidade de introduzir alguma racionalidade nas decisões, justificando a demanda crescente de informação oriunda das avaliações. É preciso que as instâncias decisórias tenham uma forte convicção para que essa informação se traduza em mudanças desejáveis de suas práticas. A percepção de um fato e, por conseguinte, sua credibilidade dependem sempre da posição de quem faz o julgamento, e aumenta à medida que é aceita por um grande número de atores em diferentes posições ou 'objetividade transposicional' (Sen, 1994). Por um lado, os avaliadores se acostumam com os processos de elaboração de políticas e de decisão dos gestores e, por outro, é preciso que as instâncias de decisão compreendam os desafios e exigências da avaliação. Os gestores precisam ter uma idéia correta da consistência e da relevância dos resultados para decidir como aplicá-los às situações e aos problemas que os preocupam.

Postula-se que o grau com que os resultados de uma avaliação são levados em conta pelas instâncias decisórias será tanto maior quanto maior for a sua credibilidade, entendida como o valor científico potencialmente capaz de ser atribuído pelas instâncias de decisão à avaliação em si mesma, um dos principais objetivos da qualificação dos ELBs no processo de meta-avaliação. A fundamentação teórica e a credibilidade tecem relações, dentro de um determinado contexto, entre a avaliação, os diferentes sujeitos envolvidos (suas características, motivações, visão de mundo) e os problemas a serem solucionados.

A (meta)avaliação distingue-se da pesquisa por incluir um caráter normativo; ela objetiva não somente dar visibilidade aos efeitos de uma intervenção e entender como foram obtidos, mas, também, julgá-la (Figura 2). É nesta área que as instâncias de decisão, de posse dos resultados da avaliação, elaboram seus julgamentos, expressam seus valores e manifestam suas intenções estratégicas. Pode-se então conceber que os resultados de uma avaliação não se traduzam automaticamente em uma decisão, mas espera-se que informações co-produzidas nesta iniciativa, agora publicizadas, contribuam para o julgamento de uma determinada situação com maior validade, influenciando positivamente as decisões.

Figura 2 – Pesquisa, avaliação e tomada de decisão



Fonte: Contandriopoulos, 2006.

É preciso lembrar que a (meta)avaliação é um exercício permanente que não se esgota em julgamentos previamente estabelecidos, pois padrões de qualidade ou de verdade jamais são completamente apropriados; quando se pensa neles como 'garantia' logo se percebe a ingenuidade desta pretensão. Mas o fato de que julgamentos são provisórios e contingentes não nos obriga a cair no relativismo se nos orientarmos por princípios e normas co-validados por consensos construídos em debates anteriores e sempre abertos à sua crítica. Essa construção conjunta do conhecimento não deve ser vista como disputa de opiniões, e sim como uma solidariedade avaliativa explorando hipóteses e contradições que renovem os temas e conceitos tratados.

Qualquer atividade reflexiva desenvolvida em uma abordagem participativa permite aos membros do grupo (re)construir o julgamento e planejamento do 'avaliando' problematizado e (re)qualificado na sua construção. Juntos com os muitos cooperadores e co-produtores desta meta-avaliação, cujas estratégias de atuação anteriormente delineadas se detalham e se enredam nos capítulos deste livro (ver apresentação), compartilhamos o engajamento em aprimorar continuamente a qualidade e utilidade deste produto incompleto que se coloca agora como 'meta-avaliando', mas com a convicção de que estamos contribuindo para o desenvolvimento de uma cultura nacional de avaliação em saúde.

REFERÊNCIAS

- BARON, G. & MONNIER, E. Une approche pluraliste et participative: coproduire l'évaluation avec la société civile. *Informations Sociales*, 110, 2003.

