

# CAPÍTULO 11 – MAIS METODOLOGIA DO QUE MÉTODOS

LOUISE POTVIN E SHERRI BISSET

Uma avaliação é muito mais do que coletar, analisar e interpretar dados científicos a fim de comparar os resultados de tratamentos. Nos últimos 40 anos, as discussões sobre métodos no campo da avaliação de programas evoluíram no sentido de incluir modelos e reflexões acerca dos múltiplos e complexos papéis associados à prática da avaliação. Na verdade, para Shadish, Cook e Leviton (1991), a base de conhecimento que diz respeito à prática da avaliação deve considerar questões relacionadas aos papéis do avaliador, assim como ao formato da avaliação. Dessa forma, a prática da avaliação requer competências metodológicas e técnicas para condução de investigações sistemáticas, além de todo um conjunto de habilidades interpessoais e de negociação, identificadas por Brown (1995) como algo que perpassa os aspectos pedagógicos e políticos. A concepção de uma avaliação como uma prática, como fazemos nesta coletânea de ensaios, se baseia na premissa de que os avaliadores são mais do que bons e rigorosos cientistas, que implementam ferramentas de investigação empíricas para estudar os programas e intervenções, enquanto está claro que a função do avaliador envolve mais do que as atividades de coleta de dados, como esses dois aspectos da prática, que são parte inerente da avaliação de programas, se juntam para construção do conhecimento da prática da avaliação? Para abordar essa questão, nós propomos revisitar as fundamentações teleológicas, epistemológicas e ontológicas com base nas quais o papel das avaliações é definido.

Consistente com a orientação deste livro, nós definimos o papel dos avaliadores *vis-à-vis* um programa como sendo algo que é modelado pela prática de avaliação. Sugerimos, no entanto, que a prática da avaliação não consiste simplesmente de um repertório de papéis a partir do qual o avaliador pode (mais ou menos) escolher arbitrariamente a fim de definirem suas funções e suas atividades de avaliação (em outras palavras, o *kit* de ferramentas metodológicas do avaliador). Neste capítulo, nós argumentamos que o papel do avaliador *vis-à-vis* os *stakeholders* de um programa e a abordagem adotada para identificar, descrever e mensurar um programa e seus efeitos, formam a prática do avaliador, que é consistente e coerente no contexto do seu trabalho.

Também argumentaremos que uma prática, assim como um paradigma, representa uma racionalidade organizada (CROZIER e FRIEDBERG, 1977), comum a grupos de indivíduos, e que possibilita a esses grupos identificar, estruturar, interpretar e resolver problemas existentes e que são identificáveis por meio das lentes de uma prática ou paradigma (KUHN, 1962).

## As quatro dimensões da prática

A metodologia é uma das quatro dimensões que caracterizam uma prática (LEVY, 1994).

As práticas, como os paradigmas, possuem uma organização coerente, de modo que, quando tomadas conjuntamente, essas quatro dimensões são altamente relacionadas entre si e constituem um conjunto coerente e racional de proposições que conectam um profissional ao mundo por meio de sua prática. Essas quatro dimensões são rotuladas de: teleológica, ontológica, epistemológica e metodológica.

A dimensão teleológica diz respeito ao objetivo central da prática: o que o profissional coloca como meta para alcançar por meio do seu trabalho e das atividades. É essa dimensão que define a intencionalidade e se refere ao projeto e a visão mais abrangente que sustentam a prática. Ela oferece significado à ação ao identificar os possíveis mundos para os quais a prática esta contribuindo. Essa dimensão é muitas vezes tida por dada e, raramente, é discutida de forma aberta e revisada criticamente no campo da avaliação. Típico de um campo prático, no entanto, os resultados das avaliações são, muitas vezes, descritos como algo que precisam ser utilizáveis e contribuir para algum projeto de transformação (MARK e HENRY, 2004). A avaliação é, muitas vezes, definida como algo que fornece dados científicos para a tomada de decisões. Esse uso instrumental dos resultados das avaliações em decisões relativas ao programa representa o critério último pelo qual uma grande maioria dos avaliadores descreve a *raison d'être* do campo (PRESKILL e CARACELLI, 1997). Tal uso instrumental, no entanto, quase nunca é possível. Patton (1997) e muitos teóricos da avaliação propuseram outros tipos de uso para os resultados de uma avaliação (HARTZ, DENIS, MOREIRA e MATIDA, 2008), como o esclarecimento, ou seja, a contribuição dos resultados das avaliações para as explicações teóricas sobre o funcionamento do mundo (WEISS, 1998). Para Mark, Henry e Julnes (2000) todos esses usos específicos dos resultados das avaliações podem ser entendidos como contribuindo para a melhoria social à luz de que

[...] mesmo na ausência de um uso direto, os resultados das avaliações, muitas vezes, parecem ter ajudado a moldar os pressupostos das pessoas, suas crenças e expectativas, e, por sua vez, eles parecem ter influenciado decisões subsequentes sobre os programas e políticas, às vezes, distantes no tempo e em lugar da avaliação original (MARK *et al.*, 2000, p. 22).

Para esses autores, a fim de contribuir para a melhoria social, os avaliadores assumem as seguintes responsabilidades com relação aos resultados dos seus estudos. Primeiramente, os avaliadores determinam que resultados podem melhor contribuir para apoiar as deliberações, decisões e ações realizadas pelas instituições, uma vez que as instituições representam os agentes legítimos da regulação e da transformação social em sociedades democráticas. Em segundo lugar, eles assumem responsabilidade pela qualidade da informação extraída dos estudos de avaliação reais. Terceiro, eles garantem que os resultados sejam disseminados pelas redes mais relevantes de profissionais e de tomadores de decisão.

A segunda dimensão da prática tem relação com seu objeto. Considerando que a prática é o trabalho transformador que um profissional realiza no mundo que o envolve, esse trabalho coloca o foco sobre uma classe de objetos específica. Não há prática sem os objetos que uma prática busca regular reproduzir ou transformar. Essa é a dimensão ontológica da prática. Com a exceção de Weiss (1998), há pouca discussão sobre a realidade ontológica dos programas e políticas que, normalmente, são apresentados como objetos de avaliação (POTVIN, GENDRON e BILODEAU, 2006). Poucas vezes perguntou-se em que consistem os programas. Como discutido por POTVIN e MCQUEEN (2008), a maioria das discussões e representações dos programas não trazem a figura dos atores, tomados no sentido de agentes que exercitam um poder causal em uma dada situação; atores cujas ações induzem uma reação de outros atores. Nas raras ocasiões em que os atores são alvo de discussão juntamente com objetos materiais, como em Weiss (1998), a articulação e as conexões entre os elementos humanos e não humanos dos programas não são bem desenvolvidas. Representar programas como uma série de ações ou eventos na ausência de atores que operem es-

sas ações coloca o ônus da ação nas dimensões técnicas e materiais dos programas. Neste capítulo, nós argumentamos que as concepções dos avaliadores acerca da natureza dos programas e sobre o que constitui sua realidade exercem influência fundamentais sobre a prática das avaliações e, em última instância, definem as outras dimensões da prática.

