

UNIVERSIDADE ESTADUAL DE CAMPINAS
PÓS-GRADUAÇÃO EM SAÚDE COLETIVA - MESTRADO

**A TEMÁTICA DA AVALIAÇÃO NO CAMPO DA SAÚDE
COLETIVA:
UMA BIBLIOGRAFIA COMENTADA**

ANA CECÍLIA BASTOS STENZEL

Campinas, junho de 1996

ANA CECÍLIA BASTOS STENZEL

**A TEMÁTICA DA AVALIAÇÃO NO CAMPO DA SAÚDE
COLETIVA:
UMA BIBLIOGRAFIA COMENTADA**

DISSERTAÇÃO APRESENTADA AO CURSO DE PÓS-GRADUAÇÃO EM SAÚDE COLETIVA, DO DEPARTAMENTO DE MEDICINA PREVENTIVA E SOCIAL DA FACULDADE DE CIÊNCIAS MÉDICAS, DA UNIVERSIDADE ESTADUAL DE CAMPINAS, PARA OBTENÇÃO DO TÍTULO DE MESTRE EM SAÚDE COLETIVA.

ORIENTADOR: PROF^o DR. NELSON RODRIGUES DOS SANTOS

CO-ORIENTADOR: PROF^o DR. HELENO RODRIGUES CORRÊA FILHO

Campinas, junho de 1996

Ao meu pai e à minha mãe.

A cada um de meus amados irmãos: Maria
Emília, Maria do Carmo, Daniel, Juliana e
Raquel.

À Ruth, amiga, companheira e exemplo, nesta
“aventura” do mestrado.

Ao Heitor, pelo carinho e apoio.

Ao Dr. Miguel, parteiro deste “filho”.

AGRADECIMENTOS

Aos antigos amigos de Porto Alegre, que me estimularam a sair para estudar: Pedro, Aloísio, Salete, Carminha, Luíza.

Aos colegas e amigos da Divisão de Municipalização: Neide, Pressoto, Bernadete, Márcia, Nadir.

Aos novos amigos de Campinas, que me acolheram com tanto carinho: Dário, Angélica, Sheila, Erly, Iguti, Nazira, Lenir, Ivo, Helenice, Erna, Maria, Luís e Ivone, Eugênia.

Aos colegas e amigos do Departamento de Medicina Preventiva e Social, especialmente Rose, Eduardo e Zezé.

Aos professores Heleno Rodrigues Corrêa, meu co-orientador, Ana Maria Segall Corrêa e Solange L'Abbate, pela confiança, estímulo, exemplo, e, sobretudo pelo carinho.

Ao professor Nelson Rodrigues dos Santos, meu orientador, pela confiança e pela paciência.

Aos professores Zulmira Maria de Araújo Hartz e Gastão Wagner de Souza Campos, pela compreensão e apoio.

À Dra. Alice, cujo apoio deu início a esta jornada.

Ao Governo do Estado do Rio Grande do Sul, que viabilizou a realização deste mestrado, concedendo-me licença remunerada. No primeiro período, através de Sandra Collares, minha chefe imediata, Júlio Hoczman, Secretário Estadual da Saúde e Alceu Collares, Governador do Estado. No segundo, respectivamente, através de Maria Ivanilda Vernier, Germano Mostardeiro Bonow e Antônio Brito.

À Coordenadoria de Aperfeiçoamento do Ensino Superior (CAPES), pelo apoio financeiro através da bolsa de estudos.

“Senhor

dai-me coragem,

para mudar o que pode ser mudado,

serenidade,

para aceitar aquilo que não pode ser

mudado;

e sabedoria,

para saber a diferença.”

RESUMO

Este trabalho é uma reflexão sobre a temática da avaliação no campo da Saúde Coletiva.

O interesse pela avaliação tem crescido muito nos últimos tempos no Brasil, indicando que está conformando-se um novo campo de conhecimentos e de práticas na Saúde Coletiva. Isto se faz, necessariamente, através das relações interdisciplinares, por duas razões: porque a Saúde Coletiva é uma área interdisciplinar e porque a amplitude e complexidade dos objetos da avaliação em saúde exigem o concurso de conhecimentos produzidos em várias disciplinas para que possam ser compreendidos.

A revisão bibliográfica indica a fase incipiente em que se encontra o campo da avaliação em saúde no país, quando comparada com a produção científica na América do Norte e na Europa. Entretanto, tem havido um crescimento da produção científica nacional, com uma diversidade de enfoques que varia desde os trabalhos de revisão teórica até relatórios de pesquisas de avaliação dirigidas à uma diversidade de objetos, em níveis de amplitude e complexidade crescentes: procedimentos e cuidados individuais, tecnologias, serviços e programas, sistemas de atenção em saúde, políticas de saúde.

Desenvolvemos o trabalho em torno de dois eixos: 1) aspectos teórico-metodológicos; 2) aspectos históricos.

A estrutura do primeiro capítulo é a mesma dos projetos de pesquisa: finalidade, objeto, metodologia, sujeito. Na metodologia, discutimos as grandes questões da pesquisa em avaliação, tanto aquelas que derivam da problemática das Ciências Sociais, como subjetividade *versus* objetividade, qualitativo *versus* quantitativo e validade do conhecimento, quanto as questões específicas da avaliação: a escolha do tipo de avaliação e os problemas típicos da pesquisa neste campo, rigor científico *versus* demandas políticas, “avaliabilidade” dos programas e questões éticas, a maioria derivadas do contexto político onde ocorre o processo avaliativo. Identificamos duas linhas discursivas na Saúde Coletiva: a avaliação de qualidade e a avaliação de programas. Além disto, discutimos o conceito e a tarefa da avaliação e os principais conceitos operacionais dos estudos em avaliação, quais sejam: estrutura-processo-resultados; eficácia-efetividade-eficiência; indicador-critério-padrão e qualidade.

No segundo capítulo, fazemos uma revisão da temática da avaliação nas três áreas básicas de conhecimento em Saúde Pública: Ciências Sociais, Planejamento e Epidemiologia. Verificamos que o campo da avaliação está bem desenvolvido e institucionalizado nos países anglo-saxões, especialmente no Canadá, Estados Unidos da América do Norte e Inglaterra, principalmente devido ao interesse estatal na avaliação das políticas públicas. Na América Latina em geral, e no Brasil especificamente, o padrão clientelista e autoritário do Estado tem restringido o desenvolvimento da avaliação.

Concluimos que o contexto atual do Sistema Único de Saúde favorece o desenvolvimento da avaliação no Brasil.

RESUMO

Este trabalho é uma reflexão sobre a temática da avaliação no campo da Saúde Coletiva.

O interesse pela avaliação tem crescido muito nos últimos tempos no Brasil, indicando que está conformando-se um novo campo de conhecimentos e de práticas na Saúde Coletiva. Isto se faz, necessariamente, através das relações interdisciplinares, por duas razões: porque a Saúde Coletiva é uma área interdisciplinar e porque a amplitude e complexidade dos objetos da avaliação em saúde exigem o concurso de conhecimentos produzidos em várias disciplinas para que possam ser compreendidos.

A revisão bibliográfica indica a fase incipiente em que se encontra o campo da avaliação em saúde no país, quando comparada com a produção científica na América do Norte e na Europa. Entretanto, tem havido um crescimento da produção científica nacional, com uma diversidade de enfoques que varia desde os trabalhos de revisão teórica até relatórios de pesquisas de avaliação dirigidas à uma diversidade de objetos, em níveis de amplitude e complexidade crescentes: procedimentos e cuidados individuais, tecnologias, serviços e programas, sistemas de atenção em saúde, políticas de saúde.

Desenvolvemos o trabalho em torno de dois eixos: 1) aspectos teórico-metodológicos; 2) aspectos históricos.

A estrutura do primeiro capítulo é a mesma dos projetos de pesquisa: finalidade, objeto, metodologia, sujeito. Na metodologia, discutimos as grandes questões da pesquisa em avaliação, tanto aquelas que derivam da problemática das Ciências Sociais, como subjetividade *versus* objetividade, qualitativo *versus* quantitativo e validade do conhecimento, quanto as questões específicas da avaliação: a escolha do tipo de avaliação e os problemas típicos da pesquisa neste campo, rigor científico *versus* demandas políticas, “avaliabilidade” dos programas e questões éticas, a maioria derivadas do contexto político onde ocorre o processo avaliativo. Identificamos duas linhas discursivas na Saúde Coletiva: a avaliação de qualidade e a avaliação de programas. Além disto, discutimos o conceito e a tarefa da avaliação e os principais conceitos operacionais dos estudos em avaliação, quais sejam: estrutura-processo-resultados; eficácia-efetividade-eficiência; indicador-critério-padrão e qualidade.

No segundo capítulo, fazemos uma revisão da temática da avaliação nas três áreas básicas de conhecimento em Saúde Pública: Ciências Sociais, Planejamento e Epidemiologia. Verificamos que o campo da avaliação está bem desenvolvido e institucionalizado nos países anglo-saxões, especialmente no Canadá, Estados Unidos da América do Norte e Inglaterra, principalmente devido ao interesse estatal na avaliação das políticas públicas. Na América Latina em geral, e no Brasil especificamente, o padrão clientelista e autoritário do Estado tem restringido o desenvolvimento da avaliação.

Concluimos que o contexto atual do Sistema Único de Saúde favorece o desenvolvimento da avaliação no Brasil.

MOTIVAÇÕES

Este trabalho é fruto das inquietações de uma trajetória de vinte anos no funcionalismo público, treze dos quais no setor da saúde.

Em primeiro lugar, reflete a convicção na relevância do papel do Estado, cujas instituições devem garantir os direitos de cidadania, entre eles a saúde; em segundo lugar, a convicção de que o servidor público pode exercer seu papel de modo eficaz, eficiente e, sobretudo, com dignidade.

Desde quando cursava a Faculdade de Medicina, fui sendo paralelamente "formada" pela experiência concreta de funcionária pública, pois já trabalhava em uma secretaria de estado, no Rio Grande do Sul. Logo após a formatura, fui percorrendo o que acredito ser a trajetória típica do sanitarista de nosso tempo, que começa com a especialização em Saúde Pública, passa pelo trabalho como clínico, em um pequeno Posto de Saúde, localizado no prédio da associação de moradores de uma favela, segue na chefia de uma Unidade de Saúde de maior porte, depois nas equipes de supervisão de em um serviço regional, e finalmente chega ao setor de planejamento do nível central da Secretaria de Saúde do Estado. Além disto, paralelamente, inclui um curta experiência de docência, como professora-horista em uma universidade pública federal.

Nessa trajetória, totalmente dedicada ao setor público - o que nos tempos atuais soa como um demérito -, foi possível conhecer o sistema público de saúde em todos os níveis hierárquicos de um estado da federação - local, regional e central - e do ponto de vista de duas instituições fundamentais deste sistema - uma Secretaria Estadual de Saúde, com tradição de Saúde Pública, e o INAMPS, voltado para a assistência individual curativa.

Além disto, pude participar diretamente da implantação dos projetos inovadores da política de saúde nos últimos anos, originados no chamado Movimento da Reforma Sanitária -

primeiro, as Ações Integradas de Saúde (AIS), depois, o Sistema Unificado e Descentralizado de Saúde (SUDS), e finalmente o Sistema Único de Saúde (SUS) -, tendo acompanhado o processo de reordenamento institucional do setor, especialmente a municipalização da saúde.

Esta experiência profissional fez surgir muitos interesses e inquietações, que me motivaram a buscar o Mestrado em Saúde Coletiva, como um espaço para a reflexão e para a instrumentalização na busca de respostas para algumas questões. A principal delas era motivada pela angústia que eu vivia na recém criada Divisão de Municipalização da Saúde, onde deveria fazer a análise dos Planos Municipais de Saúde, emitindo um parecer técnico sobre os mesmos. Ou seja, meu papel era avaliar, sem que os critérios de avaliação estivessem claros.

Não havia experiência institucional acumulada para orientar-me sobre o desempenho deste papel de avaliadora, até porque a situação era nova para todos. Um ponto que parecia crucial era como avaliar a participação da Epidemiologia no processo de planejamento, visto que a norma legal previa o uso de critérios epidemiológicos na definição de prioridades, na orientação da organização de serviços de saúde, bem como na alocação de recursos financeiros.

Por isto, minha intenção original de fazer o Mestrado era para ter contato com "metodologias simplificadas" para a elaboração de planos e para o acompanhamento e avaliação das políticas municipais de saúde, visando a capacitação de pessoal nos municípios e nas instâncias regionais da Secretaria Estadual de Saúde.

Em função deste interesse, aceitei o convite da área de Epidemiologia do Departamento de Medicina Preventiva e Social da Faculdade de Ciências Médicas da Universidade Estadual de Campinas, para participar como colaboradora nas atividades de avaliação do Convênio CENEPI/UNICAMP, através do qual este departamento passava a funcionar como um dos núcleos de capacitação do Programa de Capacitação de Recursos Humanos em Epidemiologia, do Centro Nacional de Epidemiologia, órgão da Fundação Nacional de Saúde, vinculada ao Ministério da Saúde.

Em razão da avaliação do Convênio CENEPI/UNICAMP, acabei participando da realização de três pesquisas em avaliação:

- a primeira, “Avaliação do Convênio CENEPI/UNICAMP”, realizada pela área de Epidemiologia do Departamento de Medicina Preventiva e Social da Faculdade de Ciências Médicas da UNICAMP, e coordenada pela Prof^a. Ana Maria Segall Corrêa;

- a segunda, denominada “Diagnóstico das condições de uso da Epidemiologia no SUS”, cuja finalidade era avaliar as condições de uso da Epidemiologia nos serviços municipais da região de Campinas, São Paulo, com vistas à subsidiar o Programa de Capacitação em Epidemiologia, no âmbito do Convênio CENEPI/UNICAMP, coordenada por mim;

- a terceira, “Pesquisa de Acompanhamento e Avaliação do Curso de Especialização em Saúde Pública”, do referido departamento, realizada pela área de Ciências Sociais do mesmo e coordenada pela Prof^a. Dr^a. Solange L'Abbate.

Além da participação direta no planejamento do estudo, na coleta e análise dos dados destas pesquisas, acompanhamos, à distância, a realização de uma quarta pesquisa - “Avaliação do Programa de Capacitação em Epidemiologia do CENEPI” -, realizada por uma equipe de assessores contratados pelo próprio Centro Nacional de Epidemiologia, com a finalidade de avaliar o referido programa, com abrangência nacional.

O tema para este estudo foi então surgindo de uma forma natural, na medida em que eu me questionava sobre a forma de conduzir uma avaliação.