- BEYWL, W. The role of evaluation in democracy: can it be strengthened by evaluation standards? *Journal of Multidisciplinary Evaluation*, 6: 10-29, 2006.
- BEZZI, C. Evaluation programmatic. *American Journal of Evaluation*, Special issue, 12(1): 56-76, 2006.
- BUSTELO, M. The potential role of standards and guidelines in the development of an evaluation culture in Spain. *American Journal of Evaluation*, 12(4): 437-453, 2006.
- BRASIL. Ministério da Saúde. Projeto de Expansão e Consolidação do Saúde da Família - Proesf (Componente III) Termo de Referência para o Sub-componente D Linha de Ação 3 - Estudos de Linha de Base: desenvolvimento de estudos de linha de base nos municípios selecionados para o componente 1. Brasília: Ministério da Saúde, 2004. Disponível em: <www.saude.gov.br>.
- CENTERS FOR DISEASES CONTROL (CDC). Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report*, 48 (RR-11): 1-40, 1999.
- CENTRE OF EXCELLENCE FOR EVALUATION/TREASURY BOARD OF CANADA SECRETARIAT (CEE/TBS). Report on effective evaluation practices, 2004a. Disponível em: <www.tbs-sct.gc.ca>.
- CENTRE OF EXCELLENCE FOR EVALUATION/TREASURY BOARD OF CANADA SECRETARIAT (CEE/TBS) Review of the quality of evaluation across departments and agencies, 2004b. Disponível em: <www.tbs-sct.gc.ca>.
- CONTANDRIOPOULOS, A. P. Avaliando a institucionalização da avaliação. *Ciência & Saúde Coletiva*, 11(3): 705-712, 2006.
- COOK, T.D. & GRUDER, C. L. Meta-evaluation research. *Evaluation Quarterly*, 2(1): 5-51, 1978.
- COOSKY, L. The meta-evaluand: the evaluation of project TEAMS. *American Journal of Evaluation*, 20(1): 123-136, 1999.
- DEMARTEAU, M. A theoretical framework and grid for analysis of programme-evaluation practices. *Evaluation*, 8(4): 454-473, 2002.
- GRASSO, P. G. Meta-evaluation of an evaluation of reader focused writing for the veteran benefits of administration. *American Journal of Evaluation*, 20: 355-370, 1999.
- HARTZ, Z. M. A Princípios e padrões em meta-avaliação: diretrizes para os programas de saúde. *Ciência & Saúde Coletiva*, 11(3): 733-738, 2006.
- HARTZ, Z. M. A. *et al.* Multi strategy in health promotion community interventions evaluation: an indicator of quality. (Mimeo.)
- HUDSON, J.; MAYNE, J. & THOMLISON, R. *Auditing Effectiveness Evaluations: action oriented evaluation in organizations*. Toronto: Wall & Emerson (WE), 1992.

- INDEPENDENT EVALUATION GROUP/WORLD BANK (IEG/WB). Sourcebook for evaluating global and regional partnership programs. Indicative principles and standards, 2007. Disponível em: <www.worldbank.org/ieg/grpp>- abril>.
- LOVE, A. & RUSSON, C. Evaluation standards in an international context. *New Directions for Evaluation*, 104: 5-14, 2004.
- MARCEAU, R. Chartes, principes, lignes directrices, éthique et standards: une réponse internationale diversifiée à des exigences de qualité des évaluations de programme. *Bulletin Société Québécoise d'Évaluation de Programme*, 18: 13-16, 2005.
- MARK, M. M. Evaluation's future: furor, futile or fertile. *American Journal of Evaluation*, 22(3): 457-479, 2001.
- MINAYO, M. C. S.; ASSIS, S. G. & SOUZA, E. R. (Orgs.) *Avaliação por Triangulação de Métodos: abordagem de programas sociais*. Rio de Janeiro: Editora Fiocruz, 2005.
- MOHAN, R. & SULLIVAN, K. Managing the politics on evaluation to achieve impact. *New Directions for Evaluation*, 112: 7-23, 2006.
- MONNIER, E. *Évaluation de l'Action des Pouvoirs Publics*. Paris: Editora Economica, 1992.
- PICCIOTTO, R. The value of evaluation standards: a comparative assessment. *Journal of Multidisciplinary Evaluation*, 3: 30-59, 2005.
- SANDERS, J. R. Metaevaluation of 'the effectiveness of comprehensive, case management interventions: evidence from the national evaluation of the comprehensive child development program'. *American Journal of Evaluation*, 20(3): 577-582, 1999.
- SCOTT-LITTLE, C.; HAMANN, M. S. & JURS, S. G. Evaluations of after-school programs: a meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation*, (4): 387-419, 2002.
- SEN, A. Objectivity and position: assessment of health and well-being. In: CHEN, L. C. Kleinman, A. & Ware, N. C. (Eds.) *Health and Social Change in International Perspective*. Boston: Harvard University Press; 1994.
- SOCIÉTÉ FRANÇAISE DE L'ÉVALUATION (SFE) Charte de l'Évaluation des politiques publiques et des programmes publics, 2003. Disponível em: <www.sfe.asso.fr>
- STAME, N. Governance, democracy and evaluation. *American Journal of Evaluation*, Special issue, 12(1): 7-16. 2006
- STUFFLEBEAM, D. L. *Metaevaluation*. Kalamazoo: Western Michigan University Center, 1974. (Occasional Paper, 3)
- STUFFLEBEAM, D. L. Evaluation Models. *New Directions for Evaluation*, 89 (spring): 7-98, 2000.

