

Construção de Indicadores Qualitativos para Avaliação de Mudanças

The Construction of Qualitative Indicators for the Evaluation of Changes

Maria Cecília de Souza Minayo¹

PALAVRAS-CHAVE:

- Educação Médica.
- Avaliação em Saúde.
- Sociologia Médica.

KEYWORDS:

- Medical Education.
- Health Evaluation.
- Medical Sociology.

RESUMO

Apresenta-se uma reflexão teórico-metodológica focada na construção de indicadores qualitativos, voltados principalmente para a avaliação de mudanças no campo educacional. Foi escrito para subsidiar o processo de autoavaliação de um conjunto de escolas médicas que desenvolvem um programa de transformação de seus métodos tradicionais para outros fundamentados na filosofia e na prática de aprendizagem ativa. Nele são discutidos e problematizados: o conceito de indicador; a produção de indicadores para pesquisa avaliativa e a construção de indicadores qualitativos. Esses últimos são tratados sob duas formas mais frequentes: dentro da lógica quantitativa e da lógica qualitativa. Indicadores qualitativos a partir da visão quantitativa da realidade são construídos principalmente por meio de escalas. Dentro da ótica qualitativa, são elaborados a partir das representações e práticas dos sujeitos envolvidos na investigação. Enfatizando a necessidade de produção de indicadores de mudança que levem em conta estruturas, processos, relações e a contribuição da subjetividade, conclui-se que a validade interna desses indicadores pode ser alcançada quando as categorias empregadas em sua construção são fundamentadas teórica e contextualmente e possuem significados mútuos e partilhados entre os participantes.

ABSTRACT

This essay presents a theoretical and methodological approach to the construction of qualitative indicators, mainly for evaluation of changes in the educational field. It was prepared as a contribution to the self-evaluation process of a group of medical schools that are implementing a program for transformation of their traditional methods to other fundaments in philosophy and to the practice of active learning. The following concepts are discussed: indicator, production of indicators for evaluative research and construction of qualitative indicators, the latter in their two most frequent aspects: within the quantitative logic and the qualitative logic. Qualitative indicators within the quantitative view are mainly constructed by scales. In the qualitative perspective they emerge from the representations and practices of the participants of an investigation. The author emphasizes the need for constructing indicators for change that take into account structures, processes, relationships as well as the contribution of subjectivity of all participants. The conclusion is that internal validity of this kind of indicators can be achieved when the categories employed in this construction are theoretically and contextually founded and contain mutual meanings shared between the participants.

Recebido em: 11/12/2008

Reencaminhado em: 08/01/2009

Aprovado em: 09/01/2009

INTRODUÇÃO

O objetivo deste ensaio é apresentar uma reflexão teórico-metodológica focada na construção de indicadores qualitativos, voltados principalmente para avaliação de mudanças. Foi escrito para subsidiar o processo de autoavaliação de um conjunto de escolas médicas que estão implementando um programa de transformação de seus métodos tradicionais para outros fundamentados na filosofia e na prática de aprendizagem ativa. Nesse sentido, este artigo faz parte de um conjunto de textos que estão sendo produzidos no interior do projeto conduzido pela Comissão de Avaliação do Ensino Médico (Caem).

São três os termos principais que compõem o campo semântico deste artigo: construção de indicadores, mudança e avaliação. A ênfase será dada ao primeiro termo, e os outros dois constituem o pano de fundo da reflexão e por isso apenas serão mencionados oportunamente.

Embora o conceito de mudança seja um dos mais fortes e complexos tratados pela sociologia¹⁻⁵ neste estudo, seu sentido será resumido como o impacto de determinadas intervenções sociais ou pedagógicas sobre instituições e atores: as que são produzidas rapidamente e às vezes em cadeias sequenciais; as que procedem de uma vontade deliberada; as que ocorrem por influxo de tecnologias materiais e sociais; e as que são cumulativas e afetam muitos indivíduos ou aspectos cruciais da sociedade⁶.

Igualmente, quando falo de avaliação neste texto, estou me referindo ao conceito de investigação avaliativa que exige determinados cuidados acadêmicos como fundamentação teórica e contextualização ao lado dos processos normativos tradicionais, tendo como base as reflexões de Contandriopoulos *et al.*⁷

Neste texto serão tratados os seguintes temas: o conceito de indicador; construção de indicadores de avaliação; o que são e como construir indicadores qualitativos.

Conceito de indicador

A definição do termo “indicador”, do ponto de vista científico, varia pouco de um autor para outro. Em geral, os pesquisadores consideram que os indicadores constituem parâmetros quantificados ou qualitativos que servem para detalhar se os objetivos de uma proposta estão sendo bem conduzidos (avaliação de processo) ou foram alcançados (avaliação de resultados). Como uma espécie de sinalizadores da realidade, a maioria dos indicadores dá ênfase ao sentido de medida e balizamento de processos de construção da realidade ou de elaboração de investigações avaliativas.