INTRODUÇÃO

"Ou seja, nada pode ser intelectualmente um problema, se não tiver sido, em primeira instância, um problema na vida prática. Isto quer dizer que a escolha de um tema não emerge espontaneamente, da mesma forma que o conhecimento não é espontâneo. Surge de interesses e circunstâncias socialmente condicionadas, frutos de determinada inserção no real, nele encontrando suas razões e seus objetivos" (Minayo: 1992, 90).

Este trabalho é uma reflexão sobre a temática da avaliação no campo da Saúde Coletiva. Portanto, está voltado para uma visão geral teórico-metodológica do campo da avaliação, e não para uma revisão de métodos e técnicas para elaborar estudos nesta área. Tampouco, consiste da apresentação de resultados de uma pesquisa em avaliação.

Uma vez que toda a produção teórica nasce de um problema concreto, também este trabalho é o fruto de uma *práxis*, no sentido de que consiste de uma reflexão sobre uma prática concreta - nossa participação direta em três pesquisas de avaliação, uma das quais como coordenadora, e o acompanhamento, à distância, de uma quarta avaliação.

Estas experiências concretas propiciaram o aparecimento de uma série de questões, que surgiram durante a realização das pesquisas, como um processo natural de "vigilância epistemológica" sobre os modelos de estudo em avaliação, seus procedimentos e instrumentos de coleta de dados, os modos de análise.

Segundo BRUYNE e cols., a prática científica faz com que o pesquisador coloque para si, consciente ou inconscientemente, uma série de questões, de ordem epistemológica, que "podem ajudá-lo a resolver seus problemas práticos e a elaborar soluções teóricas válidas". Citando R. Blanché, os autores colocam que esta epistemologia interna "nasce de algum modo sob

os próprios passos do cientista, ...ela é exigida por problemas que se colocam no próprio interior da ciência." Neste sentido, "todo pesquisador torna-se filósofo, é obrigado a isto pelos problemas que encontra em suas abordagens: natureza da explicação, natureza dos fatos, validade dos procedimentos, etc" (Bruyne et al: 1982, 35-36; 44).

Ou seja, deste processo de vigilância epistemológica, que é interno e intrínseco ao trabalho do pesquisador, foram surgindo muitas questões sobre o modo de operacionalizar uma pesquisa em avaliação. Posteriormente, a reflexão sobre estas questões foi instrumentada por uma revisão bibliográfica sobre o assunto, dando origem ao trabalho que ora apresentamos.

É, portanto, por referência à uma função epistemológica "intrínseca" a toda prática de pesquisa que definimos nosso objeto de estudo como uma **reflexão** sobre avaliação. Mais ou menos como coloca Jean LADRIÈRE, sobre aqueles que "já tendo a experiência da pesquisa, encontram-se num daqueles momentos em que, voltando-se criticamente sobre o adquirido, experimentam a necessidade de recolocar em questão os procedimentos dos quais se serviram e de assegurar-se novamente das bases sobre as quais construíram" (Ladrière apud Bruyne et al: 1982, 21).

Quando entramos na literatura específica da área, tivemos a sensação de que estávamos em um campo que abacava tudo, pois avaliar diz respeito a tudo. A abrangência do tema e o seu caráter interdisciplinar trouxeram uma dificuldade extraordinária para que nos localizássemos no campo da avaliação, totalmente novo para nós.

No campo da Saúde Coletiva, tradicionalmente, a temática da avaliação é vinculada à área do Planejamento em saúde, como a etapa final da implementação dos planos, projetos ou programas. Entretanto, a maior parte da literatura disponível era oriunda da área das Ciências Sociais, sendo voltada para a avaliação de programas sociais. Por outro lado, a Epidemiologia também tratava do assunto, onde falava-se em fazer pesquisas para avaliar o impacto do serviços de saúde. Afinal, onde estávamos metidos?

A bibliografia que consultamos indicava que que o campo da avaliação em Saúde Coletiva era relativamente incipiente, no Brasil, quando comparado com a produção da literatura anglo-francesa. Entretanto, parecia que a temática da "avaliação em saúde" vinha

despertando cada vez mais interesse, tanto por parte do pessoal vinculado aos serviços quanto daqueles ligados à academia, evidenciado-se pelo aumento de trabalhos relacionados à temática, em eventos científicos, como congressos e encontros, e na publicação de artigos.

A produção científica no campo da avaliação em saúde vem crescendo, tanto no que concerne aos relatos de pesquisas quanto aos artigos de revisão. Em recente artigo de revisão, são mencionados diversos estudos, com diferentes enfoques: avaliação de serviços de saúde, incluindo estudos sobre a utilização e a qualidade dos serviços, avaliação de programas, avaliação do impacto de ações de atenção primária, avaliação da implementação de políticas públicas, incluindo os processos de descentralização das ações de saúde, avaliação da equidade, avaliação de tecnologias e estudos sobre a questão da qualidade (Silva & Formigli: 1994, 87).

Em eventos recentes, tem sido colocado, inclusive, que a “avaliação” talvez esteja conformando-se como uma nova área de conhecimentos e práticas no campo na Saúde Coletiva¹. Inclusive, já se coloca em discussão a possibilidade de formação de “profissionais em avaliação” no campo da Saúde Coletiva (Hartz & Camacho: 1995).

A possibilidade de ver a temática da avaliação como uma nova área de conhecimentos, levou-nos a esta tentativa de delimitar a conformação da mesma no campo da Saúde Coletiva. Esta tentativa implicou em duas posturas teóricas: primeiro, colocar a Avaliação lado a lado com as tradicionais áreas de conhecimento que informam o campo da Saúde Pública - Ciências Sociais, Planejamento e Epidemiologia -, buscando identificar seus conteúdos específicos; segundo, buscar elementos do discurso da avaliação nestas três áreas. Com isto, visamos conhecer o duplo sentido das relações interdisciplinares que regem a formação das novas áreas de conhecimento, onde um discurso específico vai conformando-se a partir da contribuição de diversas áreas para, em seguida, passar a influir nas mesmas com os elementos típicos do seu discurso próprio.

¹ Conforme colocações da Prof^a. H. Maria Novaes, na abertura da Mesa Redonda "A Epidemiologia na Avaliação dos Serviços de Saúde: A discussão da Qualidade", no I Congresso Paulista de Epidemiologia, realizado de 9 a 11 de março de 1995, na cidade de São Paulo (utilizamos nossas anotações pessoais, visto que os Anais deste congresso ainda não foram publicados) . Também na fala da Prof^a Lígia M. V. da Silva, na abertura da Comunicação Coordenada 47-"Novas tendências da pesquisa em Avaliação de Serviços", no III Congresso de Epidemiologia da ABRASCO, realizado em Salvador ,de 24 a 28 de abril do mesmo ano.

Ou seja, a avaliação, enquanto área específica de “saberes e práticas” no campo da Saúde Coletiva, é formada por discursos oriundos das diversas disciplinas específicas deste campo, especialmente as três mencionadas anteriormente. Porém, poderíamos afirmar que o fenômeno ocorre de modo contrário: é o discurso específico da avaliação que se expressa nas diversas disciplinas, passando a compor suas práticas discursivas². Seja como for, no capítulo referente aos aspectos históricos da temática da avaliação em Saúde Coletiva, buscamos identificar e apreender as características do discurso da avaliação nas três disciplinas básicas que informam este campo, em uma perspectiva histórica do desenvolvimento destes campos disciplinares.

Em consequência desta postura, nesta nossa primeira abordagem da área temática, recorreremos a autores de diferentes campos disciplinares, com referenciais teórico-metodológicos diferentes, reproduzindo o caráter “indisciplinar” da avaliação. O resultado deste “vôo” teórico é que, do ponto de vista bibliográfico e parafraseando Robert Farr, este trabalho pode ser tomado como um exemplo de “intertextualidade” (Robert M. Farr In Jovchelovitch & Graeschi: 1994, 48). Talvez tenhamos pecado por um excesso de ecletismo, mas, em determinados momentos, a compreensão do tema exigia que juntássemos idéias que, depois, não conseguimos mais separar. A medida que fomos aprofundando nossa reflexão sobre avaliação, percebemos que este tema exige um esforço para integrar as abordagens das disciplinas básicas da Saúde Coletiva.

A estrutura deste trabalho reflete nossa intenção original de fazer uma “avaliação da avaliação”, comparando aspectos metodológicos das pesquisas das quais havíamos participado, que mencionamos no início. Ao tentar elaborar nosso projeto de pesquisa, fomos percebendo que o campo da avaliação, por ser essencialmente aplicado, colocava em evidência todas as questões cruciais da pesquisa em Ciências Sociais, de onde, aliás, extrai suas teorias e métodos, agudizando todos os impasses e conflitos da pesquisa neste campo.

Isto levou-nos a desenvolver nossa reflexão em torno da estrutura clássica dos projetos de pesquisa: finalidade, objeto, metodologia e caracterização do sujeito do estudo em

²Fazemos este raciocínio e utilizamos os conceitos de “práticas discursivas”, “saberes” e “disciplina” conforme as idéias de Michel Foucault, que vemos em: FOUCAULT, Michel. *A arqueologia do saber*. Petrópolis, Editora Vozes, 1972.

avaliação. Além disto, para entender melhor a avaliação, necessitávamos conhecer alguma coisa sobre a forma como este campo desenvolveu-se.

Em função disto, e para facilitar nosso trajeto em um tema tão amplo e complexo, desenvolvemos nosso estudo em dois capítulos: 1) aspectos teórico-metodológicos; 2) aspectos históricos.

No primeiro capítulo, buscamos caracterizar os aspectos gerais do quadro teórico-metodológico do campo da avaliação: o conceito, os objetos, as metodologias, os sujeitos da avaliação. No segundo capítulo, apresentamos uma sucinta historiografia do desenvolvimento da área temática, buscando caracterizar uma “história institucional” do campo, referida às três áreas de conhecimento que tradicionalmente subsidiam a Saúde Coletiva - Ciências Sociais, Planejamento e Epidemiologia.

Nos aspectos teórico-metodológicos, buscamos responder as perguntas clássicas: o que é avaliação e qual a sua finalidade (conceito)? o quê se pode avaliar (objeto)? como avaliar (metodologia)? para quem avaliar, e em que contexto (dificuldades específicas da pesquisa em avaliação)? quem avalia (sujeito)?

O conceito de avaliação - atribuir juízo de valor - comporta uma ambigüidade intrínseca de sentido - subjetivo e objetivo - que permeia a prática e a produção de conhecimentos no campo. Disto derivam todas as questões metodológicas deste campo de pesquisa.

Em nosso entendimento, o campo da avaliação expõe, de forma crucial, todas as grandes questões metodológicas das Ciências Sociais, até porque a avaliação é considerada usualmente uma forma de “ciência social aplicada”. E é justamente quando tentamos aplicar uma teoria em uma situação concreta que percebemos melhor seus limites e possibilidades.

Por isto, a seção referente à metodologia foi desenvolvida em duas partes. Em primeiro lugar, as questões gerais que caracterizam a problemática das Ciências Sociais - subjetividade *versus* objetividade, qualitativo *versus* quantitativo, validade do conhecimento, que a avaliação expõe de forma crucial. Em segundo lugar, desenvolvemos algumas questões específicas da pesquisa em avaliação - a tipologia das avaliações, as grandes linhas discursivas da avaliação em saúde (avaliação de qualidade e avaliação de programas), e as dificuldades específicas do processo avaliativo, que somam-se às dificuldades gerais da pesquisa em ciências sociais.

Estas dificuldades específicas referem-se à avaliabilidade” dos programas, isto é, dificuldade de avaliar resultados diante da falta de clareza dos objetivos; à questão do rigor científico *versus* demandas políticas; as questões éticas que permeiam as práticas em avaliação.

Além disto, buscamos caracterizar o que consideramos ser as tríades conceituais básicas da avaliação: estrutura-processo-resultados; eficácia-efetividade-eficiência; indicador-critério-padrão. Verificamos que estes conceitos, embora bastante utilizados na linguagem cotidiana, tem significados de difícil apreensão, muitas vezes parecendo sobreponíveis.

A avaliação é vista como um processo que tem duas dimensões: técnica e social. A primeira, reproduzindo o esforço de objetivação da ciência; a segunda, carregada padrões de valores construídos subjetivamente na vida social.

Por outro lado, ao percorrer o campo da avaliação, fomos sendo guiados, inconscientemente, por uma série de dicotomias “clássicas” do campo da Saúde Coletiva, com as quais vínhamos nos debatendo em toda nossa trajetória profissional, e que a avaliação possibilita discutir e superar: individual *versus* coletivo; reais necessidades de saúde *versus* demandas por serviços; Clínica *versus* Saúde Pública; técnico *versus* político.

De certa forma, a medida que fomos entendendo o campo da avaliação através destas questões, fomos verificando, dialeticamente, que era neste campo onde encontrávamos elementos para melhor compreendê-las, superando as dicotomias.

Ao discutir as grandes linhas discursivas da avaliação em saúde - a Avaliação de Qualidade e a Avaliação de Programas -, verificamos que o tradicional eixo individual/coletivo que separa as práticas clínicas das práticas em Saúde Pública delimitavam, de certa forma, os campos destes discursos. Entretanto, encontramos na literatura canadense, mais uma vez, a força integrativa da avaliação, na proposta de integrar estas linhas em uma estrutura conceitual única.

Na revisão histórica, verificamos que embora tradicionalmente compreendida como uma questão de planejamento em saúde, a avaliação enquanto campo de estudos desenvolveu-se principalmente dentro da área das Ciências Sociais. A incorporação da estatística e dos métodos experimentais nesta área deu origem, na década de 60, nos EUA, à chamada “Pesquisa Avaliativa”, e por esta razão os desenhos de estudo propostos por esta corrente de pesquisa são muito semelhantes aos da Epidemiologia

Constatamos que o campo da avaliação está bem desenvolvido e institucionalizado em alguns países da Europa e na América do Norte, especialmente os EUA e o Canadá, onde existem inclusive associações de profissionais em avaliação, bem como inúmeros periódicos e publicações específicas sobre o tema. Nestes dois países, houve uma institucionalização da avaliação dentro do aparelho de Estado, com a constituição de órgãos cuja finalidade é avaliar as políticas públicas. A situação que vemos na literatura estrangeira não se verifica no Brasil, principalmente devido ao padrão clientelista e autoritário que, historicamente, tem caracterizado a atuação do Estado brasileiro, gerando fortes resistências políticas à avaliação dentro do setor público.