- STUFFLEBEAM, D. L. The metaevaluation imperative. *American Journal of Evaluation*, 2(2): 183-209, 2001.
- UUSIKILÄ, P. & VIRTANEN, P. Meta-evaluation as a toll for learning. A case study of the Structural Fund Evaluations in Finland. *Evaluation*, 6(1): 50-65, 2000.
- VAN DER MEER, F. B. & EDELENBOS, J. Evaluation in multi-actor policy processes: accountability, learning and co-operation. *American Journal of Evaluation*, 12(2): 201-218, 2006.
- VIEIRA DA SILVA, L. M. *et al.* Análise da implantação da gestão descentralizada em saúde: estudo comparado de cinco casos na Bahia-Brasil. *Cadernos de Saúde Pública*, 23(2): 355-370, 2006.
- WHORTHEN, B. R.; SANDERS, J. R. & FITZPATRICK, J. L. Meta avaliação. In: WORTHEN, B. R.; SANDERS, J. R. & FITZPATRICK, J. L. *Avaliação de Programas: concepções e práticas*. São Paulo: Edusp, 2000.
- WIDMER, T. The development and status os evaluation standards in Western Europe. In: RUSSON, C. (Ed.) International Perspective on evaluation standards. *New Directions for Evaluation*, 104: 31-54, 2004.
- WIDMER, T. The development and status os evaluation standards in Western Europe. In: RUSSON, C. (Ed.) International perspective on evaluation standards. *New Directions for Evaluation*, 104: 31-54, 2004.
- WINBERG, A. Maximizing the contribution of internal evaluation units. *Evaluation and Program Planning*, 14: 167-172, 1991.

Bibliografia Complementar

- COOKSY, L. J. & CARACELLI, V. J. Quality, context and use: issues in achieving the goals of metaevaluation. *American Journal of Evaluation*, 26(1): 31-42, 2005.
- COUSINS, B & AUBRY, T. Roles for government in evaluation quality assurance: discussion paper. In: CENTRE OF EXCELLENCE FOR EVALUATION/TREASURY BOARD OF CANADA (CEE/TBS), jun. 2006. Disponível em: <www.tbs-sct.gc.ca/aval>.
- CRACKNELL, B. E. A Neglected Area: the role of organizational dynamics and organizational learning. In: CRACKNELL, B. E. *Evaluating Development Aid: issues, problems and solutions*. London: Sage, 2000.
- CROZIER, M. & FREIDBERG, E. *L'Acteur et le Système*. Paris: Seuil, 1977.
- DAVIES, M. V. Teaching practical public health evaluations methods. *American Journal of Evaluation*, 27(2):247-256, 2006.