Instituições públicas – como o Instituto Brasileiro de Geografia e Estatística (IBGE), o Ministério da Educação, por meio

de seus institutos como a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e o Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), e o Ministério da Ciência e Tecnologia, por meio da Financiadora de Estudos e Projetos (Finep) e do Conselho Nacional de Pesquisa (CNPq), entre outros – utilizam maciçamente indicadores considerando seu sentido como especificações quantitativas e qualitativas para medir o alcance de determinados objetivos, metas e resultados.

Do ponto de vista internacional, a Organização para a Cooperação e o Desenvolvimento Econômico (OCDE)⁸ possui várias publicações sobre indicadores e os considera como séries de dados, definidos para responder perguntas sobre determinado fenômeno ou sistema específico. Igualmente a *International Organization for Standardization* (ISO), que avalia a qualidade de determinados processos e de ações de empresas, estabelece que indicadores são expressões numéricas, simbólicas ou verbais, empregadas para caracterizar atividades ou eventos, em termos quantitativos e qualitativos, com o objetivo de determinar seu valor⁹.

Autores e instituições coincidem em que indicadores podem ser construídos para medir ou revelar aspectos relacionados a diversos planos em observação: níveis individuais, coletivos, associativos, políticos, econômicos e culturais, entre outros. Podem, por exemplo, ser instrumentos para mensurar a disponibilização de bens e atividades, assim como para conceber parâmetros de acesso de diferentes atores a um programa, a relevância que ele possui para a vida de cada um, sua intensidade e seu sentido.

Do ponto de vista de sua utilidade, além de ser um dispositivo para medição, para o estabelecimento de parâmetros e para avaliação, os indicadores são importantes instrumentos de gestão, pois permitem ao administrador operar sobre dimensões-chave de sistemas e de processos, monitorando situações que devem ser mudadas, incentivadas ou potencializadas desde o início de uma intervenção até o alcance do que foi pretendido e previsto como resultado.

Mas a utilidade de um bom indicador depende de algumas condições: (a) que estejam normalizados e que sua produção histórica (sua temporalidade) se atenha sempre à mesma especificação ou forma de medida, permitindo a comparabilidade; (b) que sejam produzidos com regularidade, visando à formação de séries temporais e permitindo visualizar as tendências dos dados no tempo; (c) que sejam pactuados por quem (grupos, instituições) os utiliza e quem pretende estabelecer comparabilidade no âmbito nacional e até internacional, como é feito, por exemplo, pela Capes e pelo CNPq na avaliação de cursos, de projetos e de periódicos científicos; (d) e

que estejam disponíveis para um público amplo e de forma acessível, propiciando à opinião pública um formato simples de acompanhamento do desempenho de instituições e de políticas públicas ou que recebam financiamento público.

Para poder eleger indicadores, devemos ter uma concepção precisa das organizações e sistemas que queremos gerenciar ou transformar, fato este que pode apresentar diferentes níveis de facilidade ou complexidade. Portanto, os vários tipos de indicadores existem para que possamos efetuar adequações nos objetivos, nas metas e até na missão de uma organização, permitindo ao gestor ou avaliador acompanhar todo o processo. Os indicadores assinalam tendências. No entanto, nenhum indicador pode aportar certeza absoluta quanto aos resultados de uma ação ou de um processo, pois sua função é apenas ser um sinalizador: indicadores são instrumentos, não operam por si mesmos, indicam o que devem indicar.

Quando observados do ponto de vista quantitativo e qualitativo, os indicadores se referem aos aspectos tangíveis e intangíveis da realidade. Tangíveis são os elementos facilmente observáveis, como renda, escolaridade, forma de organização e gestão, legislação, mecanismos de divulgação. Intangíveis são os atributos que só podem ser captados indiretamente por meio de suas formas de manifestação, como é o caso do incremento da consciência social, da autoestima, de valores, de atitudes, de liderança, de protagonismo e de cidadania. “Como essas são dimensões complexas da realidade, processos não-lineares ou progressivos, demandam um conjunto de indicadores que apreendam algumas de suas manifestações indiretas, cercando a complexidade do que pretendemos observar”¹⁰ (p. 4).

Construção de indicadores para avaliação

A escolha de indicadores para avaliação deve ser feita em função dos aspectos que queremos analisar – por exemplo, se queremos medir ou acompanhar a eficiência, a eficácia, a efetividade ou o impacto a partir de determinada política organizacional. Eficiência diz respeito à boa utilização de recursos financeiros, materiais e humanos em relação às atividades e resultados alcançados. Eficácia se refere à relação entre as ações realizadas e os resultados obtidos. Efetividade é a observação da incorporação das mudanças geradas por determinado programa na realidade da população-alvo. Impacto corresponde ao poder de influência e de irradiação de um projeto realizado^{11,12}.