Definir prática como o trabalho de um profissional sobre um objeto implica a presença de um *rap-port* entre o profissional e a classe dos objetos que caracteriza a prática. A dimensão epistemológica da prática tem a ver com esse relacionamento entre o profissional e o objeto de sua prática. Ela faz a pergunta sobre que tipo de *rapport* que deveria existir entre um profissional e o mundo a fim de que o primeiro influencie as transformações e a regulação do último. Para um avaliador, essa dimensão diz respeito ao relacionamento existente entre ele/ela e os programas ou políticas a seres avaliados. Muitas vezes reduzida a uma oposição entre avaliação externa e interna (por exemplo, pesquisa participativa *versus* não participativa, observação participante *versus* não participante), uma resposta adequada para essa pergunta requer que o avaliador, como sujeito, se posicione com relação ao programa. A prática da avaliação pode variar de um simples *rapport* sujeito-objeto, no qual o avaliador, como um sujeito, controla as circunstâncias do programa, idealmente concebido como destituído de poder no relacionamento (as decisões do avaliador não são influenciadas pelo programa), até um *rapport* muito mais complexo entre sujeito-objeto-projeto, no qual tanto o avaliador quanto o programa são agentes ativos na avaliação e no qual esse *rapport* também é percebido como sendo um ingrediente ativo na prática do avaliador. Nós mostramos que quando uma avaliação é pensada para ser responsiva às várias fases pelas quais um programa passa, tanto o programa quanto a avaliação influenciam um ao outro e se tornam cada vez mais parte do contexto no qual o programa inicialmente foi desenvolvido (POTVIN, CARGO, MCCOMBER, DELORMIER e MACAULAY, 2003).

Para a prática de avaliação baseada em pesquisa, essa dimensão epistemológica (que, em última instância, coloca a questão do *status* do avaliador em relação ao programa) se preocupa com como o conhecimento sobre os programas é possível dada sua natureza.

Por fim, a dimensão metodológica de uma prática se refere diretamente ao tipo de ação que os profissionais conduzem para alcançar seus objetivos. Para um avaliador, a dimensão metodológica da prática está primeiramente, mas não exclusivamente, relacionada a como o conhecimento sobre programas é produzido, mas ela não pode ser reduzida apenas a esse aspecto.

De fato, na medida em que a finalidade visada por uma avaliação se expande além da mera produção de conhecimento científico, a dimensão metodológica da prática de avaliação pode incluir uma ampla gama de ações que dizem respeito principalmente à análise da possibilidade de avaliação, que define a pergunta e o projeto de avaliação (THURSTON, GRAHAM e HATFIELD, 2003; THURSTON e POTVIN, 2003), e ao uso do conhecimento (MARK e HENRY, 2004; PATTTON, 1997).

## Três tipos ideais de práticas de avaliação

Uma prática é a operacionalização de um conjunto coerente de posições acerca das quatro dimensões descritas acima. Essa coerência é a fundamentação da prática e é vista como uma consistência e uma continuidade da prática ao longo do tempo e do espaço. Embora seja concebível que as várias combinações de respostas que os profissionais fornecem àquelas quatro perguntas

possam produzir uma infinidade de variedades de práticas, nós acreditamos que existem correlações inerentes entre essas dimensões e que elas, assim, limitam as variações nos tipos de prática de avaliação. A esse respeito, o conceito de Weber de tipo-ideal é útil (WEBER, 1952). Um tipo-ideal é uma construção mental de um fenômeno social que combina e simplifica características que, normalmente, não são encontradas em uma única instância, mas que fornece um caso exemplar representativo de toda uma categoria de fenômeno social. Existem muitas taxonomias de tipos-ideais de avaliação, dependendo da característica que se queira destacar. Nesta seção, nós apresentamos três tipos-ideais de práticas de avaliação que enfatizam as consequências metodológicas de três paradigmas de avaliação coerentes, conforme apresentado na Tabela 1.

Tabela 1 – Três tipos-ideal de prática de avaliação

	Teleológica	Ontológica	Epistemológica	Metodológica
Avaliação como Experimentação	Teste das hipóteses causais acerca dos efeitos que os programas exercem no contexto	Programas como entidades técnicas com uma realidade objetiva	Distanciamento objetivo entre avaliador e programas	Experimentação; excluir possíveis hipóteses rivais
Avaliação como Negociação	Melhora no programa por meio da conscientização dos atores a respeito de suas ações nos programas	Programas como representações que descrevem e/ou orientam as ações dos atores	Co-construção construtivista sujeito-sujeito da representação do programa	Construir teoria do programa por meio das representações dos atores e de círculos hermenêuticos
Avaliação como Reflexividade Organizada	Compreender o papel mediador dos programas nas transformações sociais	Programas como sistemas de ações que conectam entidades	Transformação reflexiva do sujeito-objeto-projeto	Acompanhar ações dos programas e as conexões que elas operam

Note-se que algumas práticas bem conhecidas de tipos-ideais de avaliação, tal como a avaliação com foco na utilização (PATTON, 1997) ou a avaliação participativa (SPRINGETT, 2001), não são diretamente circunscritas pela nossa tipologia, principalmente porque essas práticas não são discutidas prioritariamente em termos de suas implicações metodológicas. A primeira é geralmente apresentada com relação à dimensão teleológica da avaliação, ou seja, seus usos por parte dos tomadores de decisão, enquanto a última (a avaliação participativa) é mais apresentada em relação à questão epistemológica do valor relativo de várias formas de conhecimento. Nessa perspectiva, a avaliação participativa também pode ser compreendida como uma forma de negociação.

A principal advertência com relação ao uso dos tipos-ideais para caracterizar a prática de avaliação no campo da promoção da saúde é que, pelo fato deles não existirem de fato na realidade, os tipos-ideais são representações simplistas. Como uma ferramenta heurística, um tipo-ideal não consegue reproduzir todas as sutilezas e nuances da realidade social que ele representa. Ele é requisitado principalmente para contrastar o fenômeno social por meio da exacerbação de suas diferenças em um número limitado de dimensões. A prática é algo muito mais rico do que qualquer tipo-ideal que possa ser construído para representá-la. É por isso que este capítulo precisa ser complementado pelos capítulos relatam e discutem questões práticas da condução de avaliações que também contribuem para promover a agenda da promoção da saúde.

## Avaliação como experimentação

De acordo com Pawson e Tilley (1997), esse tipo-ideal de prática de avaliação tem sido a marca característica das avaliações desde o surgimento deste campo.

Na base de tudo no começo estava a lógica da experimentação [...] O profissional, o assessor de políticas e o cientista social estão em consenso quanto a apreciação da beleza do *design*. Em um aspecto, ela tem raízes profundas no discurso filosófico sobre a natureza da explicação, como em *A System of Logic*, de John Stuart Mill; por outro lado, essa é a marca característica do senso comum, entranhada nas campanhas publicitárias (PAWSON e TILLEY, 1997, p. 4).

A avaliação como uma prática de experimentação é caracterizada pela finalidade de conduzir testes sobre as hipóteses causais acerca dos efeitos de um programa no ambiente em que a iniciativa é realizada. Essa finalidade está profundamente enraizada no campo da avaliação. Já em 1967, em um dos primeiros livros publicados cujo título trazia a palavra “avaliação”, Suchman (1967) explicava que a avaliação é, em última instância, um mecanismo de teste de hipóteses por meio do qual demonstrar um efeito positivo alcançado por um programa voltado à solução de um problema é oferecer evidências acerca da validade da teoria que fundamenta o programa.

Na prática, como profissionais que testam hipóteses, os avaliadores, muitas vezes, são também pesquisadores acadêmicos, profundamente envolvidos na formatação e planejamento de programas. Os programas, como tratamentos, são vistos como ferramentas para testar proposições teóricas sobre a etiologia das doenças. Esse foi o caso com os primeiros *trials* de intervenção comunitária para doenças cardiovasculares, como o *Stanford Five City Project* (FARQUHAR *et al.*, 1985), o *Pawtucket Heart Health Program* (CARLETON, LASATER, ASSAF, LEFEBVRE e MCKINLAY, 1987) e o *Minnesota Heart Health Program* (BLACKBURN *et al.*, 1984). Nesses programas, como naqueles que foram formatados e avaliados em seus resultados de acordo com o mesmo paradigma de avaliação há uma forte participação de pesquisadores/avaliadores na definição e formatação do programa que, normalmente, termina antes de sua implementação, mantendo a avaliação longe de questões ligadas à implementação, mas não da implicação de demonstrar os efeitos do programa.