- DENIS, J. L.; LEHOUX, P. & CHAMPAGNE, F. A knowledge utilization perspective on fine-tuning dissemination and contextualization knowledge. In: LEMIEUX-CHARLES, L. & CHAMPAGNE, F. *Using Knowledge and Evidence in Health Care*. Toronto: University of Toronto Press, 2004.
- LEMIEUX-CHARLES, L. & CHAMPAGNE, F. *Using Knowledge and Evidence in Health Care*. Toronto: University of Toronto Press, 2004.
- MAYNE, J. Audit and evaluation in public health management: challenges, reforms and different roles. *Canadian Journal of Program Evaluation*, 21(1): 11-45, 2006.
- POLLIT, C. Performance information for democracy. The missing link? *American Journal of Evaluation*, Special Issue, 12(1): 38-55, 2006.
- SCRIVEN, D. Key evaluation checklist, 2005. Disponível em: <www.wmich.edu>. Acesso em: abr. 2006.
- UNITED KINGDOM EVALUATION SOCIETY. Guidelines for good practice in evaluation. Disponível em: <www.evaluation.org.uk/Pub_library/Good-Practice.htm>.
- YARBROUGH, D. B.; SHULHA, L. M. & CARUTHERS, F. Background and history of the Joint Committee's Program Evaluation Standards. *New Direction for Evaluation*, 104: 15-30, 2004.

ANEXO 1 – REFERÊNCIAS BIBLIOGRÁFICAS DE META-AVALIAÇÃO PUBLICADAS EM PERIÓDICOS INTERNACIONAIS, SELECIONADAS DA COMPILAÇÃO DE PATRICIA ROGERS PARA O PROJETO AREOL (ACTION RESEARCH AND EVALUATION ON LINE)⁴

Abordagens conceituais e operacionais

- BEDELL, J. R. *et al.* An empirical evaluation of a model of knowledge utilization. *Evaluation Review*, 9: 109-126, 1985.
- CARON, D. J. Knowledge required to perform the duties of an evaluator. *The Canadian Journal of Program Evaluation*, 8(1): 59-78, 1993.
- CHELIMSKY, E. The politics of program evaluation. In: CORDRAY, D. S.; BLOOM, H. S. & LIGHT, R. J. (Eds.) *Evaluation Practice in Review*. San Francisco: Jossey-Bass, 1987.
- COOK, T.D. & GRUDER, C. L. Meta-evaluation research. *Evaluation Quarterly*, 2(1): 5-51, 1978.
- GALLEGOS, A. Meta-evaluation of school evaluation models. *Studies in Educational Evaluation*, 20: 41-54, 1994.
- GREENE, J. C. A case study of evaluation auditing as metaevaluation. *Evaluation and Program Planning*, 15: 71-74, 1992.
- HOUSE, E. Multicultural evaluation in Canada and the United States. *Canadian Journal of Program Evaluation*, 7(1): 133-156, 1992.
- MCCORCLE, M. D. The operation was a success but the patient died: a critique of 'the implementation and evaluation of a problem-solving training program for adolescents'. *Evaluation and Program Planning*, 7: 193-198, 1984.
- MCTAGGART, R. When democratic evaluation doesn't seem democratic. *Evaluation Practice*, 12(1): 9-21, 1991.
- O'SULLIVAN, E.; BURLESON, G. W. & LAMB, W. E. Avoiding evaluation co-optation. *Evaluation and Program Planning*, 8: 255-259, 1985.
- PALUMBO, D. J. & NACHMIAS, D. The preconditions for successful evaluation: is there an ideal paradigm? *Policy Sciences*, 16: 67-79, 1983.
- SMITH, N. L. Toward the justification of claims in evaluation research. *Evaluation and Program Planning*, 10: 309-314, 1987.
- SMITH, N. L. & HAUER, W. J. The applicability of selected evaluation models to evolving investigative designs. *Studies in Educational Evaluation*, 16: 489-500, 1990.
- TORRES, R. T. Improving the quality of internal evaluation: the evaluator as consultant-mediator. *Evaluation and Program Planning*, 14: 189-198, 1991.