A nosso ver, a criação do SUS constitui um fator que deve impulsionar grandemente o desenvolvimento do campo da avaliação no Brasil. Alguns conceitos do campo da avaliação, como "acesso", "integralidade", "qualidade", que foram tomados como objetivos de luta do Movimento da Reforma Sanitária, tornaram-se diretrizes da Lei Orgânica de Saúde (Silver: 1992, 197).

Alguns autores colocam, inclusive, que avaliar certos atributos de qualidade das práticas em serviços de saúde, tais como "cobertura" e "equidade", corresponderia a "verificar o cumprimento do preceito consitucional do direito à saúde" (Silva & Formigli: 1994, 88; 80).

Uma rápida revisão da Lei Federal 8080/90, a Lei Orgânica da Saúde (LOS), revela que a maioria das atribuições do SUS são definidas através de conceitos próprios do campo da avaliação. Os quatro artigos onde são definidas as atribuições dos diferentes níveis de direção do SUS (art. 15 a art. 18), contêm sessenta e seis (66) incisos, dos quais aproximadamente 45% se referem a "controle e avaliação", "acompanhamento e fiscalização", "elaboração de normas técnicas", definição de "critérios" e "parâmetros", estabelecimento de "padrões de qualidade", organização de "sistemas de informações" em saúde. Isto sem contar os artigos que referem-se à vigilância em saúde: epidemiológica, sanitária, nutricional, saúde do trabalhador e ambiente do trabalho, meio ambiente em geral. A LOS estabelece que deverão ser definidas "as instâncias e mecanismos de controle, avaliação e fiscalização das ações e serviços de saúde" (art. 15, inciso I) e também de "acompanhamento" (art. 16, inciso XVII). Além disto, a direção nacional deveria estabelecer o "Sistema Nacional de Auditoria e coordenar a avaliação técnica e financeira do SUS, em todo o território nacional, em

cooperação técnica com os estados, municípios e Distrito Federal” (art. 16, inciso XIX) (CONASEMS: 1990, 9-13).

Entretanto, o cumprimento destas proposições de avaliação fica dificultado na medida em que o significado dos conceitos não ficou claro no texto legal (Silver: 1992, 197). Neste sentido, a garantia da proposta de reorganização das práticas de saúde no SUS passaria pela necessária "definição de estratégias para avaliação, sintonizadas com as modificações propostas para um novo modelo assistencial" (Paim apud Silva & Formigli: 1994, 88).

Enfim, o objetivo deste trabalho é fazer uma revisão teórica da temática da avaliação no campo da Saúde Coletiva, a partir dos conhecimentos produzidos nas três áreas de conhecimento que informam este campo - Ciências Sociais, Planejamento e Epidemiologia.

Do ponto de vista metodológico, o estudo caracteriza-se como uma reflexão teórica, através da qual buscamos estabelecer inter-relações entre as diversas áreas disciplinares envolvidas com a temática da avaliação. Ao propormos fazer uma “reflexão” em um campo tão “prático” quanto este, sentimos-nos respaldados por BACHELAR, quando afirma que: “A realidade visada pelas ciências, seus objetos, deve ser tanto descoberta prospectivamente quanto provada reflexivamente” (G. Bachelar apud Bruyne et al.: 1982, 201).

Diante da forma abrangente como tratamos o tema, corremos o risco de que este trabalho possa parecer, em alguns momentos, “um conglomerado de coisas dispersas”, como adverte ECO (Eco: 1983, 10). Esperamos que isto não tenha acontecido, mas, se tiver, resta-nos o consolo de que possam ser coisas pertinentes àqueles que iniciam seu contato com a área da avaliação, como foi o nosso caso.

Considerando a carência relativa de estudos em avaliação voltados para as intervenções de saúde, no contexto do SUS, e considerando a também relativa “novidade” do campo da avaliação no Brasil, acreditamos que o estudo que desenvolvemos tem alguma relevância, não apenas porque organiza de uma certa forma o assunto, mas principalmente porque estamos buscando contribuir para o desenvolvimento de uma “cultura de avaliação” nas instituições públicas no Brasil, voltada para a prestação de contas à população.

Para finalizar, cabe esclarecer alguns pontos que permeiam este trabalho.

Em primeiro lugar, algumas idéias fundamentais deste trabalho não foram colhidas nas fontes formais, ou seja, na literatura, mas em eventos científicos recentes, onde a questão

da avaliação foi discutida, como congressos, e mesmo palestras em cursos de pós-graduação. Sem querer eximir-mo-nos do dever de buscar fontes mais formais, entendemos que a “novidade” relativa do campo, no Brasil, autoriza-nos a este procedimento, que, embora não seja o mais adequado ao rigor científico, expressa a busca da construção de um "saber".

Em segundo lugar, devido à relativa escassez de literatura específica no Brasil, utilizamos, talvez de forma excessiva, a citação de terceiros. Para contornar, em parte, este problema, colocamos em nota de rodapé a referência bibliográfica dos autores mencionados em “apud”, para facilitar futuras pesquisas.

Em terceiro lugar, optamos por traduzir as citações das publicações em inglês e francês, visando tornar a leitura mais fácil, embora reconhecendo a possibilidade de equívocos nestas traduções.

Por último, cabe esclarecer utilizamos indistintamente os termos "Saúde Coletiva" e "Saúde Pública", embora reconhecendo as diferentes concepções que embasam os mesmos. O termo “Saúde Coletiva” é usado especificamente no Brasil, tendo surgido a partir das críticas que o movimento da Medicina Social latino-americana fazia ao conteúdo ideológico da Saúde Pública tradicional. É um termo busca, no conceito de “coletivo”, chamar atenção para a idéia de “direitos, situação histórica, comprometimento de condições de vida social e uma crítica ao indivíduo como responsável único por sua saúde/doença” (Minayo: 1992, 79). Do ponto de vista conceitual, ambos os termos tem limitações (Minayo: 1992, 14)

Apesar de endossar boa parte das críticas ao conteúdo ideológico do termo "Saúde Pública", entendemos que este termo, certamente por ser mais antigo, é reconhecido pela maioria das pessoas e tem significado para elas, sendo inclusive utilizado pelos meios de comunicação, tanto no Brasil quando no resto do mundo, e no meio acadêmico de uma forma geral.

**1. A TEMÁTICA DA AVALIAÇÃO NO CAMPO DA
SAÚDE COLETIVA:
ASPECTOS TEÓRICO-METODOLÓGICOS**

1.1. O conceito de avaliação: o que é avaliar.

Ao tentar compreender o conceito de avaliação de uma perspectiva acadêmica, deparamo-nos com uma grande dificuldade. Aos poucos, fomos entendendo que esta dificuldade tinha raízes em quatro ordens de razão, relacionadas entre si: a primeira, porque avaliar é uma atividade inerente ao homem, daí que o conceito de avaliação está muito impregnado por noções difusas de senso comum; segundo, porque enquanto campo de investigação científica, a avaliação é essencialmente uma forma de pesquisa aplicada, que caracteriza-se muito mais como uma finalidade de estudo do que por um objeto específico de estudo; terceiro, porque o desenvolvimento teórico do campo da avaliação é relativamente incipiente; quarto, porque a tarefa de avaliar implica em um duplo e aparentemente contraditório sentido, onde a subjetividade da noção de “juízo” mescla-se e alterna-se com a objetividade da noção de “medida” dentro do processo avaliativo.

Nesta seção, tentaremos demonstrar estas idéias, buscando identificar os elementos que consideramos essenciais para caracterizar a avaliação: na primeira parte, veremos algo sobre o objeto da avaliação, seu contexto e seus objetivos; na segunda parte, tentamos analisar os termos que constituem a essência da definição de avaliação, quais seja, as noções de juízo, valor e mérito; na terceira parte, buscamos demonstrar como o processo avaliativo percorre um círculo onde alternam-se procedimentos subjetivos e objetivos, a partir de uma visão sistêmica onde a idéia central é a retroalimentação; por fim, na quarta parte, tentamos esclarecer algumas confusões entre o processos de avaliação e o de planejamento, bem como com as noções de controle e de acompanhamento, seguimento ou monitoramento.

1.1.1. Objeto, Contexto e Objetivos

O ato de avaliar faz parte da vida cotidiana. Consciente ou inconscientemente, estamos sempre avaliando, de forma que a avaliação é “parte inseparável da vida humana”, sendo inclusive considerada uma das características que distingue o comportamento do homem do comportamento dos animais, guiados pelo instinto e pelo reflexo-condicionado³ (Suchman: 1967, 11, 23; Cohen&Franco: 1994, 72; Aguilar&Ander-Egg: 1994, 18).

Segundo SUCHMAN, a avaliação é “um processo social contínuo [...] básico a quase todas as formas de comportamento social, seja o de um indivíduo singular ou de uma organização complexa”. O ser humano está sempre avaliando, por isto, “a linguagem do dia-a-dia está repleta de adjetivos descritivos que conotam valor e juízo, como bonito, prudente, inútil, mau” (Suchman: 1967, 11, 7).

Em consequência da cotidianidade da tarefa de avaliar, endentemos que o conceito de avaliação ainda está muito permeado pelo seu entendimento de senso comum. Talvez por esta razão, o primeiro aspecto apontado pelos estudiosos do campo é a diversidade e a inespecificidade dos termos que se referem à tarefa de avaliar. WEISS entende que “avaliação é uma palavra elástica, a qual se estica para cobrir juízos de muitos tipos” (Weiss: 1972, 1).

Há três décadas atrás, SUCHMAN caracterizava assim o problema:

“Atualmente, o termo ‘avaliação’, apesar de sua ampla popularidade, é pobremente definido e usado frequentemente de forma inapropriada. Em grande parte, seu significado é dado por certo, e poucas tentativas foram feitas, mesmo por aqueles mais preocupados com o assunto, para formular uma definição conceitualmente rigorosa ou para analisar os princípios fundamentais do seu uso. O resultado é um amplo desacordo, com muitos outros termos tais como ‘assessment’, ‘appraisal’, e ‘judgment’ frequentemente sendo utilizados de forma intercambiável com ‘evaluation’” (Suchman: 1967, 27).

³A idéia de que a capacidade de avaliar distingue o comportamento humano do comportamento animal consta do editorial do “American Journal of Public Health”, em saudação à Primeira Conferência de Avaliação em Saúde Pública, realizada em 1955, nos EUA, conforme vemos citado em Suchman: 1967, 23, Nota de fim de capítulo.

Ainda hoje, SILVA&FORMIGLI chamam atenção sobre o problema:

*"...sob o rótulo de 'avaliação' reúnem-se diversas atividades, cujo espectro varia desde um julgamento **subjetivo** do desenvolvimento de determinada prática social, do tipo 'fomos bem?', 'deu resultados?', 'estamos satisfeitos?' até a chamada pesquisa avaliativa, que busca responder a perguntas semelhantes recorrendo a métodos e técnicas possuidores de maior **objetividade** [grifos nossos]"*
(Silva&Formigli: 1994, 80).

Uma vez que a tarefa de avaliar é inerente ao homem, tudo que existe pode ser avaliado. Segundo WEISS, o objeto da avaliação pode ser uma pessoa, uma coisa, uma idéia, uma ação isolada qualquer, uma intervenção organizada (Weiss: 1972, 1).

No campo dos fenômenos sociais, onde insere-se a Saúde Pública, o objeto primordial da avaliação, colocado de maneira ampla, é a ação humana intencional e organizada visando uma mudança social, ou seja, uma intervenção que pode ser um projeto, um programa, ou prática social qualquer (Weilenmann: 1980. 11-15). Por esta razão, a maioria dos autores consultados trata da avaliação das intervenções sociais, concebidas de forma ampla como ações organizadas para a mudança, geralmente sob forma de programas sociais.

Para WEILENMANN, é importante “vincular as atividades dos projetos com o contexto muito mais amplo da mudança social e transformação da sociedade” (Weilenmann: 1980, 14). Ouçamos autor:

“O veículo para efetivar a mudança social é a transformação da sociedade é o projeto ou projetos. Em outras palavras, as coisas não acontecem por si mesmas; para a mudança e a transformação terem lugar, a atividade organizada é essencial”. (Weilenmann: 1980, 11)

Entretanto, não é imprescindível a existência de um “programa” ou “projeto” para fazer a avaliação, pois esta não depende da “complexidade e organicidade das ações; basta que exista um conjunto de atividades específicas que se organizam para conseguir um determinado fim” (Aguilar&Ander-Egg: 1994, 34).

Uma vez que o objeto da avaliação no campo social é definido desta forma ampla, temos algumas implicações.

A primeira, que tendo por objeto de estudo a ação de “mudança social deliberada”, a avaliação é essencialmente “o processo através do qual o homem tenta verificar sua própria habilidade para influenciar outros homens ou o seu ambiente” (Suchman: 1967, 28-9). Ou seja, a avaliação é um instrumento investigativo que produz conhecimento sobre a ação social.

Ao tomar como objeto a ação humana organizada, a avaliação produz conhecimento sobre esta ação, de forma que a produção de conhecimento sobre as intervenções sociais é um dos objetivos fundamentais dos estudos em avaliação (Contandriopoulos et al.: 1992, Weilenmann: 1980, 29). Para WEILENMANN, acima de tudo, a avaliação deve ser vista como “um meio para desenvolver, testar ou refinar teorias de mudança social, em adição ao seu propósito mais tradicional que diz respeito à preocupação com a efetividade de programas específicos” (Weilenamnn: 1980, 29).

A segunda implicação decorre do fato que a ação humana visando a mudança social é essencialmente uma ação política.

Em consequência disto, embora a pesquisa avaliativa tenha muitas semelhanças com a pesquisa científica, elas são distintas porque “é característico da avaliação constituir uma pesquisa social encaixada em um contexto de poder e política” (Cohen&Franco: 1994, 280). Como coloca GADOTTI: “Seria ingênuo pensar que a avaliação é um processo técnico. Ela é também uma questão política” (Demo: 1987, 7, Prefácio de Moacir Gadotti).

Daí que alguns autores vejam a avaliação como uma “ferramenta” para a elaboração de políticas ou como “um aspecto crítico do processo político global, contribuindo para as decisões políticas e para a alocação de recursos” (Weilenmann: 1980, 27, 29). Especialmente a avaliação de programas públicos está localizada no “adverso contexto político” (Mc Queen: s/d, 41).

Porém, é preciso salientar que embora entendendo que a pesquisa em avaliação seja útil, os autores reconhecem que “a avaliação nunca é suficiente para estabelecer políticas” (Contandriopoulos et al.: 1992).