Conforme exemplificado por aqueles programas formatados por pesquisadores acadêmicos e que incluem o conhecimento científico mais atualizado sobre etiologia e prevenção de doenças, na avaliação como experimentação, os programas são concebidos como soluções passíveis de generalização para problemas definidos objetivamente. A origem do programa reside fora da situação

problemática, ou seja, assume-se que o programa vem depois do problema e após as possíveis soluções terem sido identificadas. É essa externalidade que comporta a possibilidade de generalização. Os programas são, dessa forma, essencialmente concebidos como soluções técnicas que incorporam conhecimento científico em resposta a uma situação problemática, sem fazer referência aos atores sociais envolvidos com o problema, sua definição ou solução. O paradigma de referência é aquele do desenvolvimento das drogas (FLAY, 1986). Nesse paradigma, os programas são constituídos por objetos materiais, organizados de acordo com um conjunto de procedimentos estável e previsível. As variações contextuais e na implementação são vistos como ruídos a serem eliminados e controlados da melhor forma possível. Aqui, os mecanismos causais que produzem mudanças não são entendidos como entre os atores que operam o programa, ou nos relacionamentos que eles desenvolvem elementos não humanos do contexto em que o programa é implementado. Pelo contrário, os atores são instrumentos cujos papéis são definidos na descrição e no marco lógico do programa (POLAND, FROHLICH e CARGO, 2008).

Quando os programas são concebidos primordialmente como arranjos técnicos, normalmente, a agência dos atores em questão é vista, da mesma forma, como bastante limitada, já que suas interações com as entidades técnicas dos programas geralmente são predeterminadas por um conjunto rígido de instruções e regras. Isso é feito para assegurar uma máxima fidelidade na implementação do programa, bem como o maior impacto. De fato, na linguagem experimentalista, as pessoas que se beneficiarão dos efeitos dos programas, muitas vezes, são descritas como o alvo do programa. Essa metáfora balística não é benigna por dois aspectos. Primeiro, os alvos situam-se exclusivamente na ponta final de uma transação. Eles estão lá para serem atingidos por algo (por exemplo, uma bala mágica). Em segundo lugar, e mais importante, embora o alvo possa ser móvel, o que o faz mover não é a informação relevante. O único fato importante dos movimentos realizados pelo alvo é sua posição relativa (distância e direção) em relação àquela da pessoa que mira naquele alvo.

Na avaliação como experimentação, o relacionamento entre o avaliador e o programa, incluindo a equipe e metas, é geralmente tido por de mão única. Quanto mais específica a hipótese a ser testada pelo programa e sua avaliação, menos flexibilidade existe para expressão da agência pelos atores envolvidos. Essa “objetificação” dos programas e de seus componentes, idealmente, maximiza o controle sobre o tratamento e os potenciais “fatores de confusão”. Como sujeito conhecedor, as ações dos avaliadores são minimamente influenciadas pelo objeto (programa) em estudo. Alternativamente, uma hipótese-chave para validade externa – a capacidade de generalizar os resultados para outros cenários e alvos – é que o programa não é responsável às ações do avaliador. Em outras palavras, um programa avaliado não é diferente de outro que não foi avaliado. A dimensão epistemológica é, assim, caracterizada por distanciamento e objetividade maximizadas entre o avaliador e o programa sob avaliação.

Nessa prática de tipo-ideal, as escolhas metodológicas são fundamentalmente governadas pelo paradigma experimentalista. O conhecimento é aprimorado pela capacidade de um indivíduo de manipular as condições de produção de um fenômeno e de reproduzir esse fenômeno por meio da manipulação das condições que o criou, reduzindo essas condições ao essencial. A essência da ciência laboratorial é isolar, o mais perfeitamente possível, uma causa putativa e seus efeitos de todas as outras variáveis de confusão. Isso é feito por meio do fechamento e isolamento da situ-

ação experimental. Cook e Campbell (1979) discutiram amplamente o enigma associado com a emulação das condições laboratoriais em pesquisas conduzidas em cenários da vida real, nos quais esse isolamento é praticamente inviável. Na visão deles, o trabalho do avaliador nessas situações é o de tentar impor a delimitação da situação, ao mesmo tempo em que mantem a integridade do tratamento experimental. O teste de controle randômico é uma ferramenta que permite a delimitação de vários (mas não de todos) os aspectos da situação experimental em condições de vida real. Na verdade, de acordo com Campbell (1984), o trabalho pesado de um avaliador pode se resumir a tentar excluir hipóteses plausíveis rivais, a fim de permanecer apenas com o programa em estudo como uma explicação da diferença observada entre aqueles que foram expostos ao programa e aqueles que não o foram. Isso pode ser feito por meio da formatação de estudos que automaticamente excluem algumas ameaças conhecidas à validade interna (CAMPBELL e STANLEY, 1963), como usar um grupo de controle para excluir a possibilidade de que a diferença observada antes e depois no grupo exposto possa ser devido ao efeito de maturação. Isso também pode ser feito pela documentação *post hoc* da baixa plausibilidade de tal hipótese ou pelo controle estatístico (COOK e CAMPBELL, 1979).

A avaliação como uma prática de experimentação é caracterizada por uma forte preocupação com a concepção e implementação de ferramentas de pesquisa que garantam uma validade interna e externa ótima para os resultados da avaliação. Idealmente, isso inclui manter uma distância adequada do programa e assegurar que o trabalho e as atividades que servem à avaliação não interfiram na integridade do programa.

## Avaliação como negociação

Em contraste com a avaliação como experimentação, que foi discutida anteriormente e na qual a ênfase recaiu sobre a identidade técnica do programa, a ênfase da avaliação como negociação é colocada sobre os atores sociais, que são reunidos para formar um espaço social programado. Ao reconhecer que os programas são, essencialmente, sistemas sociais, em que as ações e interações dos atores permanecem sendo a principal dimensão de interesse, os avaliadores como negociadores focam suas práticas nas pessoas que são os alvos diretos do programa. Preocupações metodológicas dizem respeito, principalmente, à compilação das representações do programa a partir das perspectivas dos vários grupos de atores. A primazia dos atores sociais e dos processos sociais inaugura um caminho real para o construtivismo social como um paradigma para avaliação de programas.

Esse tipo-ideal de prática de avaliação foi defendido principalmente (mas longe de ter sido de forma exclusiva) por Guba e Lincoln (1989) como sendo a “Avaliação de Quarta Geração”. De fato, apesar da perspectiva construtivista de Guba e Lincoln ser provavelmente a mais reconhecida e citada na literatura de avaliação no campo da promoção da saúde, alguns dos bem conhecidos pioneiros na área de avaliação também fundaram suas abordagens na epistemologia construtivista. O trabalho de avaliadores como Wholey e Newcomer (1989) e Stake (1975) é bem conhecido no campo da avaliação. Ambos são associados com a ideia fundamental de que os programas, quando implementados, normalmente diferem da intenção do programa original pensado pelos planejadores; os programas são construídos à medida que passam por vários estágios de interação entre grupos de atores associados.