Uma pergunta fundamental na investigação avaliativa é saber qual é a combinação mais adequada entre indicadores quantitativos e qualitativos, tangíveis e intangíveis, diretos e indiretos. Valarelli¹⁰ recomenda que, no caso de avaliação de

projetos sociais e de projetos educacionais, trabalhem sempre com um sistema de indicadores, levando em conta:

- Conceções, interesses e enfoques das organizações e atores envolvidos – Um bom sistema de indicadores para avaliação deve sempre surgir do processo de diálogo e de negociação entre todos os atores. Por exemplo, numa escola de Medicina, os gestores responsáveis pela instituição poderiam estar interessados numa avaliação que enfatizasse a participação de determinados atores envolvidos no ensino e aprendizagem; os professores gostariam de saber sobre a qualidade da formação oferecida; e o MEC (por hipótese), como financiador, gostaria de ter conhecimento sobre a eficácia e a efetividade de uma proposta de mudança que estivesse sendo implementada e que ele estivesse financiando.
- O contexto – Cada investigação avaliativa deverá construir seu sistema de indicadores específicos. Pois, mesmo quando um conjunto de variáveis pode ser usado para vários projetos, os indicadores devem sinalizar condições específicas da realidade em avaliação. Por exemplo, a quantidade e a qualidade de informações preexistentes influenciam a necessidade de produzir novas informações, assim como o nível de aprofundamento de possíveis análises. Por isso, a maioria dos projetos avaliativos deve se iniciar com uma análise situacional, como ponto de partida para a construção de indicadores.
- O modo de gestão – A forma de organização de uma proposta avaliativa deve contemplar um projeto claro, com objetivos e metas, conceitos básicos, conexão entre os diferentes investigadores e outros atores e desenhos metodológicos em seu sentido mais amplo (frequentemente exigindo abordagens quantitativas e qualitativas). Nesse espaço, os indicadores são os bons sintetizadores e sinalizadores do caminho. No entanto, eles não falam por si mesmos, precisam estar articulados com as diferentes atividades que estão sendo avaliadas e servir de parâmetro para a tomada de decisões.
- Os recursos – O tempo, a duração do projeto de avaliação e os recursos disponíveis também interferem na possibilidade de utilização de determinados indicadores. Por exemplo, um indicador de impacto só pode mostrar seu sentido em investigações de longo prazo. Nossa experiência nesse sentido é que os processos sociais e educacionais frequentemente oferecem aos gestores dados importantes que são como termôme-

tros que medem os avanços ou retrocessos dos processos de intervenção. Por isso, há muitas mudanças que precisam ser acompanhadas com critérios qualitativos, pois estão em processo de acontecer. No entanto, frequentemente não é possível medir seu impacto em termos quantitativos. Por exemplo, no caso da avaliação de mudanças do ensino médico tradicional para o que utiliza métodos de aprendizagem ativa, não bastaria saber quantos alunos e quantos médicos foram incluídos ou aderiram às atividades profissionais e técnicas do programa Saúde da Família ou mesmo a determinado treinamento para tratar uma enfermidade específica em nível secundário ou terciário. É importante saber, também, como esse processo está sendo encaminhado, que nível de diálogo e de reflexão existe entre esses atores e como os estudantes estão se apropriando do conhecimento, das dúvidas e das críticas que toda interação humana compartilhada produz. E, principalmente, seria preciso criar indicadores que respondessem sobre em que medida os processos de mudança iniciados estão sendo institucionalizados, deixando para trás práticas arcaicas. Nesse particular, a sociologia nos diz que a questão mais problemática de qualquer mudança é o seu processo de incorporação às rotinas institucionais.

Construção de Indicadores Qualitativos

Esforços diversificados vêm sendo feitos para aprimorar o desenvolvimento de indicadores no âmbito da abordagem qualitativa. A tentativa de inserção de elementos subjetivos, tendo em vista as teorizações que mostram a importância do envolvimento dos diferentes atores na produção do processo e dos resultados, vem acompanhando o próprio desenvolvimento desse conceito científico, cuja finalidade primeira é aumentar a validade dos parâmetros de análise. Há pelo menos duas linhas de abordagem nessa forma de construção. Uma é proveniente da lógica quantitativa, e outra é especificamente marcada pela fundamentação hermenêutica.

Construção de indicadores qualitativos por estratégias quantitativas – A busca de mensurar valores, opiniões, relações e vivências intersubjetivas é bastante antiga no campo das Ciências Sociais. Desde 1920, diversas escalas têm sido criadas. Assis *et al.*¹¹ citam as seguintes:

- A de distância social, de Bogardus¹³, que solicita ao entrevistado para estabelecer uma hierarquia de intimidade e convivência social que ele estaria disposto a ceder para pessoas de outras nacionalidades, raças ou etnias.

- A de Thurstone¹⁴, que se baseia em questionários do tipo “concordo” ou “discordo” de certas afirmações cujos resultados serão submetidos a especialistas no tema estudado e que definirão as tendências comportamentais, das mais radicais e neutras às mais conservadoras.
- A de Lickert¹⁵, que estabelece o mesmo procedimento da escala de Thurstone, mas substitui o grupo de especialistas por um grupo social com as mesmas características.
- Os escalogramas de Guttman¹⁶, que seguem a mesma lógica da escala de Bogardus e estudam as concordâncias entre as respostas, criando-se um perfil de comportamento em relação à tolerância e aceitação de outros grupos étnicos.
- E a proposta sociométrica de Moreno¹⁷, capaz de estabelecer um sociograma, isto é, uma representação diagramática dos sentimentos de atração ou repulsa entre pessoas de determinado grupo.