⁴ < www.scu.edu.au/schools/gcm/areol/areoides.htm >.

WINBERG, A. Maximizing the contribution of internal evaluation units. *Evaluation and Program Planning*, 14: 167-172, 1991.

Foco na decisão e uso da avaliação

BEYER, J. M. & TRICE, H. M. The utilization process: a conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*, 27: 591-622, 1982.

BRASKAMP, L. A.; BROWN, R. D. & NEWMAN, D. L. Studying evaluation utilization through simulations. *Evaluation Review*, 6: 114-126, 1982.

BURRY, J. *et al.* Organizing evaluations for use as a management tool. *Studies in Educational Evaluation*, 11(1): 131-157, 1985.

CONNOLLY, T. & PORTER, A. L. A user-focused model for the utilization of evaluation. *Evaluation and Program Planning*, 4: 131-140, 1981.

COUSINS, J. B. & LEITHWOOD, K. A. Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3): 331-364, 1986.

COUSINS, J. B. & LEITHWOOD, K. A. Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: creation, diffusion, utilization*, 14(3), 305-333, 1993.

DAWSON, J. A. & D'AMICO, J. J. Involving program staff in evaluation studies: a strategy for increasing use and enriching the data base. *Evaluation Review*, 9: 173-188, 1985.

HEGARTY, T. W. & SPORN, D. L. Effective engagement of decisionmakers in program evaluation. *Evaluation and Program Planning*, 11: 335-339, 1988.

HUBERMAN, M. Steps toward an integrated model of research utilization. *Knowledge: creation, diffusion, utilization*, 8(4): 586-611, 1987.

HUBERMAN, M. & COX, P. X. Evaluation utilization: building links between action and reflection. *Studies in Educational Evaluation*, 16: 157-179, 1987.

HUBERMAN, M. & LEVINSON, N. Knowledge transfer and the university: facilitators and barriers. *Review of Higher Education*, 8: 55-77, 1984.

LAWRENCE, J. E. S. & COOK, T. J. Designing useful evaluations: the stakeholder survey. *Evaluation and Program Planning*, 5(4): 327-336, 1982.

LEVITON, L. C. & HUGHES, E. F. X. Research on the utilization of evaluations: a review and synthesis. *Evaluation Review*, 5(4): 525-548, 1981.

MACKAY, K. The use of evaluation in the budget process. *Australian Journal of Public Administration*, 51(4): 436-439, 1992.

MERTENS, D. M. Practical evidence of the feasibility of the utilization-focused approach to evaluation. *Studies in Educational Evaluation*, 16: 181-194, 1990.

- NEWMAN, D. L.; BROWN, R. D. & RIVERS, L. S. Locus of control and evaluation use: does sense of control affect information needs and decision-making? *Studies in Educational Evaluation*, 9: 77-89, 1983.
- NEWMAN, D. L.; BROWN, R. D. & RIVERS, L. S. Factors influencing the decision-making process: an examination of the effect of contextual variables. *Studies in Educational Evaluation*, 13: 199-209, 1987.
- NEWMAN, D. L. *et al.* School boards' and administrators' use of evaluative information: influencing factors. *Evaluation Review*, 7: 110-125, 1983.
- RYAN, A. G. Program evaluation within the paradigms: mapping the territory. *Knowledge: creation, diffusion, utilization*, 10: 25-47, 1988.
- SHAPIRO, J. Z. Conceptualizing evaluation use: implications of alternative models of organizational decision making. In: Conner, R. F. and Associates (Eds.) *Evaluation Studies Review Annual*. Beverly Hills: Sage, 1984. v. 9
- SIEGEL, K. & TUCKEL, P. The utilization of evaluation research: a case analysis. *Evaluation Review*, 9: 307-328, 1985.