Dessa forma, um dos papéis-chave de uma avaliação é destacar esse processo por meio do apoio a melhorias nos programas ou simplesmente pela definição de um programa em seus próprios termos. Ao enfatizar a necessidade dos avaliadores de se aproximar da ação onde quer que ela ocorra, os dois trabalhos (e outros) contribuíram, no início da década de 1970, para o declínio do monopólio do paradigma quasi-experimental no campo da avaliação. Tomando por base o conceito de programas como construções sociais, a contribuição mais importante feita por aqueles autores foi defender, com sucesso, o pluralismo de métodos e o relativismo dos valores quando se avalia um programa.

Muito embora existam múltiplas graduações de construtivismo social, a mais utilizada para embasar avaliações construtivistas é aquela que se sustenta no trabalho e na tradição acadêmica da Escola de Sociologia de Chicago, principalmente por meio do trabalho de Anselm Strauss e de suas proposições metodológicas, que estão embutidas nas investigações do tipo teoria fundamentada em dados (GLASER e STRAUSS, 1967). Embora a Escola de Sociologia de Chicago seja associada a um grande número de teóricos e especialistas em metodologia no campo da sociologia, uma de suas inovações mais críticas e relevantes é o chamado interacionismo simbólico. A primeira e mais fundamental proposição do interacionismo simbólico é a de que os seres humanos agem essencialmente como resultado de suas percepções e interpretações do significado que eles atribuem às ações de outros seres humanos. A segunda proposição (corolário) diz que é por meio das interações com outros que esses significados são desenvolvidos, transformados e são passíveis de manipulação (LE BRETON, 2004). Assim, o principal motor da ação é a representação de situações sociais e essas representações são moldadas pelos processos sociais de interação que ocorrem nessas situações.

Ao desenvolver a abordagem da “Avaliação de Quarta Geração”, Guba e Lincoln (1989) lançaram mão, de forma extensiva, do interacionismo simbólico, da metodologia da teoria fundamentada em dados, e do trabalho de Robert Stake (ABMA, 2005). Eles argumentam que os programas somente podem ser compreendidos por meio das representações que vários atores desenvolvem e sobre as quais atuam por meio das interações sociais com outros atores do programa.

Tanto na concepção de Guba e Lincoln quanto em Stake, a principal tarefa dos avaliadores é apoiar o desenvolvimento de uma representação compartilhada dos programas entre os atores e isso é, essencialmente, feito por meio da negociação. A dimensão teleológica na avaliação como prática da negociação é, dessa maneira, contribuir para melhorias nos programas à medida que ele é implementado, reconhecendo que as discrepâncias entre o programa (como ele foi planejado) e como, de fato, é praticado, são inevitáveis. Aqui, espera-se que os programas sejam aprimorados com a elaboração de uma representação comum do programa entre os *stakeholders* do programa, por meio da qual, presume-se que uma representação comum traga melhorias ao programa por meio de uma melhor coordenação e alinhamento das ações dos seus *stakeholders*. Desenvolver essa representação compartilhada é a principal tarefa dos avaliadores, que acontece, principalmente, pelo referimento das avaliações para atores do programa e, assim, permitindo que eles revisem suas representações em linha com aquelas de outros atores do programa.

Em linha com o interacionismo simbólico, a dimensão ontológica dessa prática de tipo-ideal é encontrada nas representações desenvolvidas pelos atores de um programa. Assim, os programas não possuem uma realidade objetiva, no sentido da primazia dos objetos materiais e técnicos. São

as representações que os atores do programa constroem desses objetos e o significado simbólico que eles atribuem a essas representações por meio de suas interações que orientam as ações dos indivíduos nos espaços sociais dos programas. Assim, em qualquer programa, teoricamente, existem tantas representações quanto o número de indivíduos envolvidos na iniciativa. O papel do avaliador é o de buscar revelar essas representações e ajudar na construção de um consenso entre os atores do programa.

Essa tarefa requer que o avaliador interaja estreitamente com os atores do programa. A epistemologia é construtivista, o que significa que tanto o avaliador quanto os atores do programa são sujeitos nessa construção. Não existe uma distância prescrita entre o programa e o avaliador. De fato, a representação compartilhada do programa que resulta da avaliação só pode ser construída coletivamente e o papel central do avaliador é formatar e implementar os mecanismos por meio dos quais as representações dos atores possam ser confrontadas direta ou indiretamente entre si. Por exemplo, mecanismos na forma de *feedback* extraído de diálogos, documentação ou análises observacionais, fornecem um meio para os atores do programa situarem seus próprios interesses e interpretações dos eventos de um programa em relação àqueles de outros atores.

Embora existam variações nas ferramentas técnicas específicas usadas para coletar e analisar dados, a dimensão metodológica na prática de tipo-ideal da avaliação como negociação preocupa-se, essencialmente, em produzir a representação ou teoria mais crível e confiável de um programa. Em muitos casos, os avaliadores fizeram referência à metodologia da teoria fundamentada em dados, desenvolvida por Glaser e Strauss (1967) e, mais tarde, por Strauss e Corbin (1998) a fim de criar, de forma indutiva, uma representação compartilhada a partir das representações específicas dos indivíduos acerca do programa. Técnicas específicas de coleta e análise de dados garantem um processo de confrontação contínuo entre a teoria que está emergindo na esteira dos dados levantados e o fenômeno empírico, que é iluminado pela teoria. A principal contribuição prática e metodológica de Guba e Lincoln (1989) reside na proposição de que os Círculos de Diálogo Hermenêutico devem ser implementados para que uma representação compartilhada seja construída. Como negociadores e indivíduos trabalham na construção de consensos, os avaliadores orquestram diálogos entre os atores do programa a fim de criar uma construção conjunta que inclua o máximo de elementos possíveis das representações de cada indivíduo.

No campo da promoção da saúde, muitas variações nessa prática de avaliação do tipo-ideal são descritas como de extrema relevância, dada a natureza da promoção da saúde. Apresentando a teoria de Stake de avaliação responsiva, ABMA (2005) argumenta que “A avaliação responsiva não é responsiva apenas às características únicas e às ideias emergentes no campo da promoção da saúde, mas também é sinérgica com a promoção da saúde” (p. 287).

Em particular, essa abordagem é vista como tendo o potencial de retificar os processos sociais que perpetuam as desigualdades na saúde por meio da criação de parcerias igualitárias com ampla participação de especialistas, pesquisadores, atores leigos e da comunidade. As avaliações, dessa forma, exibiriam as relações de trabalho e as parcerias existentes entre os vários atores como objeto do trabalho. As avaliações de programas comunitários, por exemplo, descrevem um conjunto de atores com uma gama de preocupações em relação às necessidades comunitárias, das organizações não governamentais, dos governos e em termos de pesquisa à medida que eles se juntam para definir uma situação problemática e construir estratégias corretivas. Aqui, o papel do avaliador é buscar rastrear esse processo, por meio

da coleta e interpretação de dados qualitativos, a fim de capturar as perspectivas, opiniões e experiências vivenciadas pelos vários atores. Na literatura disponível, os resultados analíticos desses estudos apresentam representações de forma agregada, nas quais os vários grupos de atores são identificados em algum nível de coletividade. Uma coletividade pode ser algo bastante geral, por meio da qual uma representação mais abrangente é atribuída a todos os atores do programa (LANTZ, VIRUELL FUENTES, ISRAEL, SOFTLEY e GUZMAN, 2001). De forma alternativa, as representações podem ser distintas ou específicas aos vários grupos de atores, de acordo com um critério de admissão predefinido (ex. membro da comunidade, profissional de saúde) (SCHULZ *et al.*, 2001). Além de ser organizado de acordo com a identidade dos atores de um programa, os dados podem ser interpretados em termos de sua correspondência com um estágio no planejamento ou avaliação de um programa (FARQUHAR, PARKER, SCHULZ, e ISRAEL, 2006) ou com um conjunto de conceitos associados aos princípios norteadores do programa (ex. pesquisa participativa) (SAVAGE *et al.*, 2006).