Neste sentido, a colocação de SUCHMAN é bem esclarecedora:

“Raramente é possível que todos os fatores envolvidos em uma decisão possam fazer parte de um modelo teórico. Entretanto, através desta técnica, pode-se aumentar o grau de compreensão do problema baseando-se em fatos prováveis, e o uso do método científico pode ser estendido ao domínio das decisões administrativas” (Suchman: 1967, 66).

A pesquisa em avaliação está portanto consagrada ao “princípio da utilidade”. Segundo VERGARA, não se avalia “somente por um exercício intelectual para descobrir uma verdade oculta”, senão porque se espera que da avaliação poderá haver o melhoramento de um projeto que está em marcha, ou mesmo sua eliminação, desde que “esta apareça como a medida mais razoável”. Ou seja, as conclusões que se chega a partir da avaliação devem gerar medidas que influenciem no andamento do projeto (Vergara: 1980, 16).

Esta idéia fica bem clara na pergunta de SUCHMAN: “Estamos interessados em produzir apenas a exposição do fato, ou que as pessoas tomem consciência do mesmo, ou que se interessem, e/ou que façam alguma ação?” (Suchman: 1967, 39).

O outro objetivo fundamental da avaliação, portanto, é fornecer subsídios ao processo de tomada de decisão. Segundo WEILENAMAN, a pesquisa e a análise avaliativa “são uma tarefa intelectual necessária para tornar viável o processo de decisão em uma base de continuidade” (Weilenmann: 1980, 13). Para VERGARA, a avaliação é “parte” do processo racional de tomada de decisões (Vergara: 1980, 16).

É, pois, em relação à formulação de políticas, ao processo de planejamento, programação e à tomada decisões que a avaliação é considerada, do ponto de vista acadêmico, essencialmente como “ciência social aplicada” (Weilenmann: 1980, 24-27; Aguilar&Ander-Egg: 1994, 32).

Discute-se, inclusive, se a avaliação poderia ser considerada “um ramo da ciência”, ou se deveria ser vista simplesmente como “uma aplicação da metodologia da pesquisa” (Cohen&Franco: 1994, 79).

Seja como for, o ponto de partida para qualquer avaliação é uma insatisfação com algo: se pensamos que tudo está bem, não sentimos necessidade de avaliar. Para SUCHMAN

“..... inerente à idéia de avaliação está uma atitude crítica da mente, um desafio ao status quo. Uma avaliação raramente ocorre em uma atmosfera de complacência e satisfação. O chamado para avaliar é usualmente o resultado de uma insatisfação em algum lugar” (Suchman: 1967, 41).

No campo das intervenções sociais, a atitude crítica inerente a qualquer processo avaliativo manifesta-se também na própria essência do objeto da avaliação - a ação de mudança social.

Todas as considerações feitas até aqui demonstram que a avaliação é um processo altamente complexo, tanto pela finalidade que se propõe - emitir um juízo de valor - quanto pelo objeto para a qual se volta - as intervenções sociais, pois além da complexidade da rede de determinantes que influem em qualquer problema e ação social (Suchman: 1967, 11, 20; Weilenmann: 1980, 31), estas intervenções envolvem disputas de interesse ou poder. Estas duas características remetem diretamente à subjetividade inerente à ação humana, dificultando em muito o propósito do campo da avaliação em constituir-se conforme o modelo hegemônico de ciência.

1.1.2. Os termos da definição de avaliação

Tanto porque a tarefa de avaliar faz parte da vida cotidiana, como devido ao fato de que a avaliação é considerada uma forma de pesquisa aplicada, o desenvolvimento teórico do campo ainda é incipiente, de forma que o conceito de avaliação está muito impregnado pelas noções de senso comum.

Referindo-se ao campo geral da avaliação, SUCHMAN mencionava, há trinta anos atrás, que havia na área uma “ausência de qualquer formalização teórica ou metodológica” (Suchman: 1967, 28). Recentemente, SILVA&FORMIGLI ainda referem “um incipiente desenvolvimento teórico” do campo da avaliação em saúde (Silva&Formigli: 1994, 88). Alguns artigos que consultamos são justamente tentativas de desenvolver uma estrutura ou um

modelo conceitual para o campo de avaliação em saúde, conforme se verifica no próprio título - “integrative model” (Donabedian et al.: 1982), “integrative framework” (Battista et al.: 1989), “conceptual framework” (Clemenhagen&Champagne: 1986).

Há indícios, portanto, de que o conceito de avaliação ainda está predominantemente localizado no que BRUYNE et al. denominam de “campo doxológico”, ou seja, o campo da “realidade de todos os dias”, “do saber não sistematizado, da linguagem e das evidências da prática cotidiana”, onde predomina a “doxa”, ou opinião (Bruyne et al.: 1982, 202-203, 33). Para que o o conceito de avaliação encontrasse “seu lugar teórico”, isto é, alcançasse maior especificidade enquanto conceito científico, seria necessário uma “ruptura epistemológica” mais evidente com as noções de senso comum (p. 34). E isto demandaria o amadurecimento de uma “teoria da avaliação”, em relação à qual as noções de senso comum seriam confrontadas.

Segundo SUCHMAN, existe uma “inextrincável mistura de definições conceituais e operacionais” na literatura sobre avaliação, com maior ênfase nas últimas. Ao contrário do que se esperaria de um conceito científico, a maioria das definições conceituais não atenderiam “nenhuma formalidade lógica”, mas ao invés disto, ofereceriam uma “lista de características descritivas da avaliação como um processo cognitivo e afetivo”. As definições operacionais, por sua vez, concentrar-se-iam nos “propósitos da avaliação” e nos “procedimentos” envolvidos na condução dos estudos avaliativos (Suchman: 1967, 28). Evidentemente, o predomínio das “definições operacionais” deve-se ao caráter essencialmente aplicado da pesquisa em avaliação.

Enfim, diante do estado atual do conhecimento, seria um ato de vaidade propor “uma definição universal e absoluta da avaliação” (Contandriopoulos et al.: 1992, 4). Apenas para ilustrar, recolhemos algumas tentativas de definição que encontramos na literatura, que podem ser vistas nos anexos.

Apesar de tudo, verifica-se que há praticamente um consenso entre os autores consultados, que o aspecto comum em todas as tentativas de definição conceitual da avaliação é a noção de “juízo de valor” ou de “julgamento do mérito”. Ou seja, avaliar é emitir juízo de valor ou de mérito sobre algo, que pode ser uma pessoa, uma coisa, uma idéia, uma ação ou intervenção (Suchman: 1967, 11; Weiss: 1972, 1; Hartz&Camacho: 1995).

A essência do que compõe o conceito de avaliação está portanto nucleada em torno de duas noções: a noção de “juízo” ou “julgamento” e a noção de “valor” ou “mérito”. Vamos analisar melhor cada uma delas.

Em primeiro lugar, é necessário estabelecer uma diferenciação entre as noções de “valor” e de “mérito”, feita por Scriven⁴ e mencionada por AGUILAR&ANDER-EGG:

“Assim, o valor de uma ação depende do grau de necessidade de realizá-la. Em outras palavras, quanto mais necessário é fazer algo, tanto maior tem seu valor. Por outro lado, o mérito tem a ver com a boa realização de algo, independentemente do seu valor. Uma ação é meritória quando faz bem e com eficiência aquilo que se propõe” (Aguilar&Egg: 1994, 26-27)

O autor chama atenção que uma ação pode ser meritória, isto é, ser bem executada, porém ter pouco valor, isto é, ser pouco necessária.

A noção de juízo de valor, por sua vez, leva-nos ao campo da filosofia e ao Dicionário Aurélio. Do ponto de vista filosófico, **juízo de valor** é "o que enuncia uma apreciação". Este conceito opõe-se ao de **juízo de realidade**, que é "o que enuncia um fato ou uma relação entre fatos". Os dois conceitos correspondem, respectivamente, à noção de “doxa” ou opinião, e à noção de “episteme” ou conhecimento, conforme vemos em NOGUEIRA (Nogueira: 1994, 33).

Esta clara distinção que é colocada no campo filosófico, entretanto, parece não existir nos processos concretos de avaliação, como tentaremos demonstrar a seguir.

O conceito de avaliação está situado exatamente sobre a linha que separa, filosoficamente, o mundo da “doxa” e o da “episteme”. No campo da “doxa”, isto é, da opinião, predomina o julgamento dos fatos baseado em valores subjetivos - o juízo de valor -, enquanto no campo da “episteme”, isto é, do conhecimento científico, predomina a demonstração lógica, objetiva, empírica da “verdade” dos fatos - o juízo de realidade. Em nosso entendimento, o conceito de avaliação é construído por elementos destes dois “mundos”, que reflete-se na permanente tensão entre subjetividade e objetividade do processo avaliativo.

⁴SCRIVEN, M. The methodology of Evaluation. In. *Perspectives on Curriculum Evaluation*. Chicago, Rand McNally, 1967.

A confusão começa no sentido que a avaliação tem na linguagem de senso comum, que fomos buscar no Dicionário Aurélio⁵.

Etimologicamente, o verbo avaliar significa “aproximar-se da valia”. A análise mais detalhada dos vocábulos ligados à definição de avaliação - “valia”, “valor”, “mérito” - causou a impressão de que havia algo aparentemente ambíguo e contraditório, que finalmente acabamos por entender como um duplo sentido destas palavras: “subjetivo” e “objetivo”.

No sentido subjetivo, avaliar tem o significado de "apreciar ou estimar o merecimento de; fazer idéia de; reconhecer a grandeza, a intensidade e a força de; fazer a apreciação, ajuizar". A subjetividade está caracterizada pela própria natureza do que é definido como valor - "qualidade pela qual determinada pessoa ou coisa é estimável em maior ou menor grau". Neste sentido, a idéia central de valor corresponde à noção de “mérito ou merecimento intrínseco”, e tem como sinônimos “força, audácia, coragem, valentia, vigor”. Também incluímos neste sentido o conceito de “valia”, referido à idéia de "utilidade, préstimo, serventia, valimento", ou seja, referida ao conceito marxista de “valor de uso”.

No sentido objetivo, avaliar significa "determinar a valia ou o valor, o preço". A objetividade está na noção de cálculo, de atribuir valor material a algo. O valor, neste sentido, é "o equivalente justo em dinheiro, mercadoria, etc, especialmente de coisa que pode ser comprada, ou vendida, preço, valia"; ou ainda, "a estimativa em dinheiro de um artigo, em determinado tempo, o preço do mercado". O sentido de valor aqui corresponde ao conceito marxista de “valor de troca”.

De qualquer forma, tanto enquanto “mérito” (qualidade intrínseca), “valia” (utilidade) ou “preço” (valor em dinheiro), o valor é atribuído ao objeto avaliado pelo sujeito da avaliação. Logo, qualquer valor, por mais objetivamente determinado que seja, será sempre “subjetivo”.

Também fomos verificar o sentido da palavra “avaliação” em Inglês⁶, língua na qual há uma vasta literatura sobre o tema. Verificamos que diferentes termos são traduzidos por “avaliação” para o português: “evaluation”, “assessment”, “appraisal”, “judgment”.

⁵ FERREIRA, Aurélio Buarque de Holanda. Novo Dicionário da Língua Portuguesa. Rio de Janeiro, Editora Nova Fronteira, 1986, 2ª ed., revista e aumentada.

⁶ VALLANDRO, Leonel & Vallandro, Lino. “*Dicionário Inglês-Português*”. Porto Alegre, Ed. Globo, 1979.

Nos vocábulos aparentemente mais utilizados - “evaluation” e “assessment”⁷-, predomina o “sentido objetivo” do termo avaliação. Assim, "evaluate" refere-se à avaliação no sentido de “computar, calcular; exprimir numericamente”; e "assessment" refere-se à avaliação para fins de “tributação” ou "determinação de taxa, quota ou multa". No verbete "estimate", dado como sinônimo, novamente encontramos a vertente "financeira" do termo: "avaliar, calcular, orçar" ou "fazer uma estimativa ou orçamento". Ou seja, não tem nada a ver com o sentimento de estima por algo ...

O termo ‘judgment’ (juízo), relacionado à “judge” (juiz), obviamente tem uma conotação menos matemática. No mesmo sentido de julgamento, descobrimos na etimologia do termo “assessment” uma referência à “seated beside a judge”, isto é, estar sentado ao lado do juiz. Por outro lado, na origem de “appaisal” está “praise”, que significa ato de expressar aprovação, admiração, louvor. Ou seja, ainda que um tanto perdido no passado, encontramos também na língua inglesa o sentido “subjetivo” da avaliação⁸.

1.1.3. A concepção sistêmica do processo avaliativo

Ao nuclear o conceito de avaliação em torno da noção de “juízo de valor”, os estudiosos do campo aparentemente estariam indicando uma concepção mais consoante com o aspecto "subjetivo" da palavra. Porém, a literatura que consultamos é pobre no tratamento teórico da questão dos valores sociais. Ao invés disto, o enfoque principal são os procedimentos metodológicos, onde a subjetividade, que é admitida teoricamente ao formular o conceito, cede lugar à busca da objetividade. Na descrição do “processo avaliativo”, os aspectos centrais são a construção de indicadores que possam ser traduzidos em números, a

⁷Fazemos esta afirmação baseados não somente na impressão subjetiva que ficamos das leituras em inglês, mas também na consulta ao dicionário de termos utilizados como “descritores” no banco de dados informatizados ERIC (Educational Research Information Center), da “American Educational Association”, onde encontramos os artigos sobre avaliação classificados sob estes dois grandes termos descritores.

⁸Consulta feita no Dicionário Webster’s: *Webster’s Encyclopedic Unabridged Dictionary of the English Language*. Gramercy Books, New York/Avenel/New Jersey, 1989 edition.

definição de critérios e parâmetros para comparar “objetivamente” estes indicadores e a validação dos instrumentos de medida das variáveis que compõe os indicadores.

Ou seja, embora a “definição conceitual” de avaliação enfatize a questão do “juízo de valor”, a definição “operacional”, predominante na literatura, indica que é a “medida”, e não o “juízo”, a questão central da avaliação. Para WEISS, avaliar implica na noção de que alguém está “examinando e pesando” algo (Weiss: 1972, 1). Ou seja, alguém está voltado para o “fato”, para o “juízo de realidade”.

A lógica que rege esta aparente incongruência entre a subjetividade do conceito e a objetividade dos procedimentos de avaliação pode ser assim resumida: se avaliar é emitir juízo de valor, para conformar este juízo é preciso “medir” o valor que está em questão. Daí a importância da “validade das medidas”, pois estas vão fundamentar o juízo.