A maioria desses instrumentos foi produzida entre os anos 1920 e 1930. Todos os seus organizadores atribuíam valor numérico ou ordenação percentual a respostas sobre atitudes e comportamentos frente a situações reais ou hipotéticas¹⁸. Tais dispositivos seguiam os mesmos modelos de análise qualitativa desse período histórico, os quais são intensivamente analisados por Lazarsfeld & Baron¹⁹, evidenciando a lógica do pensamento estatístico que os preside em suas preocupações epistemológicas e metodológicas: (a) construção de amostras reunidas de maneira sistemática; (b) busca de validação dos procedimentos de coleta de dados e de resultados; (c) criação de codificadores para medir regularidades; (d) utilização de análise de frequência como critério de objetividade e de cientificidade; (e) construção de dispositivos para medir validade, fidedignidade e produtividade da análise⁶.

Um exemplo bastante conhecido de avaliação qualitativa pela lógica quantitativista é a que analisa programas e serviços por meio do chamado “indicador de satisfação do usuário”. Nesse dispositivo, os pesquisadores buscam expressar um *quantum* de contentamento sobre relações travadas com os profissionais, informações recebidas e eficácia ou competência da ação dos profissionais para atender às necessidades do cliente.

Convencionou-se denominar “mensuração de atitudes” ou “escalas de atitudes” a referida linha de investigação avaliativa qualitativa que utiliza instrumentos e resultados típicos da abordagem quantitativa. Apesar de serem uma fórmula muito antiga de estudo de percepções e subjetividade, as escalas que quantificam atitudes e comportamentos continuam

por reduzir a questão da singularidade e da individualidade a uma série de proposições estabelecidas pelo pesquisador. Aparelhos considerados objetivos, em geral os indicadores estabelecidos em escalas permanecem externos aos sujeitos investigados. Por exemplo, uma escala de “satisfação do usuário de saúde” é um artefato fechado que não leva em conta o grau de consciência de direitos por parte dos clientes. Usuários sem acesso a níveis básicos de cidadania, por exemplo, tendem a assinalar maior grau de satisfação nas respostas aos questionários, pelo simples fato de serem atendidos ou de serem ouvidos por uma assistente social, uma diretora ou um médico. No mesmo sentido, Williams²⁰ ressalta que a quantificação do qualitativo sob a forma de *surveys* é empobrecedora, pois não leva em consideração os critérios dos usuários quando julgam a atenção recebida, como, por exemplo, seus problemas socio-culturais, de gênero e étnicos, entre outros.

Não pretendo aqui fazer uma crítica destrutiva a respeito das escalas, mas apenas ressaltar que o tipo de informação que podem gerar está vinculado à filosofia que as fundamenta. Sua limitação precisa ser levada em conta por pesquisadores das áreas de psicologia, sociologia, da área de epidemiologia e da avaliação que fazem rotineiro uso delas para enfatizar percepções, sentimentos e reações dos indivíduos em relação a atributos pessoais, como autoestima e depressão; ou sociais, como é caso da avaliação de qualidade de vida. Considero que haveria um enriquecimento muito grande desses pesquisadores se buscassem se aproximar do universo cultural dos atores que avaliam.

Construção de indicadores qualitativos por estratégias qualitativas – Chamo de indicadores qualitativos propriamente ditos os que expressam a voz, os sentimentos, os pensamentos e as práticas dos diversos atores que compõem o universo de uma pesquisa ou de uma avaliação⁶. Eles evidenciam a adoção ou a rejeição de certas atitudes, valores, estilos de comportamento e de consciência¹⁰ e se fundamentam na necessidade de ressaltar as dimensões das relações vividas intersubjetivamente, entendendo que elas fazem parte de qualquer processo social e o influenciam.

Para construir indicadores qualitativos, é importante assumir sua natureza hermenêutica²¹⁻²³. Ou seja:

- A origem da produção de indicadores deve ser a realidade empírica e concreta (no caso, a dinâmica das escolas médicas). Os indicadores não devem ser colocados como um produto a ser provado pela realidade e, sim, a partir da realidade social para a construção do sistema de indicadores de saúde.
- É preciso entender que os atores sociais, ao se comunicarem sobre qualquer assunto do mundo da vida,

o interpretam e julgam a partir de um estoque de referências, como classe, gênero, grupo de pertinência, etnia, idade e função social, entre outros²⁴.

- Igualmente, os pesquisadores que analisam tais interpretações (como construto de segunda ordem, como diria Shutz²⁴) devem se colocar dentro do processo de construção da realidade, buscando valorizar a relação entre observador e observado que convivem no mesmo tempo histórico e estão marcados por diversos fatores sociais comuns²⁵.
- Geralmente não vamos encontrar listagens de indicadores qualitativos, pois, diferentemente das escalas de mensuração, esses são construídos com os próprios atores que fazem parte do estudo avaliativo e, a não ser em casos culturais mais ou menos homogêneos, não podem e não devem ser replicados.

As técnicas mais comuns para construção de indicadores qualitativos são: Grupos Focais, *Delphi* e Grupos Nominais. Mas muitas outras, como a técnica de narrativa, costumam ser utilizadas, e neste texto apresento um exemplo a respeito.