Mais especificamente, uma iniciativa como o *Detroit Community-Academic Urban Research Center* (URC) representa um exemplo em que a avaliação exerce um papel negociador ao construir uma representação coletiva do processo de construção de parcerias entre a gama de atores envolvidos. As avaliações realizadas nessa iniciativa foram descritas no âmbito regional, em que os atores comunitários e acadêmicos formaram um comitê (o URC), requereram (e alocaram) fundos para iniciativas locais, e também, no âmbito local, no qual as iniciativas e projetos abordaram um problema específico em um local específico. As avaliações, tanto no âmbito regional quanto no âmbito local, tiveram por base as representações dos atores acerca do processo de construção de parcerias, bem como a operação dos serviços e aquisição de recursos.

A *East Side Village Health Worker Partnership* (ESVHWP) foi um de 12 projetos financiados pelo *Detroit Community-Academic URC*, entre os anos de 1995 e 1999 (LANTZ *et al.*, 2001). A avaliação desse projeto se apoiou na derivação das representações do programa levantadas com base em entrevistas de informantes, observações e revisão documental em relação a quatro estágios do processo de planejamento do programa (FARQUHAR *et al.*, 2006). Enquanto os dados foram coletados ao longo do tempo, a análise de um conjunto de dados de um período de tempo foi interpretada de acordo com sua relação com quatro estágios do planejamento (ou seja, avaliação e análise do problema, metas e objetivos, desenho e implementação, avaliação final). A avaliação final expandiu seu papel como um processo de negociação ao fornecer *feedback* das representações construídas aos vários atores do programa, o que serviu tanto como aprendizado e quanto para o processo de tomada de decisão como uma fonte de legitimação para as experiências e representações vivenciadas pelos que conduziam as intervenções (ou seja, pelos profissionais da saúde) e pelos membros do comitê gestor.

Por exemplo, ao oferecer *feedback* às interpretações das reuniões dos profissionais de saúde, a avaliação valida a presença de quatro problemas distintos com base nos quais os profissionais da saúde se dividiram em quatro subgrupos, cada um com seu foco individual e conjunto de atividades. Essa abordagem é de algum modo distinta do círculo de diálogo hermenêutico, no sentido de que ela não impõe uma declaração de problema consensual.

Outras avaliações do *Detroit Community-Academic URC* (LANTZ *et al.*, 2001) e também do ESVHWP (SCHULZ *et al.*, 2001) identificaram uma gama diversa de atores que se uniram para formar um centro de pesquisa (*i.e.*, URC Board) ou um comitê gestor.

A avaliação do Conselho do URC teve por base uma representação coletiva das forças internas e externas da parceria, nomeadamente em que medida a parceria conseguiu caminhar como um grupo na direção dos objetivos, e como o grupo encontrou condições que facilitaram ou atuaram como barreiras aos objetivos da pesquisa em saúde pública. Os resultados da avaliação foram organizados em categorias que identificavam a satisfação geral, fatores facilitantes e barreiras ao alcance dos objetivos. Por exemplo, o grupo identificou algumas questões como preocupações, tais como gestão do tempo, distribuição de recursos e o equilíbrio entre necessidades da comunidade e da pesquisa. Como fatores facilitadores ao alcance do objetivo, a avaliação apontou a representação comunitária no Conselho do URC, a confiança, a existência de conhecimentos relevantes e apoio organizacional. Alternativamente, as avaliações extraíram informações das entrevistas de informantes concedidas pelo *Village Health Workers* e pelo comitê gestor como parte do ESVHWP, que visavam identificar como a parceria contribuiu para aprimoração da pesquisa, das atividades programáticas, das relações comunitárias e da participação comunitária, acadêmica e das organizações e instituições da prática. As entrevistas foram interpretadas de acordo com esses temas primordiais e subcategorias específicas a esses temas foram induzidas a partir das entrevistas e os resultados foram apresentados como uma representação geral coletiva do comitê gestor e do *Village Health Workers*. Por exemplo, o aprimoramento das relações comunitárias foi descrito como um fortalecimento das redes sociais no *Village Health Workers*, como relações mais fortes entre eles e os membros do comitê gestor e relações mais fortes entre as organizações acadêmicas, da prática e de base comunitária. Dentro dessas subcategorias, questões específicas, como a confiança, diferenças de poder e comunicação, além de algumas estratégias usadas para fazer frente aos desafios (ex., piqueniques anuais e退iros) foram apresentadas.

Duas críticas são normalmente associadas a esse paradigma construtivista de avaliação. O primeiro é o da natureza essencialmente descritiva dos resultados dessas práticas avaliativas. Ou seja, enquanto é possível que o consenso seja alcançado, já que as condições mudam ao longo do tempo e do espaço, também o é a capacidade de um programa alcançar o consenso entre os atores do programa. Conforme ilustrado acima, o consenso talvez não seja algo desejável. Essa descrição, portanto, não permite que o conhecimento seja construído a partir do processo pelo qual negociações e compromissos, entre os atores do programa, ocorrem com vistas ao alcance dos objetivos comuns e acordados. De fato, na versão renovada dos estudos de avaliabilidade, a construção de um entendimento ou representação comum do programa e de seus componentes é apenas o ponto inicial a partir do qual a formatação da avaliação do projeto tem início (THURSTON *et al.*, 2003; THURSTON e POTVIN, 2003). A outra crítica tem a ver com a subestimação da influência e do poder políticos de que gozam os avaliadores nesses processos. Como o negociador-mestre, se poderia ter a compreensão errônea de que o avaliador é neutro nesse processo de construção, o que, certamente, não é o caso. O uso de mecanismos formais que asseguram uma justa distribuição de papéis e de responsabilidades na condução de projetos de avaliação é cada vez mais visto como uma garantia necessária.

## Avaliação como reflexividade organizada

Há mais outra prática de avaliação de tipo-ideal que está surgindo na literatura de avaliação e que está, aos poucos, adentrando o domínio da avaliação de programas nos campos da saúde pública e

da promoção da saúde. Pawson e Tilley (1997) desenvolveram as bases de uma abordagem realista crítica de avaliação, lançando mão da obra do filósofo Roy Bhaskar (1978). “A avaliação realista, em sua essência, tem seu foco no desenvolvimento de explicações sobre as consequências das ações sociais que contribuem para uma melhor compreensão do por que, onde, e para quem os programas funcionam ou não funcionam. Para esse fim, os avaliadores realistas colocam grande ênfase em (a) identificar os mecanismos que produzem efeitos programáticos que podem ser observados e (b) testar esses mecanismos e as outras variáveis contextuais... que possam gerar impactos sobre os efeitos que são observados” (PAWSON e TILLEY, 2004, p. 359). Trabalhar sob o pressuposto realista de que os programas são um de muitos sistemas de ações que operam simultaneamente em um dado momento em um contexto, a tarefa do avaliador realista é, idealmente, identificar os mecanismos operativos do programa e como eles interagem com as condições contextuais (incluindo com outros programas) para ser associado aos padrões de resultados pretendidos e não pretendidos (PAWSON e TILLEY, 2004). Em um mundo complexo de interações constantes, é por meio das conexões que os programas criam e mantêm com as condições contextuais que seus impactos podem ser compreendidos e identificados. Assim, os avaliadores de programa são, fundamentalmente, projetistas e operadores de aparelhos de reflexividade organizada, por meio dos quais os atores dos programas podem compreender como os programas contribuem para transformar a vida e as condições das pessoas.