O processo de avaliação pode ser comparado a um julgamento: para condenar ou absolver, é preciso apresentar as “provas” do crime. Ao historiar o desenvolvimento da avaliação enquanto campo de pesquisa, SUCHMAN menciona, em vários momentos, a necessidade de “provar” a efetividade e eficiência das intervenções públicas (Suchman: 1967). Não é à-toa, pois, que a simples idéia de ser avaliado provoca uma desagradável sensação de constrangimento na maioria das pessoas...

Entretanto, a questão da avaliação evidentemente não se esgota na “prova” objetiva, na “medida”. Os índices estatísticos “não são suficientes”, servem apenas para “descrever o estado” das coisas, mas não permitem qualquer avaliação de como as coisas estão indo, pois para fazer uma avaliação é preciso comparar (Suchman: 1967, 14).

ANDER-EGG esclarece que embora a avaliação esteja ligada à idéia de medição, não se trata da mesma coisa. A medição seria o ato de “determinar a extensão e/ou quantificação” de alguma coisa, enquanto a avaliação refere-se ao ato ou processo de “determinar o valor” de alguma coisa, que vai além de simplesmente medir ou quantificar (Ander-Egg: 1984, 19).

Para fazer um “juízo discricionário” sobre o que está sendo avaliado, é necessário, portanto, estabelecer uma comparação. Segundo SILVA&FORMIGLI, independentemente do nível de complexidade do que esteja sendo avaliado, “toda avaliação implica um julgamento, cujo veredicto baseia-se numa comparação entre aquilo que se deseja avaliar e uma situação ideal ou equivalente” (Silva & Formigli: 1994, 84).

A necessidade de definir uma “situação ideal” para comparar com a “situação real”, evidenciada pelas “medidas”, coloca outra questão crucial dentro do processo avaliativo: a existência de um critério de “bom”. Para SILVER, avaliar “implica primeiro a existência de uma noção de **bom**, e esta definição nunca foi fácil” (Silver: 1992, 195).

Enquanto os procedimentos de “medida” são objetivos, a noção do que é “bom” ou “ideal” está, evidentemente, no campo subjetivo dos valores. O juízo de valor implica em explicitar qual a noção de “bom” ou de “qualidade” ou de “justo” que está por trás deste juízo, para que este seja aceito como legítimo. Isto recoloca a avaliação novamente no campo subjetivo dos valores.

No raciocínio que fizemos até aqui é possível perceber uma noção de circularidade: juízo de valor > medir e comparar = juízo de realidade > critério de bom = juízo de valor.

De fato, o processo de avaliação é caracterizado na literatura por uma imagem de circularidade, oriunda da Teoria de Sistemas, cujo aspecto central é a noção de retroalimentação (Ver Anexo I).

Ao descrever o processo de avaliação, SUCHMAN utiliza a imagem de um círculo, que começa e termina na “formação de valores”. É no campo dos valores que se estabelece o “goal setting”, isto é, a situação onde os objetivos da ação são definidos, bem como o “goal measuring”, isto é, a definição dos critérios para medir o alcance dos objetivos (Suchman: 1967, 32-34). Ou seja, é a partir dos valores ou interesses que se define a situação ideal a ser alcançada, bem como os critérios para medir este alcance.

Na forma como ele descreve o processo avaliativo, podemos identificar basicamente três momentos: o estabelecimento de objetivos valorados (que podem ser explícitos ou implícitos) e das atividades para alcançá-los; a “medida do alcance dos objetivos”; e o “julgamento” se as atividades “valeram a pena”, isto é, o juízo de valor sobre a ação dirigida a um objetivo.

Vemos que este autor deixa bem claro como o processo avaliativo começa na subjetividade dos valores, onde definem-se os objetivos, busca a objetividade das medidas para verificar o alcance dos mesmos e volta à subjetividade do juízo de valor, para comparar o que se conseguiu alcançar com o que se buscava atingir.

Segundo WEILENMANN, a comparação das metas estabelecidas com os efeitos da ação é um tema proeminente da cibernética: “Um sistema seria cego se não fosse capaz de avaliar as informações sobre os efeitos de sua ação prévia”. Para ele, “o conceito de sistema e a perspectiva cibernética chamam atenção para a condição de circularidade e a característica de **rede** do fenômeno social” (Weilenmann: 1980, 13). É neste sentido que ANDER-EGG vê a avaliação essencialmente como um “instrumento de retroalimentação” (Ander-Egg: 1984, 19).

O processo avaliativo reproduz, de forma sintética, todo o esforço das ciências do homem para constituir-se enquanto “ciência”, conforme a concepção hegemônica de cientificidade. É o esforço de buscar o conhecimento “objetivo” da realidade, eliminando, tanto quanto possível, a “subjetividade” do sujeito que produz este conhecimento.

Este esforço fica demonstrado no próprio histórico do campo, conforme vemos nas quatro fases da avaliação, identificadas por Guba e Lincoln⁹, mencionadas em CONTANDRIOPOULOS et al.:

“O primeiro estágio é baseado na medida (dos resultados escolares, da inteligência, da produtividade dos trabalhadores). O avaliador é essencialmente um técnico que tem que saber usar e saber construir os instrumentos que permitem medir os fenômenos estudados. O segundo estágio se fortalece nos anos 20 e 30. Ele trata de identificar e descrever como os programas permitem atingir seus resultados. O terceiro estágio é baseado no julgamento. A avaliação deve permitir o julgamento de um intervenção. O quarto estágio está emergindo. A avaliação é então feita como um processo de negociação entre os atores envolvidos na intervenção a ser avaliada” (Contandriopoulos et al: 1992).

Percebe-se, claramente, que as duas primeiras fases correspondem a tentativas de “objetivar” o processo avaliativo, que na terceira fase acabam cedendo à subjetividade implícita na noção de julgamento. No quarto estágio, ainda em formação, há um reconhecimento do papel inevitável da subjetividade, na medida em que propõe a “negociação” entre os atores envolvidos no processo avaliativo.

⁹GUBA, E.G; LINCOLN, Y.S. *Fourth Generation Evaluation*. Newbury Park, Sage, 1990.

Esta fase corresponde ao surgimento da chamada “avaliação emancipatória”, que para GADOTTI corresponde a uma “concepção dialética da avaliação”, na qual, ao invés de constituir-se como “um exercício autoritário do poder de julgar”, a avaliação busca, ao contrário, constituir-se “num processo e num projeto em que avaliador e avaliando buscam e sofrem uma mudança qualitativa” (Demo: 1987, 7-8, Prefácio de Moacir Gadotti).

A maior parte da literatura disponível poderia ser enquadrada no terceiro estágio - a avaliação como julgamento -, estando fundamentada na concepção sistêmica da avaliação.

Neste ponto, caberia então destacar alguns aspectos que observamos na literatura que, em nosso entendimento, são decorrentes dos limites da visão sistêmica.

O predomínio da concepção sistêmica no campo da avaliação certamente contribuiu para que a questão dos valores tenha sido pouco trabalhada na literatura consultada. A questão dos valores implica em trabalhar com a contradição e o conflito, enquanto a visão sistêmica enfatiza a funcionalidade, a busca da harmonia.

Em uma crítica franca à concepção sistêmica da avaliação, MCQUEEN lembra que a avaliação envolve o “exercício do julgamento”, logo, é um processo carregado de valores, e não simplesmente um “controle cibernético” (McQueen: 1992, 43).

Embora a maioria dos autores reconheça o “contexto altamente carregado de valores da pesquisa avaliativa”, vimos somente em SUCHMAN, um dos precursores da avaliação de programas públicos, uma preocupação maior em desenvolver a questão dos valores. Ele apresenta uma breve definição de valor, como “qualquer aspecto de uma situação, evento ou objeto que está investido de um **interesse preferencial** [grifo nosso]”. Os valores seriam os princípios através dos quais se estabelecem as “prioridades e hierarquias de importância entre necessidades, demandas e objetivos”. O autor lembra que campo dos valores é um campo conflitivo, onde as “forças que estabelecem os objetivos estão sempre em competição entre si por dinheiro, recursos e esforços”. Ele chega a afirmar que “a pré-condição para um estudo avaliativo é a presença de alguma atividade cujos objetivos tem algum valor assumido”, embora este valor possa não estar explicitado (Suchman: 1967, 29; 32-34).

Uma vez que o círculo avaliativo começa na formação de valores, onde se define os objetivos da ação, e termina novamente neste campo, onde se define o critério de “bom”, que

vai fundamentar o juízo avaliativo, a omissão em tratar diretamente a questão dos valores tem duas consequências.

A primeira é a colocação em plano secundário do diagnóstico, ou avaliação inicial, ou avaliação *ex ante*, que justifica a intervenção a ser avaliada. Praticamente nenhum dos autores tratou a fundo esta questão, ou então, restringiu a avaliação inicial a uma questão econômica de “comparar a eficácia ou a rentabilidade” de diferentes alternativas de ação (Aguilar&Ander-Egg: 1994, 41). Ou seja, a literatura sobre o processo avaliativo acompanha o curso da ação. A avaliação inicial que precede esta ação, ou seja, o diagnóstico do problema que justifica a intervenção social organizada, praticamente não é tratado na literatura específica do campo.

Alguns autores destacam a importância de definir qual a transformação social que se está querendo, ou a importância em estabelecer o ponto de partida da ação e o ponto de chegada desejado: mudança do quê para o quê (Weilenmann: 1980, 12; Suchman: 1967, 39). Isto pressupõe que “o problema social atacado pelo programa deveria estar bem fundamentado em uma teoria” sobre a sociedade, que, entretanto não é tratada na literatura do campo da avaliação, até porque é considerada como “ciência aplicada”, que dispensa teorias (Weilenmann: 1980, 29).

A segunda consequência da omissão em tratar a questão dos valores é tentativa de restringir a avaliação a uma questão de “comparar medidas”, que vimos anteriormente.

De qualquer forma, sem uma teoria social que possa fundamentar uma discussão sobre os valores subjacentes à ação de mudança social, não há “base para a comparação com o passado nem com programas similares, e desta forma não haveria base para o refinamento e o melhoramento do programa; grosseiramente, significaria que o aprendizado não poderia ocorrer” (Weilenmann: 1980, 29). Com isto, a finalidade essencial da avaliação - “aprender da experiência, tanto exitosa ou fracassada, ajudando a racionalizar a tomada de decisões” - fica bastante comprometida (Cohen&Franco: 1994, 279).

O sentido da retroalimentação no processo avaliativo não deve resumir-se a um controle cibernético, apenas para ver se o sistema está funcionando conforme o previsto, mas implica em uma visão crítica sobre o próprio sistema, através da qual pode processar-se um aprendizado verdadeiro sobre a ação, contribuindo desta forma para a construção de uma teoria da ação social.

1.1.4. Esclarecendo algumas confusões

Nesta seção, pretendemos apontar as diferenças entre o conceito de avaliação e o de planejamento/programação, bem como em relação às noções correlatas de controle, monitoramento, seguimento ou acompanhamento. Em nosso entendimento, apesar da tentativa dos autores de esclarecer as diferenças, observa-se ainda na literatura uma certa sobreposição na definição destes conceitos, indicando que persiste algum grau de confusão entre a avaliação e estes termos.

A mesma imagem de circularidade que habitualmente é utilizada para descrever o processo de planejamento é também utilizada para descrever o processo de avaliação.

Há trinta anos atrás, quando o campo da avaliação tomou impulso, SUCHMAN fez uma caracterização da processo avaliativo que é em tudo e por tudo superponível ao processo de planejamento/programação. Utilizando a imagem do círculo, o autor coloca como primeira fase do processo a definição dos objetivos e metas do programa, bem como a definição dos critérios para medir o alcance dos mesmos. Em seguida, seriam identificadas as atividades necessárias para alcançar as metas, fase que ele denomina explicitamente de “planejamento do programa”. Segue-se a fase de realização destas atividades, denominada de “operação do programa”, e por último, a avaliação dos efeitos das atividades, na fase de “avaliação do programa”, que traz de volta o processo para o campo dos valores, onde foram definidos os objetivos (Suchman: 1967, 34).

Fica evidente que esta caracterização do processo avaliativo é em tudo e por tudo superponível ao que KNOX et al. denominam de “princípio circular” ou “círculo cibernético” do planejamento em saúde, cujo primeiro ponto é a “análise da situação” inicial e o último é a “avaliação” da situação final (Knox et al.: 1986, 13).

Segundo AGUILAR&ANDER-EGG, a confusão entre os conceitos de “avaliação” e “programação” é bastante comum. Porém, embora a prática da avaliação inclua também as

quatro fases da “estrutura básica de procedimento”, quais sejam - o diagnóstico, a programação, a execução e a avaliação -, esta prática não se confunde com a programação propriamente dita. A idéia de que “tudo é programar” ou “tudo é avaliar” seria um “erro metodológico e conceitual” muito comum, cuja origem está no corporativismo profissional de programadores e avaliadores (Aguilar&Ander-Egg: 1994, 22).

A tarefa de estabelecer a diferença entre o planejamento (ou programação) e a avaliação fica por conta de ANDER-EGG:

“Se planejar é introduzir organização e racionalidade na ação para a consecução de determinadas metas e objetivos, a avaliação é um modo de verificar essa racionalidade, medindo o cumprimento - ou perspectiva de cumprimento - dos objetivos e metas previamente estabelecidos e a capacidade para alcançá-los”
(Ander-Egg: 1984, 20).

Acreditamos que a confusão tenha ocorrido devido à própria história do surgimento da Pesquisa Avaliativa, essencialmente vinculada ao desenvolvimento dos programas sociais nos EUA na década de 60, conforme veremos no histórico da avaliação. Nesta fase inicial, a avaliação separou-se das funções de planejamento e implantação de programas, e passou a ser considerada “uma função em si e para si mesma” (Weilenmann: 1980, 24). Certamente por influência da matriz conceitual do planejamento de programas, os estudos avaliativos eram essencialmente “estudos baseados em objetivos”, ou seja, a avaliação limitava-se a determinar se os objetivos foram ou não alcançados, e em caso positivo, concluía-se se houve ou não êxito no trabalho avaliado, que alguns autores classificam de “quase-avaliações” (Aguilar&Ander-Egg: 1994, 39). Ao diferenciar-se do planejamento como um campo específico de estudos, a avaliação levou consigo a visão sistêmica, daí que ambos os processos tem a mesma estrutura básica.

Atualmente, de uma forma geral, a avaliação é vista de dois modos:

- como a “última etapa ou fase da estrutura básica de um procedimento”, que é considerada como uma “avaliação final”;
- como um “processo permanente de controle operacional”, que é considerada como “avaliação parcial”.