Grupos Focais – Por grupos focais, como o próprio nome indica, entendemos um dispositivo para aprofundar o significado de determinado tema. Vários autores coincidem em que grupos focais como uma técnica de pesquisa qualitativa derivada das entrevistas grupais são organizados para coletar informações por meio das interações frente a um tópico específico sugerido por um pesquisador que é ao mesmo tempo coordenador ou moderador do grupo.

Relegada a segundo plano durante décadas, a partir de 1980 a técnica dos grupos focais passou a ocupar um lugar fundamental em investigações qualitativas no campo social, cultural, psicológico, organizacional, educacional e da avaliação. O foco da reflexão é aprofundado pelas opiniões surgidas no próprio grupo, a partir do jogo de influências mútuas que emergem entre os participantes e se desenvolvem no contexto das interações. Ou seja, o grupo permite manifestar divergências, explicitar e defender pontos de vista, mas também mudar de opinião. Geralmente, quando utilizados para a construção de indicadores, os grupos focais são encaminhados para consensualizar temas-chave que posteriormente são analisados pelo investigador e transformados em dispositivos balizadores da realidade.

A organização de um grupo focal segue mais ou menos a seguinte orientação: o coordenador, também chamado moderador, esclarece aos participantes o motivo da pesquisa e qual o tema-chave que pretende colocar em discussão. A seu lado deve estar um relator que toma nota de todos os detalhes da

reunião, a grava ou a filma (com o consentimento dos participantes), deixando o coordenador com tempo e tranquilidade suficientes para ouvir e encaminhar as discussões.

Um bom grupo focal não deve ter mais que seis a oito pessoas, para que todos tenham possibilidade de se expressar, debater e defender suas idéias. As regras básicas de funcionamento dos grupos focais, a serem combinadas de antemão com os participantes, são: (1) solicitar que apenas uma pessoa fale de cada vez; (2) evitar discussões paralelas, para que todos possam participar; (3) pedir que cada um fale livremente o que pensa; (4) evitar que alguns participantes dominem a discussão e (5) manter a atenção de todos voltada para a temática em questão. Além de garantir essas regras, o coordenador deve estar atento para que o grupo não se disperse, introduzindo perguntas que possam aprofundar o assunto e esclarecê-lo o mais adequadamente possível.

Técnica Delphi – Essa técnica pode ser caracterizada como uma estratégia para estruturar o aperfeiçoamento de instrumentos de pesquisa ou dos resultados das análises. Geralmente, para participar, são convocados pesquisadores ou especialistas no assunto. Muito utilizada nas pesquisas avaliativas, essa técnica pode apresentar-se de várias formas. É o caso de um grupo de especialistas que se reúne para ler, criticar e sugerir sobre determinado material anteriormente preparado, com a finalidade de refinar, adequar e dar qualidade ao que foi construído por um pequeno grupo ou por uma pessoa.

O uso mais comum da técnica *Delphi*, no entanto, tem sido mais individualizado. Um grupo de pesquisa ou um pesquisador elabora um texto, contendo um conjunto de indicadores que julga serem fundamentais para determinada avaliação e compartilha sua versão preliminar com os melhores especialistas sobre o assunto. O demandante pede a colaboração para leitura e crítica, dando aos colaboradores um prazo para correção, acréscimos e supressões. Ao receber as respostas, cabe ao pesquisador ou ao grupo de investigação a incorporação das sugestões e a síntese das contribuições. Essa técnica tem um valor fundamental para validação tanto de instrumentos como de resultados de investigação. A base da técnica *Delphi* é o reconhecimento da superioridade do juízo de pessoas experientes em relação ao juízo individual ou de um pequeno grupo. É preciso ressaltar que, quando falamos aqui de especialistas, não estamos nos referindo necessariamente a títulos. Por exemplo, um médico do serviço de atenção básica dedicado e experiente em relação à orientação dos alunos pode ser um excelente colaborador para pensar indicadores de avaliação dos “cenários da prática”, junto com os pesquisadores.

Grupo Nominal – A técnica de grupo nominal²⁶ é bastante usada em processos de tomada de decisão operacional e

também para construção de indicadores. A expressão *grupo nominal* é utilizada para designar o fato de que cada pessoa convidada a participar de uma discussão mediada por essa técnica trabalhará, inicialmente, com um *nome* ou com uma *palavra* na qual colocará toda a força de sua idéia a respeito do tema em pauta. Essa é a parte mais importante do trabalho, que em seguida será socializado dentro do grupo, que não deve ultrapassar dez pessoas.

A condução do grupo nominal poderia ocorrer da seguinte forma: a coordenação distribui a cada participante uma caneta e uma folha de papel, solicitando que responda a determinada pergunta por escrito. Por exemplo: na construção de indicadores de avaliação das escolas médicas, que tema você considera que definitivamente não pode faltar? Coloque-o em apenas uma palavra. São concedidos mais ou menos três minutos para a resposta. Uma vez que, a partir da demanda do investigador, cada pessoa escreveu a palavra que sintetiza sua idéia, inicia-se a segunda etapa: sob a coordenação do pesquisador, cada um vai dizendo a sua palavra, que é imediatamente escrita num quadro ou projetada em *datashow*. Num terceiro momento, abre-se tempo suficiente para que cada um aprofunde o sentido da palavra que escolheu. A experiência mostra que, geralmente, ocorrem repetições de palavras ou de sentidos, o que é muito bom porque demonstra que existem consensos em relação aos temas-chave. Aos poucos, o coordenador vai conduzindo o grupo para resumir seu pensamento em torno de um número pequeno de temas.