Nós propomos que a dimensão teleológica de uma avaliação como uma prática de reflexividade organizada diz respeito ao entendimento das mediações operadas por um programa na transformação do mundo. O reconhecimento da complexidade inerente dos programas e de seus mecanismos operativos em sistemas abertos leva à proposição realista crítica de que a causalidade não pode ser estabelecida somente pela observação empírica das constantes conjugações de eventos (BHASKAR, 1978). De fato, é um sistema aberto no qual uma multidão de mecanismos estão constantemente interagindo e transformando o sistema, em que padrões de recorrência exatos não podem ser observados. É apenas em condições controladas, nas quais algum fechamento ou quase-fechamento é exercido no sistema, que podem ocorrer observações semelhantes a leis. Em todas as outras situações, o principal papel do cientista é identificar e considerar o efeito dos mecanismos irrelevantes (COOK e CAMPBELL, 1979). Na falta de tal fechamento, as observações empíricas somente pode fazer sentido quando relacionadas à teoria sobre configurações-padrão de mecanismo-contexto-resultado que elas ajudam a fortalecer. Como mecanismos em operação dentro de um contexto, um programa que interage com outros mecanismos existentes se torna um mediador da constante transformação do contexto. Sua presença altera o curso normal dos eventos, seja reforçando ou temperando o impacto de outras forças operativas existentes.

Nessa prática de tipo-ideal, os programas são concebidos como sistemas de ações nos quais os atores humanos e não humanos operam os mecanismos putativos do programa por meio de suas ações (POTVIN, GENDRON e BILODEAU, 2006). A realidade ontológica é, assim, uma de ação e interações. É o que é feito pelas vários atores do programa, tanto os humanos quanto os não humanos, que contribui para transformar o contexto na direção esperada pelos que iniciaram o programa. Assim, a realidade mais relevante do programa reside nas ações realizadas no espaço do programa. Essas, como todas as ações, envolvem dois tipos de realidades interativas. Um tipo é tangível e objetivo, e consiste das pessoas, objetos, recursos e espaços físicos que situam a ação e a torna viável. Normalmente, é o que está agrupado sob o plano de trabalho

e marcos lógicos associados ao desenvolvimento do programa. O segundo tipo é simbólico e inclui os significados que os atores atribuem às ações e papéis designados, de forma implícita ou tácita, aos atores em relação às ações. Isso é o que é normalmente capturado pelas representações dos programas. Uma ontologia realista dos programas tem que abranger esses dois tipos de realidade e suas interações.

A fim de entender como um programa media as transformações que acontecem em um espaço e tempo específicos, o avaliador não tem escolha que não seguir as ações que estão sendo desenvolvidas e se desenrolando em nome do programa, bem como suas ramificações e impactos no contexto social do programa. Acompanhar a ação é a dimensão metodológica mais importante de uma ferramenta reflexiva. É por meio de suas ações e prática que os atores ativam suas intenções muito mais do que por meio do discurso que carregam sobre suas intenções (representações). Esse é um dos limites bastante reconhecido da fenomenologia; que a representação do indivíduo a respeito de sua ação é normalmente cega quanto às condições estruturais que restringem e tornam a ação possível (BOURDIEU, 1972). Idealmente, é por meio da descrição e da análise dessas ações que instanciam o programa que os atores podem refletir de forma crítica sobre o programa e, racionalmente, realizar as constantes correções que são necessárias para fortalecer a configuração padrão de mecanismo-contexto-resultados. Por exemplo, à medida que os profissionais ou os pesquisadores aumentam o grau de envolvimento com um grupo de atores em questão, a representação deles sobre o que é um problema e sobre como ele pode ser resolvido, provavelmente, muda. Às vezes, devido à natureza da pesquisa, comunidade ou indivíduos envolvidos, os planos de um programa são adaptados (BISSET, CARGO, DELORMIER, MACAULEY e POTVIN, 2004), outras vezes, no entanto, apesar de reconhecer a necessidade de mudança ou expansão dos planos, a latitude dos profissionais é restringida por conta de uma necessidade imposta para que se mantenha um curso (HAWE e RILEY, 2005). A avaliação como reflexividade organizada acompanha “o porquê” e “como” ocorrem transformações dessa sorte, de maneira a oferecer inteligibilidade às adaptações locais.

O princípio subjacente sobre o qual se baseia a metodologia de acompanhamento das ações defende que um programa é mais precisamente e mais amplamente compreendido por meio de um estudo aprofundado dos executores do programa em ação. Ao acompanhar a ação dos executores do programa, nós obtemos uma visão do programa a partir do seu funcionamento essencial. Essa abordagem coloca em contraste uma compreensão do programa baseada em relatos do que o programa pode visar alcançar e exemplos de atividades exemplares. Da mesma forma que a sociologia da associação rejeita definições ostensivas do que seja o social (LATOUR, 2006), pode-se, igualmente, argumentar que ao se definir um programa com base em relatos e exemplos, nós contornamos o verdadeiro trabalho de aprendizado sobre a manutenção de um programa, sobre o que mantém um programa coeso e operativo.

Duas características distintas podem ser identificadas em uma metodologia que se baseia no acompanhamento das ações. Em primeiro lugar, ela não se limita a investigar ações que creem fazer parte de uma cadeia causal de eventos que leva a um conjunto predefinido de resultados. Embora certamente sejam importantes, essas ações, no entanto, capturam uma porção limitada do que os profissionais fazem e assumem sua parte e são movidas por um modelo de realização único. Ademais, as ações só podem ser planejadas até certo ponto, após o qual se tornam desvirtuadas e se desenvolvem espontaneamente dadas as condições que emergem.

Esse último ponto traz à tona a segunda característica singular das avaliações que acompanham a ação. Acompanhar a ação requer que as experiências vivenciadas que ocorrem dentro e fora dos limites típicos de um programa sejam consideradas. Em primeiro lugar, ações que ocorrem fora do espaço físico e do tempo da realização do programa formam opiniões e impressões que impactam como os participantes do programa e os profissionais envolvidos interagem com o programa. Em segundo lugar, no curso do tempo e espaço de um programa, sua forma não é apenas determinada pelas instruções e materiais que são entregues, mas, pelo contrário, pelas formas com que os vários atores que estejam presentes interajam com as instruções e os materiais, e entre si.

Finalmente, a dimensão epistemológica dessa prática de tipo-ideal enfatiza a posição reflexiva dos atores envolvidos no programa, bem como daqueles envolvidos na avaliação. De fato, ao acompanhar a ação à medida que ela se desenvolve e se desdobra, o avaliador se torna, necessariamente, uma parte importante do contexto do programa. Como tal, não se pode excluir a possibilidade de que as conexões entre o programa e sua avaliação, que são necessárias a fim de que ela acompanhe as ações que acontecem no programa, estejam, elas mesmas, mediando as transformações no programa, no contexto e também na avaliação. Não só um programa avaliado difere do que ele poderia ser caso não tivesse sido submetido a uma avaliação, como também os projetos de avaliação também diferem de acordo com os programas e versões de programas aos quais se vinculam. Assim, uma avaliação precisa do papel mediador do programa nas transformações contextuais leva em consideração o papel transformador da avaliação no programa. Em última instância, uma avaliação é uma intervenção exercida sobre um programa (MARK *et al.*, 2000).

## Conclusão

Neste capítulo, nós exploramos três práticas de tipo-ideal de avaliação. Fundamentamos nossa discussão desses tipos-ideais em uma concepção pragmática da prática que supõe quatro dimensões coerentes subjacentes: teleológica, ontológica, epistemológica e metodológica, conforme indicado na Figura 1, adaptada de Gendron (2001).