No primeiro caso, a avaliação corresponde à fase final do círculo do planejamento, que ANDER-EGG desdobra em quatro momentos: 1) estudo-investigação-diagnóstico; 2) programação; 3) execução ou implementação; 4) avaliação. No segundo caso, a avaliação é entendida como um seguimento, que a ciência da Administração denomina “controle operativo” (Ander Egg: 1984, 24; 36-37).

É em relação ao segundo modo de ver a avaliação que se estabelece a confusão com as noções de “controle” e “acompanhamento”, “monitoramento” ou “seguimento”. Segundo AGUILAR&ANDER-EGG, estes são “termos similares” à avaliação, porém com “alcances diferentes” (Aguilar&Ander-Egg: 1994, 19). Entretanto, ele próprio denomina a avaliação que realiza-se durante a execução de um programa de “monitoramento”, que também seria conhecida como avaliação contínua, avaliação concorrente ou concomitante, e avaliação da gestão (p. 41).

SILVER faz a distinção entre as noções de “controle” e de “avaliação”. Ela apresenta as definições elaborada pelo Comitê Administrativo de Coordenação do grupo de trabalho para o desenvolvimento rural das Nações Unidas (CAC), de 1984, que tinha por finalidade “sistematizar a terminologia usada em nível internacional”. Embora considere que estas definições não sejam perfeitas, ela entende que ajudam a compreender a sutileza das diferenças entre as duas atividades (Silver: 1992, 197-198):

*“Controle: (vigilância/fiscalização) É a vigilância contínua e periódica da implementação de uma atividade (e seus distintos componentes) a fim de assegurar que a entrada (**input**), o processo de trabalho e a produção prevista (**output**) e outras ações necessárias funcionem conforme o previsto.”*

“Avaliação: É um processo que tenta determinar o mais sistemática e objetivamente possível a relevância, efetividade e impacto das atividades, tendo em vista seus objetivos. É uma ferramenta orientada para a ação e aprendizagem. É um processo organizativo que visa tanto melhorar as atividades em andamento quanto planejar o futuro e orientar a tomada de decisões.”

A autora utiliza a noção de tempo para acentuar as diferenças: o controle, que ela também denomina de “monitoramento”, pressupõe “um seguimento **longitudinal** no tempo”, enquanto a avaliação seria “uma espécie de corte **transversal** e de visão mais ampla [grifos nossos]”. Ela esclarece que a avaliação no sentido de “controle” tem por fim “identificar os problemas à medida que estes apareçam, permitindo assim corrigir as distorções e assegurar o bom desenvolvimento das atividades” (Silver: 1992, 198).

AGUILAR&ANDER-EGG, que antes haviam denominado a avaliação durante a execução de “monitoramento”, entendem que, embora seja parecido, o termo controle é “de natureza bem diferente” da avaliação. Para ele, o controle seria simplesmente a “verificação” de algo, enquanto a avaliação seria um “juízo” acerca deste algo. Ouçamos o autor:

“Uma coisa é constatar o que acontece (controle) e outra julgar por que acontece o que acontece (avaliação), quer se trate de atividades, recursos ou resultados”
(Aguilar&Ander-Egg: 1994, 19).

COHEN&FRANCO, por sua vez, diferenciam a avaliação do “acompanhamento” ou “monitoramento”, dado como sinônimos. Eles apresentam uma definição de “acompanhamento”, também formulada pela ONU em 1984, que corresponde à definição de “controle” apresentada por Silver. Como esta autora utiliza o termo “monitoramento” como sinônimo de “controle”, concluimos, por dedução, que a noção de “acompanhamento” deve ser entendida como sinônimo das anteriores.

Embora também destaquem o aspecto temporal na diferenciação entre os conceitos de avaliação e acompanhamento, os autores chamam atenção para o aspecto da internalidade ou externalidade destes procedimentos em relação ao que está sendo avaliado:

*“Assim, enquanto o acompanhamento ou monitoramento é uma atividade gerencial **interna** que se realiza durante o período de execução e operação, a avaliação pode ser realizada tanto antes ou durante a implementação, como ao concluir a mesma ou mesmo algum tempo depois, para quando se prevê que o projeto provocou todo o seu impacto. Por outro lado, a avaliação tanto **interna***

como externa se ocupa com o modo, em que medida e por que tais pessoas foram beneficiadas [grifos nossos]” (Cohen&Franco: 1994, 77).

AGUILAR&ANDER-EGG preferem a utilização do termo “seguimento”, por ter um uso mais generalizado nos países de língua espanhola, ao invés de “monitoramento”, que seria uma tradução literal do inglês “monitoring”. O seguimento seria um “processo analítico que, mediante um conjunto de atividades, permite registrar, compilar, medir, processar e analisar uma série de informações que revelam o curso ou desenvolvimento de uma atividade programada”. Isto forneceria “informação retroalimentada” que permite ajustar o programa, sendo que o aspecto central do conceito é a continuidade e periodicidade (Aguilar&Ander-Egg: 1994, 20-21).

Finalizando, podemos resumir o que vimos sobre o conceito e na tarefa de avaliar nos seguintes aspectos:

- avaliar é uma atividade intrínseca da vida social;
- no campo da ciência, a avaliação é entendida basicamente como um tipo de pesquisa aplicada, dirigida para verificar o efeito das ações que visam a mudança social, através da “medição” destes efeitos;
- os dois objetivos fundamentais da avaliação são produzir conhecimento sobre a ação e contribuir para o processo de formulação de políticas e tomada de decisões;
- o processo avaliativo ocorre em um contexto de disputa entre interesses que se colocam na esfera política;
- o ponto de partida para qualquer avaliação é uma insatisfação com o **satus quo**;
- a atitude fundamental na avaliação é a atitude crítica;
- o processo avaliativo é atravessado pela disputa entre a subjetividade e a objetividade;
- os termos fundamentais da definição de avaliação são “juízo”, “valor” e “mérito”, e a noção fundamental do processo avaliativo é a de “retroalimentação”.

1.2. Os objetos da avaliação em saúde: o quê avaliar.

Como vimos anteriormente, o processo de avaliação faz parte da vida cotidiana, de forma que tudo que existe pode ser avaliado: um objeto, um ser humano, uma atividade, um programa social, uma idéia. Entretanto, no campo social, o objeto privilegiado da avaliação é a ação organizada, ou seja, as práticas ou intervenções sociais.

No campo específico da avaliação em saúde, o objeto da avaliação também é definido de forma ampla, podendo “desdobrar-se em tantas quantas forem as concepções sobre saúde e práticas de saúde”, de forma que “os recortes possíveis são inúmeros”, variando desde a avaliação de um procedimento isolado ou do cuidado individual até a avaliação ampla do sistema econômico e social (Silva & Formigli: 1994, 81).

Neste trabalho, propomos fazer dois grandes recortes da realidade para delimitar os objetos da avaliação no campo da saúde: o primeiro, refere-se à natureza deste objeto - estados de saúde ou intervenções em saúde; o segundo, refere-se ao nível de análise no qual são enfocados estes objetos - nível individual ou nível agregado.

Ao fazer o primeiro recorte, partimos do pressuposto de que os objetos da avaliação no campo da Saúde Coletiva deverão ser, evidentemente, os objetos deste campo. Para FRENK, os objetos de investigação e de práticas da “nueva salud pública” seriam de duas ordens: “las **condiciones** de salud de las poblaciones” e “la **respuesta** social organizada a esas condiciones” (Frenk: 1992: 77). Estes tem sido os objetos de estudo no campo da Saúde Pública. Em seu histórico sobre a pesquisa avaliativa, SUCHMAN lembra que, inicialmente, os esforços da Saúde Pública dirigiam-se para a avaliação das condições de saúde das comunidades e, posteriormente, com a expansão dos serviços públicos, para avaliar a efetividade dos programas (ou intervenções) de saúde (Suchman: 1967, 12-15). Mais recentemente, BATTISTA e colaboradores classificam os estudos no campo da saúde em estudos que enfocam os estados de saúde (“health states”) e estudos que enfocam as intervenções em saúde (“health interventions”) (Battista et al.: 1989, 1156).

Entretanto, tradicionalmente, o primeiro recorte no campo da saúde refere-se à diferenciação entre as ações voltadas para o indivíduo (nível individual) e aquelas voltadas

para os grupos ou as populações (nível agregado ou coletivo), que tem diferenciado a Clínica da Saúde Pública.

Ao manter este recorte, não estamos querendo endossar a clássica separação entre a dimensão individual e coletiva, que tem justificado um discurso de oposição entre a Clínica e a Saúde Pública. Entendemos da mesma forma que SOUZA CAMPOS, para quem a separação individual/coletivo é “mecânica e artificial”, sendo fruto de um “equivoco conceitual”. Estas duas dimensões da realidade estão relacionadas, devendo ser abordadas sempre em conjunto, dentro de uma proposta de “integração sanitária”. Segundo o autor

“...a Clínica e a Saúde Pública tratam do mesmo problema, ainda que enfatizando diferentes aspectos da questão saúde/doença e dando origem a modelos de atenção circunstancialmente, mas não necessariamente, diferentes” [grifos nossos] (Souza Campos: 1991, 68).

Além disto, este autor demonstra, cabalmente, a influência da Clínica nas práticas programáticas da Saúde Pública. Para ele, tanto a Saúde Pública baseia-se no conhecimento clínico, quanto a Epidemiologia, tradicionalmente considerada a disciplina básica das práticas coletivas, está estreitamente vinculada às práticas individuais da Clínica (Souza Campos: 1991, 67-68, 61).

A semelhança entre a Clínica e a Saúde Pública no que diz respeito aos objetos da avaliação em saúde já havia sido apontada há muito tempo atrás por Greenberg e Mattison¹⁰, conforme vemos em SUCHMAN:

“Greenberg e Mattison traçam um paralelo direto entre o processo clínico, dirigido à diagnosticar o que está errado e à prescrever um tratamento, o qual pode então ser avaliado em termos da melhora do paciente, e o processo na saúde pública de determinar as necessidades da comunidade, desenvolver programas de saúde pública para atender estas necessidades, e acompanhar estes programas para avaliar seu sucesso ou fracasso” (Suchman: 1967, 39).

A necessidade de superar a divisão Clínica/Saúde Pública, fundamentada na divisão individual/coletivo não é apenas uma questão teórica. A partir década de 70, as políticas de extensão de cobertura dos serviços de saúde foram criando a necessidade de desenvolver um “novo modelo tecnológico” da Saúde Pública, que incorporasse, em um mesmo modelo assistencial, a lógica da Epidemiologia tradicional, voltada para o desenho das ações preventivas coletivas e a assistência médica individual curativa (Nemes In Schraiber: 1990, 75).

Em decorrência da forma como foram expandindo-se os sistemas de saúde em toda a América Latina, a “nova” Saúde Pública transpõe esta divisão, caracterizando-se pela incorporação das atividades da clínica, voltadas para a assistência individual curativa, dentro das tradicionais ações de promoção e prevenção, voltadas para as populações ou segmentos populacionais (Frenk: 1992; OPS/OMS: 1984).

Especificamente no caso do Brasil, a integralidade da assistência à saúde é uma diretriz do próprio Sistema Único de Saúde, deixando portanto de ser uma especulação teórica e passando a ser uma exigência legal.

Por tudo isto, nossa proposta de manter o recorte individual/coletivo é apenas para chamar atenção que ele diferencia dois níveis de análise na pesquisa avaliativa: o nível individual e o nível agregado.

Enfim, em termos amplos, podemos considerar que os objetos da avaliação no campo da Saúde Coletiva são de duas ordens - os estados (ou condições) de saúde das comunidades ou populações e as intervenções (ou práticas) de saúde sobre estas comunidades ou populações -, e que este objeto pode ser abordado em dois níveis de análise - o nível individual e o nível agregado.

Uma vez definido os grandes recortes que delimitam os objetos da avaliação em saúde, passamos a tratar da forma como estes objetos vem sendo abordados na literatura. Para orientar nossa análise, utilizamos dois referenciais: o referencial tradicional da literatura clássica, que trabalha com a idéia do “círculo avaliativo”, voltado para a avaliação das intervenções coletivas na esfera social, e um referencial mais recente, específico do campo da saúde, que coloca os estudos avaliativos no papel intermediário entre os estudos que enfocam

¹⁰GREENBERG, Bernard & MATTISON, Berwyin F. The Whys and Wherefores of Program Evaluation.

os estados de saúde e os que enfocam as intervenções de saúde, em qualquer nível de análise. Além disto, buscamos demonstrar como os diferentes objetos da avaliação são tratados, preferencialmente, pelas disciplinas que compõe o campo da Saúde Coletiva.

1.2.1. Os estados de saúde

Começamos pela primeira ordem de objetos no campo da avaliação em Saúde Coletiva - os estados ou condições de saúde.

Na literatura geral sobre avaliação, vemos que o diagnóstico “em si mesmo constitui um processo de avaliação da realidade existente”, sendo o primeiro passo do processo avaliativo “que irá se prolongar através da execução de um projeto e da medição do impacto deste na transformação desta realidade” (Vergara: 1980, 81). Como vimos na seção referente ao conceito de avaliação, o processo avaliativo reproduz o “princípio circular” ou “círculo cibernético” do planejamento, cujo primeiro ponto é a “análise da situação” inicial e o último é a “avaliação” da situação final (Knox et al.: 1986, 13). Alguns autores utilizam inclusive os termos “avaliação antes”, inicial ou *ex ante* para referir-se à avaliação que é feita “antes de tomar a decisão de empreender um programa ou projeto, ou de aprovar a sua realização”, com o objetivo avaliar a conveniência ou não de levar a cabo o empreendimento, a partir da análise da sua pertinência, viabilidade e eficácia potencial (Aguilar&Ander-Egg: 1994, 41).

Tradicionalmente na Saúde Pública, esta questão é tratada nos famosos “diagnósticos de saúde”, “diagnóstico de comunidade”, “diagnóstico populacional”, “diagnóstico da situação de saúde”, “perfil epidemiológico”, que nada mais são do que “avaliações” das condições ou estado de saúde de comunidades ou populações. Utilizando exemplos deste campo, SUCHMAN discute as características dos roteiros de avaliação (“evaluation guides” ou “evaluation schedules”) destinados a estes diagnósticos, denominados em Inglês por termos como “community self-surveys of public service programs”, “health survey”, “community

appraisal forms”, “surveying community needs” e “public health diagnosis”(Suchman: 1967, 8, 13, 24). Na língua espanhola, os termos utilizados referem diretamente à avaliação: “evaluación de las condiciones” ou “evaluación de los estados de salud”.