É preciso ficar claro que um grupo nominal ou focal ou a técnica *Delphi* não têm como obrigação construir indicadores. Essa é uma tarefa teórica do pesquisador ou do grupo de pesquisa, que, no entanto, pode ter seu trabalho altamente potencializado se ouvir os diferentes atores envolvidos no processo ou especialistas que têm experiência na área. O que os participantes oferecem são temas cruciais que precisam ser transformados em indicadores.

Indicadores qualitativos, portanto, devem ser construídos de forma participativa e considerados como balizas avaliativas, que permitem mapear com mais profundidade a natureza das mudanças ocorridas e em processo²⁷. Reforçando esse princípio, Demo²⁸ ressalta que a participação dos sujeitos é essencial para a construção de indicadores, pois a participação política no cotidiano é o cerne da dimensão humana de qualidade, da capacidade de se autogerir, da criatividade que desenha caminhos futuros, da autodeterminação e da autopromoção dos sujeitos.

Existem muitos investimentos na construção de indicadores analisáveis por métodos tipicamente qualitativos. Darei dois exemplos nesse sentido.

O primeiro, bastante elucidativo, fala da construção de indicadores qualitativos para avaliação do programa *Family Resource Programs*, do Canadá, que mantém 2 mil projetos comunitários. Os avaliadores criaram indicadores utilizando narrativas das famílias assistidas²⁹ a partir de histórias individuais. Na teorização de suas técnicas, os pesquisadores apresentam o movimento em que falas individuais são agrupadas em códigos, subcódigos e temas de análise, até que sejam definidos os melhores indicadores a serem testados em protocolos de investigação operacionais voltados para avaliação dos serviços. O exemplo a seguir mostra a construção de um indicador baseado em falas como a dessa mulher, cujo depoimento é citado abaixo³⁰.

O programa ajudou a que ele ficasse um pouco mais independente e um pouco mais acostumado a lidar com outras crianças. Isso nos deu uma chance para entender que algumas das coisas vividas por ele iam ser motivo para que nós fôssemos capazes de ajudá-lo e estimular suas habilidades... Por exemplo... lidar com níveis altos de ruído e com muitas crianças, e ensinar que, em alguns momentos em que precisar, ele pode achar um canto quieto, se sentar e depois se reunir de novo com os outros. Ele agora sabe que tem que fazer isso. A situação também nos permitiu poder falar para os professores que, se começarem a ver certos comportamentos dele, o levem para um canto e deem alguns minutos para ele ficar quieto (Lily).

Um código foi extraído dessa fala: “benefício para a criança”. Igualmente foi extraído um subcódigo: “acesso a recursos para cuidar dos filhos”; e alguns temas fortes: “as práticas e experiências de engajar famílias” e “de empoderar participantes”. A partir do depoimento de Lily, os autores elaboraram dois indicadores: (1) meu filho está pronto para a escola por causa do Programa Comunitário; (2) eu sou capaz de pedir aos professores para ajudarem meu filho.

Os dois indicadores criados, segundo seus autores, serviram de base para a construção de roteiros testados nos serviços de todo o Canadá, visando implementar tecnologias (*soft*) de avaliação em nível nacional, baseadas na experiência dos participantes.

Outra tentativa importante é a experiência avaliativa do Centro Latino-Americano de Estudos sobre Violência e Saúde (Claves), densamente narrada no livro *Avaliação por triangulação de métodos*⁶. Essa obra condensa as experiências teóricas do grupo nos últimos anos^{6,25,31,32}, ao mesmo tempo em que mostra como, na prática, a avaliação foi realizada. Geralmente, nas pesquisas citadas, todos os atores interessa-

dos no processo são chamados a participar na construção de indicadores. Com eles são trabalhadas diversas técnicas para afloração dos temas relevantes, sobretudo as de grupo focal e de grupos nominais. A experiência do Claves mostra que quanto mais atores participam na construção dos instrumentos que balizam as avaliações, mais é possível ter legitimidade e consenso na apresentação dos resultados parciais e finais do trabalho.

Validade interna e externa dos indicadores qualitativos – Por se tratar de uma abordagem de compreensão da experiência a partir tanto do aporte de especialistas como do senso comum e do julgamento dos atores, uma das questões críticas que se coloca frequentemente ao investigador é o grau de confiabilidade e validade dos indicadores qualitativos. Para alguns autores, não cabe falar em confiabilidade, pois as percepções e as relações sociais, objetos desse tipo de operação, são dinâmicas e impossíveis de serem repetidas em sua integralidade, além de serem fruto do caminho interpretativo, eminentemente intersubjetivo.