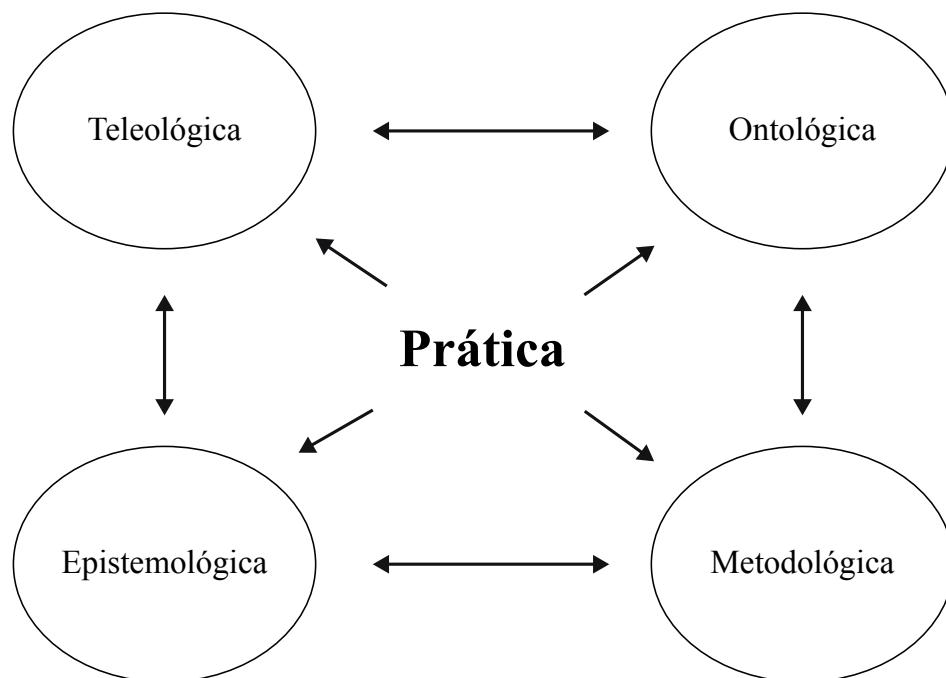
Nós, deliberadamente, escolhemos a dimensão metodológica como o ponto de entrada principal para nossa discussão. Certamente, entrar nessa exploração por meio de qualquer outra dimensão paradigmática teria levado a definição de tipos-ideais diferentes. Por exemplo, esforços para destacar a dimensão teleológica das práticas teria levado a distinção entre práticas focadas na utilização (PATTON, 1997) e de teste de hipótese/produção de conhecimento (COOK e CAM-PBELL, 1979), e uma ênfase na dimensão epistemológica teria colocado em contraste a prática participativa daquela impulsionada por especialista. *Although there should be some correlations between the outcomes of the various points of entry for exploring evaluation practices (it is more difficult to couple an evaluation as experimentation practice with a participatory practice), we do not think that there is an exact correspondence.*

Ao caracterizarmos as práticas de avaliação por meio de suas implicações metodológicas, queríamos enfatizar a riqueza e diversidade dos vários pontos de entrada que podem ser usados para esse tipo de exploração, propondo outras dimensões daquelas normalmente usadas a partir das quais definem-se questões práticas da avaliação. Também é importante notar que nós não discuti-

mos a dicotomia qualitativo/quantitativo. Para nós, isso não é uma questão metodológica, mas técnica, que diz respeito à natureza dos dados disponíveis, ao invés da maneira pela qual um sujeito pode extrair conhecimento sobre um dado objeto.

Mostramos que todos os três tipos ideais descritos em nosso artigo podem ser encontrados na literatura de avaliação no campo da promoção da saúde. Todos têm sido usados, com diferentes frequências e com variados níveis de sucesso, e achamos que existem situações nas quais a promoção da saúde é mais bem servida pelos três tipos ideais. Queremos, no entanto, apontar para o interesse em se considerar uma avaliação como uma prática de reflexividade organizada para a avaliação de promoção da saúde. Como uma área de atuação, a promoção da saúde ainda é muito recente e não pode contar com um corpo de configurações padrão de contexto-mecanismo-resultado importante solidamente estabelecido. Há, portanto, muito trabalho a ser feito a fim de criar um repertório de intervenções que possa ser operado eficientemente em vários contextos. Aprender conforme avançamos e organizar o conhecimento em teorias de intervenção coerentes é certamente uma estratégia que deve ser explorada e expandida no sentido de uma base de conhecimento adequada para a promoção da saúde. A reflexividade organizada, como prática metodológica, acomoda essa estratégia de desenvolvimento do conhecimento. Ademais, como mostrado nos vários capítulos deste livro, essa prática de tipo-ideal é compatível com muitas soluções práticas e teóricas, examinadas neste livro, para alinhar os valores da promoção da saúde ao rigor científico da avaliação.

Figura 1 – As quatro dimensões inter-relacionadas da prática (Adaptado de Gendron, 2001, p. 34)



## Referências Bibliográficas

ABMA, T.A. Responsive evaluation: its meaning and special contribution to health promotion. **Evaluation and Program Planning**, vol. 28, p. 279-289, 2005.

- BHASKAR, R. **A realist theory of science**. Hassocks, UK: Harvester Press, 1978.
- BISSET, S.; CARGO, M.; DELORMIER, T.; MACAULEY, A.; POTVIN, L. Legitimizing diabetes as a community health issue: a case analysis of an Aboriginal community in Canada. **Health Promotion International**, vol. 19, p. 317-326, 2004.
- BLACKBURN, H.; LUEPKER, R.V.; KLINE, F.G.; BRACHT, N.; CARLAW, R.; JACOBS, D. *et al.* The Minnesota heart health program: a research and demonstration project in cardiovascular disease prevention. In: MATARAZZO, J.D.; MILLER, N.E.; WEISS, S.M.; HERD, J.A.; WEISS, S.M. (Eds.). **Behavioral health**: a handbook of health enhancement and disease prevention. Silver Spring, MD: John Wiley e Sons, 1984. p. 1171-1178.
- BOURDIEU, P. **Esquisse d'une theorie de la pratique**. Geneve: Drozée, 1972.
- BROWN, P. The role of the evaluator in comprehensive community initiatives. In: CONNELL, J.P.; KUBISCH, A.C.; SCHORR, L.B.; WEISS, C.H. (Eds.). **New approaches to evaluating community initiatives**: concepts, methods and contexts. New York: Aspen Institute, 1995. p. 201-225.
- CAMPBELL, D.T. Can we be scientific in applied social science?. In: CONNOR, R.F.; ALTMAN, D.G.; JACKSON, C. (Eds.). **Evaluation studies review annual**, vol. 9. Beverly Hills, CA: Sage, 1984. p. 26-48.
- \_\_\_\_\_; STANLEY, J.C. **Experimental and quasi-experimental designs for research**. Chicago: Rand McNally, 1963.
- CARLETON, R.A.; LASATER, T.M.; ASSAF, A.; LEFEBVRE, R.C.; McKINLAY, S.M. The Pawtucket Heart Health Program: 1. An experiment in population-based disease prevention. **The Rhode Island Medical Journal**, vol. 70, p. 533-538. 1987.
- COOK, T.D.; CAMPBELL, D.T. **Quasi-experimentation: design and analysis issues for field settings**. Boston: Houghton Mifflin, 1979.
- CROZIER, M.; FRIEDBERG, E. **L'acteur et le système**. Paris: Seuile, 1977.
- FARQUHAR, J.W.; FORTMANN, S.P.; MacCOBY, N.; HASSELL, W.L.; WILLIAMS, P.T.; FLORA, J.A. *et al.* The stanford five city project: design and methods. **American Journal of Epidemiology**, vol. 122, p. 323-334, 1985.
- FARQUHAR, S.A.; PARKER, E.A.; SCHULZ, A.J.; ISRAEL, B.A. Application of qualitative methods in program planning for health promotion interventions. **Health Promotion Practice**, vol. 7, p. 234-242, 2006.
- FLAY, B.R. Efficacy and effectiveness trials (and other phases of research) in the development of -promotion programs. **Preventive Medicine**, vol. 15, p. 451-474, 1986.
- GENDRON, S. **La pratique participative en santé publique, l'émergence d'une paradigme**. 2001. Tese (Doutorado). Universidade de Montreal. Montreal, 2001. Disponível em: <<http://proquest.umi.com/pqdlink?Ver=1&Exp=01-15-2013&eFMT=7&eDID=726370931&eRQT=309&eattempt=1>> Acesso em jan. 2008.
- GLASER, B.; STRAUSS, A. **The discoveries of grounded theory**: strategies for qualitative research. Chicago: Aldine, 1967.