Os objetivos do diagnóstico seriam descrever, explicar e prever a realidade, e emitir um juízo sobre a mesma, isto é, se é satisfatória ou não (OPS/OMS: 1965, 24, 42, 47; Cohen&Franco: 1994, 79). A “explicação” tem uma importância fundamental no processo de diagnóstico, pois é com base no “modelo causal” que fundamenta esta explicação que serão elaborados os “projetos de mudança”, ou seja, as intervenções (Cohen&Franco: 1994, 79). Na Saúde Pública tradicional, este modelo explicativo é proporcionado pela Epidemiologia, refletindo o chamado “modelo biomédico” que predomina na Clínica, que em consequência, define o tipo de mudança que pode ser proposta, e o tipo de intervenção para alcançá-la.

Portanto, o diagnóstico não consiste somente da identificação e quantificação dos problemas de saúde - geralmente referidas como “avaliação das necessidades de saúde” da população -, mas também na hierarquização e escolha daqueles considerados prioritários para sofrer uma intervenção, ou seja, aqueles que serão objeto de um plano de ação ou de um programa de saúde (Paim: 1979, 43). Ou seja, a “avaliação de necessidades” é apenas um dos vários aspectos do diagnóstico (Aguilar&Ander-Egg: 1994, 22).

O ponto fundamental do diagnóstico epidemiológico em Saúde Pública é a definição de grupos de população prioritários para receber atenção em saúde, através da identificação dos grupos cujo estado de saúde está mais exposto à riscos (Galvis: 1984, 240).

É preciso lembrar que, embora um diagnóstico seja “algo mais do que fazer uma pesquisa”, não é possível fazer um diagnóstico “sem base de um estudo prévio” (Aguilar&Ander-Egg: 1994, 51).

Embora o uso do termo “diagnóstico” para se referir à “avaliação inicial” não seja exclusivo da área da saúde, entendemos que neste campo o uso do termo reflete a influência da Clínica nas práticas programáticas da Saúde Pública, que vimos anteriormente.

Esta influência fica muito clara na analogia que FERREIRA estabelece “entre os passos que se seguem para a solução de um problema de saúde individual (problema clínico) e para a solução de um problema de saúde pública.” Utilizando a mesma imagem circular dos processo de planejamento e de avaliação, que vimos na seção anterior, o autor demonstra que

os indicadores epidemiológicos devem ser usados tanto na fase do diagnóstico inicial, quanto após a intervenção, na fase de avaliação, “daí por que se pode dizer que a avaliação é, de certa forma, um novo diagnóstico” (Ferreira: 1990, 34,43). Vemos que o autor estende o uso do termo diagnóstico à própria “avaliação final”, quando habitualmente o termo é utilizado para referir-se à “avaliação inicial”.

Por esta razão, suspeitamos que a utilização do termo “diagnóstico” na área do Planejamento em Saúde tenha ocorrido por analogia às práticas clínicas, onde utiliza-se tanto a expressão "avaliação diagnóstica", no início do processo, quanto a expressão “avaliação terapêutica”, no final do processo. Não fosse essa influência da clínica médica, talvez os termos utilizados fossem simplesmente "avaliação inicial" (antes da intervenção) e "avaliação final" (após a intervenção) do estado de saúde da população.

Na Saúde Pública, tradicionalmente, a tarefa avaliativa do “diagnóstico de saúde” fica ao cargo da Epidemiologia. SUCHMAN lembra os “precursores da moderna estatística em saúde pública”, que no século XVII, através da estatística vital e dos dados de morbidade e mortalidade, “ofereceram a primeira possibilidade de uma medida objetiva do estado de saúde de uma nação” (Suchman: 1967, 13).

Além destas tradicionais fontes de dados, coletadas regularmente pelos serviços de estatística vital e epidemiológica, a Epidemiologia também promove pesquisa para coleta de dados primários sobre os estados de saúde, através dos conhecidos “inquéritos de base populacional” ou “inquéritos domiciliares”.

Mais recentemente, os estudos sobre a avaliação das condições de saúde em nível comunitário ou populacional tem exigido a concorrência de outras disciplinas, como as “ciências comportamentais, sociologia, antropologia e economia” (Battista et al.: 1989, 1157). Especialmente na linha dos estudos sobre qualidade de vida esta contribuição interdisciplinar tem sido necessária.

Também em nível individual, a avaliação das condições de saúde tem contado com a contribuição desta disciplina, através da pesquisa em Epidemiologia Clínica, seja realizada por clínicos seja por epidemiologistas (Battista et al.: 1989, 1157).

1.2.2. As intervenções em saúde

Quanto à segunda ordem de objetos da avaliação - a resposta social organizada às condições de saúde da população - consideraremos dois aspectos:

- em primeiro lugar, que vemos na literatura o uso dos termos “práticas” e “intervenções” para referir-se a esta resposta social, revelando diferentes percepções sobre este objeto de avaliação;

- em segundo lugar, que os autores consultados referem a existência de diferentes “níveis de complexidade”, “níveis de realidade” ou mesmo “níveis de objetivos” destas intervenções ou práticas, o que define diferentes objetos da avaliação em cada nível.

Em relação ao primeiro aspecto, observamos que a utilização do termo “práticas” ocorre preferencialmente na literatura latino-americana, revelando a influência do pensamento marxista no campo da Medicina Social, enquanto o termo “intervenções” parece ser o preferido na literatura anglo-saxônica sobre avaliação, especialmente a norte-americana, onde tradicionalmente tem predominado o pensamento funcionalista¹¹.

Embora tanto numa quanto noutra literatura estes termos sejam usados indistintamente, a questão obviamente não se restringe a uma preferência por palavras, mas reflete concepções diferentes da resposta social aos problemas de saúde. Ou seja, implica em diferentes formas de caracterizar este objeto da avaliação em saúde.

O termo “**prática**” reflete uma visão dialética da ação humana, e está referido ao conceito marxista de “práxis”. Embora de difícil caracterização, o essencial deste conceito é a noção de que a teoria (ou saber) está indissociavelmente vinculada à prática (ou trabalho) sobre um concreto. Neste sentido, a “práxis” compreende um “momento laborativo” e um “momento existencial” (Minayo: 1992, 245). Ou seja, a ação (trabalho) está dialeticamente vinculada à reflexão (teoria).

A utilização deste conceito fundamenta duas concepções básicas sobre as quais se pretende diferenciar a Saúde Coletiva da Saúde Pública tradicional: a primeira, que qualquer

¹¹As principais diferenças entre estas correntes de pensamento são tratadas na seção 1.3.1.

área de conhecimento é definida ao mesmo tempo como “campo de saberes e de práticas”; a segunda, que as práticas de saúde são vistas fundamentalmente como uma “prática social”, sofrendo portanto as mesmas influências do contexto econômico, político e social que qualquer outra prática, não podendo ser entendidas como práticas “técnicas”, “neutras”, livre destas influências (Donnangelo: 1983; Merhy, 1992; Gonçalves: 1994).

Nesta concepção, o aspecto fundamental da avaliação das práticas em saúde é determinar sua relação com o contexto social, econômico e político, onde as macroteorias sociais desempenham um papel fundamental. Em decorrência desta concepção, a vertente latino-americana de avaliação tem buscado demonstrar os “significados econômicos, políticos e ideológicos” das relações entre as práticas e as necessidades de saúde, apontando a “especificidade destas relações em sociedades concretas” (Silva & Formigli: 1994, 80).

O termo “**intervenção**”, por sua vez, é utilizado de forma intercambiável com o termo “programa” na literatura anglo-saxônica. Além das questões semânticas específicas da língua inglesa¹², podemos afirmar que, de uma maneira geral, estes termos refletem uma visão mais funcional da ação humana, traduzida em uma concepção sistêmica da “ação programática”.

Voltando a SILVA & FORMIGLI, vemos que nesta vertente da avaliação está mais voltada para a “capacidade das práticas modificarem uma dada situação de saúde”, onde o aspecto fundamental é a discussão sobre as “características” e os “efeitos” destas práticas ou intervenções (Silva & Formigli: 1994, 80).

Na Saúde Pública tradicional, a ação programática consiste no recorte de problemas sanitários específicos, que embora sejam definidos “supostamente” por critérios técnicos e epidemiológicos, na verdade resultam de uma seleção feita “a partir de pressões sociais, políticas ou econômicas”, e respondem a uma “indisposição do Estado em gastar em saúde, decorrendo daí a necessidade de recortes” (Souza Campos: 1991, 59-60).

Entretanto, observamos que, na literatura canadense mais recente, vem ocorrendo uma mudança no conceito de programa, que implica em dois aspectos: o primeiro, é a adoção de uma concepção ampla do mesmo, que acaba culminando com a inclusão do próprio contexto como um dos componentes do programa; o segundo, é a superação da noção de programa como uma prática exclusivamente “coletiva”, onde assume-se que a avaliação de programas

implica em um nível individual e um nível agregado de análise, onde incluem-se dimensões micro e macro teóricas. Tentaremos demonstrar estas idéias a seguir.

Inicialmente, CLEMENHAGEN&CHAMPAGNE caracterizam uma noção ampla de programa para englobar os objetos da avaliação em saúde, que seriam os serviços, as atividades e as intervenções. Embora referindo-se ao âmbito hospitalar, a idéia geral fica clara:

"Para propósitos avaliativos, qualquer serviço hospitalar, atividade ou intervenção pode ser descrito em termos de objetivos, recursos, serviços e efeitos - e estes elementos tornam possível submeter o serviço, atividade ou intervenção à um escrutínio avaliativo na forma de um 'programa'" (Clemenhagen&Champagne: 1986, 383).

Posteriormente, BATTISTA et al. definem uma intervenção como “um conjunto de atividades especificamente designadas para modular os estados de saúde”, que podem ter objetivos preventivos, curativos, reabilitatórios ou paliativos (Battista et al: 1989, 1157).

Algum tempo depois, CONTANDRIOPOULOS et al. retomam o conceito de intervenção, caracterizando-o através dos componentes típicos dos programas:

“Uma intervenção é constituída pelo conjunto de meios (físicos, humanos, financeiros, simbólicos) organizados em um contexto específico, em um dado momento, para produzir bens ou serviços com o objetivo de modificar uma situação problemática” (Contadriopoulos et al.: 1992).

Segundo estes autores, uma intervenção caracteriza-se por ter cinco componentes: “(1) Objetivos; (2) Recursos; (3) Serviços, bens ou atividades; (4) Efeitos e (5) Contexto preciso em um dado momento” (Contadriopoulos et al.: 1992).

Embora esta caracterização de intervenção, aparentemente, nada mais seja do que a definição clássica de programa, uma análise mais detalhada revela a inclusão de duas noções fundamentais: a noção de “contexto” e a noção da especificidade histórica do mesmo -

^{1,2}Segundo comunicação pessoal da Prof^a Zulmira Hartz, na língua inglesa os termos “intervention”, “program” e “treatment” são utilizados como sinônimos.

“contexto preciso em um dado momento”. Entendemos que isto significa, de certa forma, a incorporação das críticas feitas pelos autores marxistas à noção funcionalista de programa, que enfoca apenas a racionalidade interna do mesmo, omitindo seu contexto social, político e econômico e sua especificidade histórica. Este entendimento é de certa forma confirmado por HARTZ, que aponta a incorporação da “perspectiva [macroteórica] global de economia política” na avaliação de programas (Hartz: 1996).

Em outro trabalho, estes mesmos autores retomam uma definição mais frouxa de intervenção, caracterizando-a como qualquer “conjunto de atividades pressupondo a elaboração de meios organizados de maneira coerente no tempo e no espaço, para modificar uma situação” (Contandriopoulos et al.: 1994, 21).

Em relação ao segundo aspecto - a integração dos níveis micro e macro teóricos na avaliação de programas - , veremos melhor na seção que refere-se à abordagem integrada da avaliação em saúde. Por ora, destacamos o trabalho de CLEMENHAGEN&CHAMPAGNE, que acreditamos ter sido um dos pioneiros neste aspecto, onde estes autores referem-se aos aspectos “clínicos e não clínicos” dos programas, cuja avaliação implica em um “nível individual” e um “nível agregado” de análise (Clemenhagen&Champagne: 1986, 386).

As colocações de HARTZ&CAMACHO, que estão inspirados na literatura canadense, demonstram claramente como esta nova perspectiva de programa integra o nível individual e coletivo:

“O enfoque de Programa de Saúde ultrapassa os limites da prestação dos serviços, como respostas dirigidas aos problemas de saúde individuais, sem no entanto excluí-los, colocando-os na perspectiva das intervenções coletivas que interagem no processo saúde-doença e exigem uma abordagem micro/macro contextual dos profissionais com ela envolvidos” (Hartz&Camacho: 1996)

Encerramos aqui a análise das diferentes concepções da “resposta social organizada” enquanto objeto de avaliação em saúde. Passamos agora a tratar dos níveis de análise destes objetos.

Tanto na vertente das “práticas” quanto das “intervenções”, todos os autores consultados referem-se à existência de diferentes **níveis** destas intervenções ou práticas.

Embora estabeleçam diferentes gradações para estes níveis, as classificações dos diferentes autores acabam abrangendo a mesma variedade de objetos. Por outro lado, ainda que não seja explicitado, percebe-se que o eixo individual/coletivo divide as intervenções entre aquelas mais simples - as práticas individuais de prestação de cuidados - e as intervenções mais complexas - as práticas coletivas de prestação de serviços.

Iniciamos com SILVA&FORMIGLI, que distinguem duas dimensões básicas das práticas em saúde - individual (cuidado individual) e coletiva (sistemas de atenção à saúde) - reforçando o recorte individual/coletivo:

“...as práticas de saúde, à semelhança de outras práticas sociais, podem constituir-se em objeto de avaliação nas suas diversas dimensões, seja enquanto cuidado individual, seja nos níveis mais complexos de intervenção e organização, como políticas, programas, serviços ou sistemas” (Silva&Formigli: 1994, 81).

No mesmo artigo de revisão destas autoras, vemos ainda duas outras formas de classificar os níveis das intervenções em saúde.

PAIM refere-se a oito formas e níveis de “organização dos serviços de saúde: a) ações de saúde; b) serviços de saúde; c) estabelecimentos de saúde; d) distrito sanitário; e) instituições de saúde; f) sistema de saúde; g) sistema político; h) sistema econômico e social” (Paim apud Silva & Formigli: 1994, 81). DONABEDIAN distingue quatro níveis de “organização dos provedores do cuidado médico: a) o profissional individual; b) diversos profissionais; c) uma equipe organizada; d) plano, instituição e sistema” (Donabedian apud Silva & Formigli: 1994, 81).