Minayo²³, em *O desafio do conhecimento*, mostra que o maior consenso hoje existente entre os estudiosos sobre a confiabilidade e a validade dos dados qualitativos é o que se forma por meio da intersubjetividade, pelo julgamento dos pares e pelo reconhecimento dos participantes sobre o sentido e a veracidade da análise. Gadamer²¹ fala exatamente o mesmo quando trata da validade na relação intrínseca entre investigador e investigado (avaliador-avaliado). Em suas palavras, o estudo é válido quando os investigados compartilham com o investigador a veracidade e o sentido dos resultados.

Por isso, como lembram Minayo²³ e Demo²⁸, é preciso investir na objetivação, por meio de estratégias múltiplas. Entre estas, destacamos: (a) o cuidado permanente e compartilhado de discussão e troca entre os pares sobre a pertinência de determinados indicadores e procedimentos; (b) O rigor teórico-metodológico em relação à definição do objeto, dos pressupostos ou hipóteses, dos instrumentos conceituais e de campo e da análise do material recolhido. Demo²⁸ lembra, porém, que os critérios de coerência, consistência, originalidade e objetivação são o marco da crítica interna de qualquer investigação científica e não apenas do processo de produção da pesquisa qualitativa; (c) A triangulação como esforço de comunicação entre pessoas, conceitos, abordagens e elaboração de resultados. A triangulação, assim pensada e praticada, oferece mais confiabilidade aos dados e às análises.

Especificamente em relação à confiabilidade externa, em que diferentes pesquisadores deveriam encontrar os mesmos resultados investigando o mesmo tema, é necessário ter em

conta que as investigações qualitativas não buscam homogeneidades, mas, sim, diferenciações e especificidades. A elas não se aplica o critério quantitativo de comparabilidade, cuja raiz epistemológica é a observação das regularidades nos fenômenos sociais, típica do método positivista²³. Porém, nada impede que existam indicadores qualitativos capazes de expressar tanto expressões coletivas como específicas e diferenciadas.

Em outras palavras, a validade externa pretendida dentro da perspectiva qualitativa é a que se refere à possibilidade de geração de conhecimentos que contribuam para o aprofundamento conceitual ou analítico, por meio da elaboração de tipificações ou de lógicas culturais. Esses achados permitem a melhor compreensão do grupo estudado e o refinamento de categorias empíricas que podem ser utilizadas em estudos de outros grupos sociais e de instrumentos padronizados²¹, como é a pretensão do esforço canadense tratado acima.

Em relação ao incremento da confiabilidade interna, em que diferentes pesquisadores envolvidos no mesmo estudo concordam entre si em suas análises, privilegiamos algumas iniciativas: (a) constante troca de impressões e informações entre os pesquisadores, visando a interpretações consensuais; (b) checagem da adequação das impressões obtidas no trabalho de campo com os atores, interlocutores e participantes do processo de pesquisa; (c) descrição de situações e relatos verbais sem muitas interpretações *a priori* do pesquisador; (d) análise em que se busque a lógica interna e hermenêutica do grupo²³; (e) exame do produto final da pesquisa pelos mesmos pares e atores que ajudaram a construir os instrumentos^{6,27,33}. No entanto, lembra Popper³⁴ em relação a qualquer tipo de estudo, seja ele qualitativo ou quantitativo, os critérios tradicionais de validade do conhecimento precisam ser relativizados, pois nenhuma hipótese é verificável, uma vez que a acumulação de casos afirmativos não pode garantir uma generalização teórica.

CONSIDERAÇÕES FINAIS

Em síntese, a validade dos estudos de avaliação qualitativa é concebida não como um dispositivo que espelha a realidade e sim como uma “produção reflexiva”, em que o observador é parte e parcela do contexto e da cultura que busca entender e representar. Mais validade interna de uma pesquisa qualitativa pode ser obtida quando as categorias conceituais empregadas têm significado mútuo e partilhado entre os participantes e o pesquisador. Para tanto, alguns fatores são fundamentais: *tempo*, pois, quando o processo e as mudanças são focalizados, o fenômeno observado pode não se manter constante; variações nos informantes selecionados, que podem agir de forma diferente e demonstrar interpretações diferenciadas sobre os vários aspectos em questão; diversidade de informantes, o que

amplia a variedade de informações; ocorrência de eventos-limite que podem mudar o convívio e o cotidiano do contexto sob observação.