- GUBA, Y.; LINCOLN, E. **Fourth generation evaluation**. Thousand Oaks CA: Sage, 1989.
- HARTZ, Z.A.; DENIS, J-L.; MOREIRA, E.; MATIDA, A. From knowledge to action: challenges and opportunities for increasing the use of evaluation in health promotion policies and practices. In: POTVIN, L.; McQUEEN, D.V. (Eds.). **Health promotion evaluation practices in the Americas: values and research**. New York: Springer, 2008
- HAWE, P.; RILEY, T. Ecological theory in practice: Illustrations from a community-based intervention to promote the health of recent mothers. **Prevention Science**, vol. 6, p. 227-236, 2005.
- KUHN, T.S. **The structure of scientific revolutions**. Chicago: University of Chicago Press, 1962.
- LANTZ, P.M.; VIRUELL-FUENTES, E.; ISRAEL, B.A.; SOFTLEY, D.; GUZMAN, R. Can communities and academia work together on public health research? Evaluation results from a community-based participatory research partnership in Detroit. **Journal of Urban Health**, vol. 78, p. 495-507, 2001.
- LATOUR, B. **Changer de societe: refaire de la sociologie**. Paris: La decouverte, 2006.
- LE BRETON, D. **L'interactionisme symbolique**. Paris: Presses universitaires de France, 2004.
- LEVY, R. Croyance et doute: une vision paradigmatic des methodes qualitatives: ruptures. **Revue transdisciplinaire en sante**, vol. 1, p. 92-100, 1994.
- MARK, M.M.; HENRY, G.T. The mechanisms and outcomes of evaluation influence. **Evaluation**, vol. 10, p. 35-57, 2004.
- \_\_\_\_\_; \_\_\_\_\_; JULNES, G. **Evaluation: an integrated framework for understanding, guiding, and improving policies and programs**. San Francisco CA: Jossey-Bass, 2000.
- MERCILLE, G. A realist approach to synthesizing evaluation results. In: POTVIN, L.; McQUEEN, D.V. (Eds.). **Health promotion evaluation practices in the Americas: values and research**. New York: Springer, 2008.
- PATTON, M.Q. **Utilization-focused evaluation: the new century text**. Thousand Oaks: Sage, 1997.
- PAWSON, R.; TILLEY, N. **Realistic evaluation**. London, UK: Sage, 1997.
- \_\_\_\_\_; \_\_\_\_\_; Realistic evaluation. In: MATHISON, S. (Ed.). **Encyclopedia of evaluation**. Newbury Park, CA: Sage, 2004. p. 362-367.
- POLAND, B.; FROHLICH, K.L.; CARGO, M. Context as a fundamental dimension of health-promotion program evaluation. In: POTVIN, L.; McQUEEN, D.V. (Eds.). **Health promotion evaluation practices in the Americas: values and research**. New York: Springer, 2008.
- POTVIN, L.; CARGO, M.; McCOMBER, A.; DELORMIER, T.; MACAULAY, A.C. Implementing participatory intervention and research in communities: lessons from the Kahnawake schools diabetes prevention project. **Social Science e Medicine**, vol. 56, p. 1295-1305, 2003.
- \_\_\_\_\_; GENDRON, S.; BILODEAU, A. Três posturas ontológicas concernentes a natureza e aos programas de saúde: implicações para a avaliação. In: BOSI, M.L.M.; MERCADOU, F.J. (Eds.). **Avaliação qualitativa de programas de saúde: enfoques emergentes**. Petrópolis: Vozes, 2006.
- \_\_\_\_\_; McQUEEN, D.V. Practical dilemmas for health-promotion evaluation. In: POTVIN, L.;

McQUEEN, D.V. (Eds.). **Health promotion evaluation practices in the Americas**: values and research. New York: Springer, 2008.

PRESKILL, H.; CARACELLI, V. J. Current and developing conceptions of use: evaluation use topical interest group survey results. **Evaluation Practice**, vol. 18, p. 209-225, 1997.

SAVAGE, C. L.; XU, Y.; LEE, R.; ROSE, B.L.; KAPPESER, M.; Anthony, S.P. A case study in the use of community-based participatory research in public health nursing. **Public Health Nursing**, vol. 23, p. 472-487, 2006.

SCHULZ, A. J.; ISREAL, B.A.; PARKER, E.A.; LOCKETT, M.; HILL, Y.; WILLS, R. The east side village health worker partnership: Integrating research with action to reduce health disparities. **Public Health Reports**, vol. 116, p. 548-557, 2001.

SHADISH, W.R.; COOK, T.D.; LEVITON, L.C.L. **Foundations of program evaluation**. Thousand Oaks, CA: Sage, 1991.

SPRINGETT, J. Participatory approaches to evaluation in health promotion. In: ROOTMAN, I.; GOODSTADT, M.; HYNDMAN, B.; McQUEEN, D.V.; POTVIN, L.; SPRINGETT, J.; ZIGLIO, E. (Eds.). **Evaluation in health promotion**: principles and perspectives. Copenhagen: WHO Regional Publications, 2001. p. 83-105. European Series, n. 92.

STAKE, R.E. (Ed.). **Evaluating the arts in education: a responsive approach**. Columbus Ohio: Merril, 1975.

STRAUSS, A.; CORBIN, J. **Basics of qualitative research**: techniques and procedures for developing grounded theory. Thousand Oaks, CA: Sage, 1998.

SUCHMAN, E. **Evaluative research**. New York: Russel Sage Foundation, 1967.

THURSTON, W.E.; GRAHAM, J.; HATFIELD, J. Evaluability assessment a catalyst for program change and improvement. **Evaluation and the Health Profession**, vol. 26, p. 206-221, 2003.

\_\_\_\_\_; POTVIN, L. Evaluability assessment: a tool for incorporating evaluation in social change programs. **Evaluation**, vol. 9, p. 453-469, 2003.

WEBER, M. The essentials of bureaucratic organization: an ideal-type construction. In: MERTON, R.K.; GRAY, A.P.; HOCKEY, B.; SELVIN, H. (Eds.). **Reader in bureaucracy**. New York: Free Press, 1952. p. 18, 21-22.

WEISS, C.H. Evaluation. **Methods for studying programs and policy**. 2nd edition. Upper Saddle River NJ: Prentice Hall, 1998.

WHOLEY, J.S.; NEWCOMER, K.E. (Eds.). **Improving government performance**: evaluation strategies for strengthening public agencies and programs. San Francisco: Jossey-Bass, 1989.