REFERÊNCIAS

1. Nisbet RA. Social Change and History. New York: New York University Press;1969.
2. Rocher G. Le changement social. Paris: Éditions de Seuil; 1970.
3. Tourraine A. Crítica de La Modernidad. México: Fondo de Cultura Económica; 1995.
4. Terrén E. Posmodernidad, legitimidad y educación. Educación & Sociedad. 20(67):10-26; 1999.
5. Atlan H. Entre o cristal e a fumaça: ensaio sobre a organização dos seres vivos. Rio de Janeiro: Zahar; 1992.
6. Minayo MCS, Assis SG, Souza ER, (Org.). Avaliação por triangulação de métodos: abordagem de programas sociais. Rio de Janeiro: Ed. Fiocruz; 2005.
7. Contandropoulos AP, Champagne F, Denis JL, Pineault T. Avaliação na área de saúde: conceitos e métodos. In: Hartz, ZMA (Org.). Avaliação em saúde: dos modelos conceituais a práticas da implantação de programas. Rio de Janeiro: Ed. Fiocruz; 1997. p.28-48.
8. Organização para a Cooperação e o Desenvolvimento Econômico. Knowledge and Skills for Life. Paris: OECD; 2001.
9. The International Standard Organization. The ISO Standards Glossary/ISO; 1998.
10. Valarelli LL. Indicadores de resultados de projetos sociais. 2004. [online]. [Acesso em: 16 jun.2004]. Disponível em: <http://www.rits.org.br/gestão>
11. Assis SG, Deslandes SF, Minayo MCS & Santos NC. Definição de objetivos e construção de indicadores visando à triangulação. In: Minayo MCS, Assis SG, Souza ER (Org.). Avaliação por triangulação de métodos: abordagem de programas sociais. Rio de Janeiro: Ed. Fiocruz. p.105-131.
12. Aguilar MJ, Anger-Egg E. Avaliação de serviços e programas sociais. Petrópolis: Vozes; 1994.
13. Bogardus E. Distância social na cidade. In: Pierson D (Org.). Estudo de Organização Social. São Paulo; 1970. p.38-52.
14. Thurstone LL. Attitudes can be measured. American Journal of Sociology. 1928;33: 529-44.
15. Likert R. A Technique for the measurerement of attitudes. Archives of Psychology. 1932;140:1-55.
16. Guttman L. The problem of attitude and opinion measurement: the basis for scalograma analysis. Measurement and Prediction. Princeton:Princeton University Press; 1950.

17. Moreno JL. *Fondaments de la sociométrie*. Paris: Presses Universitaires de France; 1954.
18. Osgood C.E. *The measurement of meaning*. Illinois: University of Illinois; 1957.
19. Lazarsfeld P, Barton AH. *Qualitative measurement in the social sciences: classification, typologies and indices*. In: Lerner D, Lasswell HD (Ed.). *The policy sciences*. Stanford: Stanford University Press; 1951. p.155-92
20. Williams B. Patient satisfaction: a valid concept? *Soc. Sci. Med.* 1994;38(4):509-516.
21. Gadamer H. *Verdade e método*. Petrópolis: Vozes; 1999.
22. Minayo MCS. *Hermenêutica-Dialética como caminho do pensamento social*. In: Minayo MCS, Deslandes SF (Org.). *Caminhos do Pensamento*. Rio de Janeiro: Ed. a Fiocruz; 2002. p. 83-108.
23. Minayo MCS. *O desafio do conhecimento*. 11^a. ed. São Paulo: Hucitec; 2008.
24. Shutz AA. *Common sense and Scientific Interpretations of Human Actions*. In: *Collected Papers I*, 2^o ed. Hague: Martinus Niphoff; 1982.
25. Minayo MCS, Deslandes SF. *Análise diagnóstica da Política Nacional de Redução da Morbimortalidade por acidentes e violências*. Rio de Janeiro: Ed. Fiocruz; 2007.
26. Schermerhorn JR, Hunt JG, Osborn RN.(Org.). *Fundamentos de comportamento organizacional*. Porto Alegre: Bookman; 1999.
27. Deslandes SF, Assis SG. *Abordagem quantitativa e qualitativa em saúde: o diálogo das diferenças*. In: Minayo MCS, Deslandes SF (Org.). *Caminhos do Pensamento. Epistemologia e Método*. Rio de Janeiro: Ed. Fiocruz; 2002. p.195-226.
28. Demo P. *Avaliação qualitativa*. São Paulo: Cortez; 2002.
29. *Family Resources Programs. Participant's voices*. [página eletrônica]. Toronto: Ryerson University. [Acesso em: 02 dez.2008]. Disponível em: <http://www.ryerson.ca/voices>
30. Silver S, Wilson S, Berman R, Salazar CC, Soklaridis S, Karim A. *Program. For Narratives to Likert Scales. the long and winding road of developing evaluation indicators from qualitative data*. In: *The Tenth International Qualitative Health Research Conference*. April 30-May 4, 2004. Alberta: University of Alberta; 2004. p.39.
31. Minayo MCS, Souza ER. *Missão Investigar: entre o ideal e a realidade de ser policial*. Rio de Janeiro: Ed. Garamond; 2003.
32. Minayo MCS, Souza ER, Constantino P. *Missão Prevenir e Proteger. Condições de trabalho, saúde e qualidade de vida dos policiais militares do Rio de Janeiro*. Rio de Janeiro: Ed. Fiocruz; 2008.
33. Le Compte MD. *Problems of reliability and validity in ethnographic research*. *Review of Educational Research*. 1982;52:31-60.
34. Popper K. *La logique de la découverte scientifique*. Paris: Éditions Payot; 1973.

CONFLITOS DE INTERESSE

Declarou não haver

ENDEREÇO PARA CORRESPONDÊNCIA

Maria Cecília de Souza Minayo
Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública, Sérgio Arouca Claves.
Av. Brasil, 4036 – sala 700 – Mangueiras
CEP. 21040-361 – Rio de Janeiro – RJ
E-mail: cecilia@claves.fiocruz.br