BATTISTA et al. discriminam quatro “níveis de complexidade” dos objetos da pesquisa em saúde, entre as quais incluem-se as pesquisas com propósito avaliativo: técnicas, práticas, programas ou políticas. Os objetos de avaliação em cada nível são definidos da seguinte maneira:

- uma técnica é o “procedimento que resulta do desenvolvimento tecnológico” e é utilizado pelos profissionais de saúde (health-care professionals);

- uma prática é um conjunto organizado de “atividades ou procedimentos” que são “implementados individualmente” por alguém que presta cuidados de saúde (health-care provider);

- um programa é “um conjunto de atividades usualmente encaminhadas por uma organização ou instituição”, tendo como alvo uma grupo de indivíduos ou clientes;

- uma política é “um conjunto de ações determinadas por um corpo legislativo ou uma autoridade administrativa” (Battista et al.: 1989, 1157).

Observa-se que as técnicas e as práticas são intervenções conduzidas individualmente e dirigidas para o nível individual, enquanto os programas e as políticas são conduzidas por formas mais complexas de organização do trabalho, e dirigidas para os níveis comunitário e populacional.

Inspirados nos autores anteriores, CONTANDRIOPOULOS et al. alteram um pouco a classificação, passando a vincular a idéia de política à idéia de organização - “organizações políticas”, e acrescentando um nível superior de análise - os “sistemas de saúde”. A classificação destes autores inclui portando cinco níveis de intervenções em saúde: técnicas, práticas, programas, organizações políticas e sistemas de saúde (Contandriopoulos et al.: 1994, 21).

A cada nível de complexidade da “resposta social aos problemas de saúde” corresponde um tipo de pesquisa em avaliação, que pode envolver uma ou mais disciplinas. Além disto, coerentemente com as diferentes concepções deste objeto de estudo, a literatura enfoca diferentes aspectos do mesmo: a literatura latino-americana enfoca principalmente a relação das práticas de saúde com o seu contexto, enquanto a literatura anglo-saxônica enfoca principalmente as características internas das intervenções e os seus efeitos (impacto).

No primeiro caso, tradicionalmente o enfoque tem dirigindo-se para as macro-análises das instituições públicas, das políticas e dos sistemas de saúde, buscando estabelecer relações entre seu funcionamento e o sistema político, econômico e social. A maior parte dos autores estão vinculados às áreas de Ciência Sociais e Planejamento em saúde¹³. A análise da prestação de cuidados em nível individual vinha sendo abordada em trabalhos que enfocam a

¹³Nesta linha, destacamos alguns trabalhos clássicos: DONNANGELO, Maria Cecília & PEREIRA, Luiz. *Saúde e Sociedade*. São Paulo, Livraria Duas Cidades, 1976; LUZ, Madel Terezinha. *As instituições médicas no Brasil*.

prática médica como prática social liberal, autônoma, de prestação de cuidados em nível individual, apontando os fatores contextuais e históricos que condicionam a mesma¹⁴. Mais recentemente, alguns estudos tem dirigindo-se para a análise do “processo de trabalho em saúde” enquanto “prática tecnológica”, discutindo a “organização tecnológica do trabalho” como prática coletiva dentro dos serviços de saúde¹⁵.

Na vertente das “intervenções”, observa-se que a Epidemiologia tem um papel predominante. Embora tradicionalmente esteja vinculada ao diagnóstico de saúde, observa-se uma tendência desta disciplina em voltar-se para a avaliação das intervenções em saúde, desde práticas individuais, serviços e programas até a avaliação das políticas de saúde, que veremos nos aspectos históricos.

BATTISTA e colaboradores delimitam um tipo de pesquisa para cada nível das intervenções. A avaliação de técnicas¹⁶ e práticas ficaria ao cargo da pesquisa em “epidemiologia clínica”, ambas em nível individual. A avaliação de programas e de políticas de saúde, em nível comunitário-populacional, exigiria, pela complexidade das intervenções neste nível, o uso de uma “variedade de disciplinas que inclui epidemiologia, ciências do comportamento, sociologia, antropologia, economia, teoria organizacional e análise de decisão” (Battista et al.: 1989, 1156).

Porém, HARTZ&CAMACHO entendem que, embora a pesquisa avaliativa receba contribuição de diferentes disciplinas, as linhas de investigação ainda espelham as divisões, que devem ser minimizadas, entre a “pesquisa biomédica e organizacional; a atenção individual e coletiva, as dimensões macro e micro teóricas da modelagem” dos programas, o que dificulta apreciá-los em sua complexidade (Hartz&Camacho: 1996).

Rio de Janeiro, Graal, 1982; MERHY, Emerson Elias. *A Saúde Pública como Política. Um estudo de formuladores de políticas*. São Paulo, Ed. HUCITEC, 1991.

¹⁴Nesta linha destacamos: DONNANGELO, Maria Cecília Ferro. *Medicina e Sociedade: o Médico e seu Mercado de Trabalho*. São Paulo, Pioneira, 1975; SOUZA CAMPOS, Gastão Wagner. *Os médicos e a política de saúde*. São Paulo, Ed. HUCITEC, 1988.

¹⁵O trabalho pioneiro nesta linha foi a Tese de Doutorado de Ricardo Bruno Mendes Gonçalves, de 1986, publicada recentemente: GONÇALVES, R.B.M. *Teconologia e Organização Social das Práticas de Saúde: características tecnológicas do processo de trabalho na rede estadual de Centros de Saúde de São Paulo*. São Paulo, Ed. HUCITEC/ABRASCO, 1994.

¹⁶As técnicas comportam ainda um nível molecular-celular e um nível tissular, cujo estudo ficaria a cargo da pesquisa básica (Battista et al.: 1989, 1158).

Em qualquer nível das intervenções em saúde, a pesquisa em avaliação busca, usualmente, verificar a eficácia, a efetividade e a eficiência das mesmas. Ou seja, busca verificar se a intervenção funciona, e a que custo (Battista et al.: 1989, 1158).

De uma forma geral, podemos afirmar que a maior parte dos estudos em avaliação no campo da saúde referem-se ao cuidado individual, aos serviços (hospitais, centros de saúde) ou aos programas (Silva & Formigli: 1994, 81).

No Brasil, segundo SILVER, a pesquisa em avaliação de serviços tem, tradicionalmente, dirigindo-se para dois objetos:

- a avaliação da “qualidade dos serviços prestados por um médico individual, ou dentro das quatro paredes de uma instituição, através de critérios técnico-profissionais”;
- a avaliação da “produção numérica das unidades ou dos médicos”, como parte do “planejamento normativo dos serviços ou das atividades de auditoria do sistema”.

Ambas as perspectivas deixam de lado “os problemas fundamentais de equidade e distribuição de bens e serviços de saúde” e as “necessidades da população”. Além disto, poucos instrumentos de avaliação “olham” a rede como um todo, e não apenas “para um único serviço” (Silver: 1992, 198-99).

1.2.3. O papel intermediário da avaliação

Vimos até aqui, como a literatura sobre avaliação em saúde vem abordando os objetos da Saúde Pública - os estados de saúde e as intervenções de saúde -, tanto no nível individual como no nível agregado de análise.

Voltando à imagem de circularidade do processo avaliativo, que vimos na parte referente ao conceito da avaliação, poderíamos dizer que, idealmente, o ciclo da avaliação em saúde tem seu **início** na avaliação dos estados de saúde (situação inicial) e seu **fim** na avaliação dos efeitos das intervenções sobre este estado (situação final). Entretanto, a avaliação tem um papel fundamental também no **meio** do processo, isto é, durante a implantação da intervenção ou execução das atividades.

Este papel intermediário pode ser entendido de duas maneiras. Tradicionalmente, refere-se ao próprio acompanhamento e controle da execução da intervenção, que em geral fica ao cargo das áreas de Administração e Planejamento.

Entretanto, mais recentemente, a partir do trabalho de BATTISTA et al., a literatura canadense vem desenvolvendo a idéia de que a avaliação, enquanto pesquisa, teria um papel intermediário, seria o “pivô” que vincula as pesquisas sobre os estados de saúde com as pesquisas sobre as intervenções em saúde (Battista et al.: 1989, 1156).

Estes autores desenvolveram um complexo modelo para classificar as pesquisas em saúde, que utiliza três critérios: (1) o **foco da pesquisa** - os estados de saúde ou as intervenções de saúde; (2) o **nível da pesquisa** - no caso dos estudos sobre os estados de saúde, seriam os níveis molecular, tecidual, individual, comunitário-populacional, e no caso dos estudos sobre as intervenções de saúde seriam, como já vimos, os níveis das técnicas, práticas, programas e políticas; (3) o **propósito dos estudos** - desenvolvimento, descrição, explicação ou avaliação. O cruzamento destes três critérios delimita diversas áreas específicas de pesquisa, que no caso dos estudos que enfocam as intervenções e com propósito avaliativo, acabam confluindo para uma grande área de “pesquisa sobre o cuidado à saúde”, que engloba todos os níveis, como vimos anteriormente.

Ou seja, além de organizar o campo dos estudos em saúde, o modelo teórico proposto por estes autores definiu para a avaliação um papel que ultrapassa o tradicional foco nos dois objetos de estudo do campo - os estados de saúde e as intervenções de saúde - , colocando-a em um papel intermediário cujo foco é um objeto “virtual”, que localiza-se no espaço que vincula estes dois objetos.

Eles definem desta forma o papel intermediário da pesquisa em saúde com propósito avaliativo:

“A pesquisa avaliativa tem importância central na integração das pesquisas relacionadas à saúde porque é o pivô entre os estudos dos estados de saúde e os estudos das intervenções em saúde: os estados específicos de saúde levam ao desenvolvimento de intervenções desenhadas para modulá-los” (Battista et al.: 1989, 1158).

Posteriormente, CONTANDRIOPOULOS e colaboradores retomam este esquema teórico, caracterizando desta forma a pesquisa avaliativa em saúde: “A característica principal desta pesquisa consiste no fato dela ser o eixo entre a pesquisa sobre os estados de saúde e a pesquisa sobre as intervenções em saúde” (Contandriopoulos et al.: 1994, 21). Para ter uma visualização deste complexo modelo teórico, ver Anexo II.

Ou seja, o papel da pesquisa avaliativa seria justamente relacionar o efeito das intervenções com as mudanças nos estados de saúde, visto que elas foram “desenhadas para modulá-los”.

Voltando ao ciclo avaliativo, vemos que a avaliação dos estados de saúde está no início (avaliação inicial ou diagnóstico), e a avaliação dos efeitos da intervenção está no fim (avaliação final). Ao unir os dois objetos de estudo em saúde, a pesquisa avaliativa une o “início” das intervenções - o problema que a justifica - com o “fim” das intervenções - os seus efeitos -, buscando estabelecer uma relação causal entre a intervenção e as mudanças no estado de saúde. Veremos melhor estas questões na seção referente às metodologias.

Para encerrar, transcrevemos SILVA & FORMIGLI, que analisam a forma como a literatura vem tratando o papel intermediário da avaliação nas relações entre as necessidades de saúde (estados de saúde) e as práticas de saúde (intervenções em saúde):

“As relações entre as práticas e as necessidades de saúde podem ser apreendidas a partir de duas vertentes principais. A primeira delas consiste na busca dos seus significados econômicos, políticos e ideológicos, bem como da especificidade dessas relações em sociedades concretas. A segunda diz respeito à capacidade das práticas modificarem uma dada situação de saúde, atendendo ou não às necessidades de saúde de determinada população, ou seja, refere-se à discussão sobre suas características e seus efeitos (Silva & Formigli: 1994, 80).

Para as autoras, é nesta segunda vertente que “a temática da avaliação ganha relevância”, tanto diante da “possibilidade e necessidade” de intervenções que sejam capazes de mudar certos quadros sanitários, quanto diante das “dificuldades” enfrentadas por estas intervenções para alterarem os indicadores de morbi-mortalidade.

Entendemos que esta colocação demonstra o reconhecimento da história institucional do campo da avaliação, que veremos no histórico. A primeira vertente mencionada pelas autoras provavelmente refere-se à linha de avaliação mais desenvolvida na América Latina, qual seja, a análise crítica das políticas do campo da saúde, baseada em um referencial macroteórico. A segunda vertente provavelmente refere-se à avaliação de programas, cujo desenvolvimento deveu-se em grande parte aos estudos dos programas públicos voltados para a área social, que ocorreu na década de 60 na América do Norte, especialmente nos EUA, e que deram origem ao surgimento da chamada Pesquisa Avaliativa, conforme veremos nos aspectos históricos.

Resumindo, o objeto da avaliação no campo da Saúde Coletiva/Saúde Pública pode estar em qualquer ponto de um esquema conceitual no qual, partindo-se da avaliação inicial dos estados de saúde - que, se não correspondem ao desejado, caracterizam um problema a ser identificado na “avaliação de necessidades” -, encontra-se justificativas para promover uma intervenção - resposta social a estas necessidades -, cujos efeitos sobre a situação inicial serão posteriormente observados, para verificar se o objetivo desejado foi alcançado, qual seja, se houve uma mudança positiva no problema de saúde. Como vimos no conceito de avaliação, este esquema lógico tem origem na Teoria de Sistemas, cuja noção fundamental é a de retroalimentação. No contexto da pesquisa em avaliação, este esquema pode ser traduzido pela busca de relação causal entre a intervenção e o efeito observado nos estados de saúde.

1.3. Metodologia da avaliação em saúde: como avaliar.

Neste capítulo, não pretendemos descrever os passos através dos quais poderia conduzir-se uma pesquisa em avaliação. Nosso objetivo é abordar as questões mais gerais que determinam as opções metodológicas do campo, que apresentaremos em três seções.

Na primeira, tratamos da problemática do campo maior das Ciências Sociais, onde colocam-se as questões fundamentais da metodologia de pesquisa: objetividade *versus* subjetividade na produção do conhecimento, metodologias quantitativas *versus* qualitativas, validade do conhecimento produzido.

Na segunda seção, tratamos das diferentes tipologias que classificam os estudos em avaliação, apontando também alguns aspectos metodológicos específicos do campo. Buscamos ainda caracterizar o que entendemos ser as duas grandes linhas discursivas da avaliação no campo da Saúde Pública - a avaliação de programas e a avaliação de qualidade -, que vem sendo unificadas em um modelo teórico integrado para a avaliação em saúde, desenvolvido na literatura canadense.

Na terceira seção, buscamos caracterizar as principais categorias analítico-operacionais que norteiam o processo avaliativo.

1.3.1. Questões metodológicas gerais: a problemática das Ciências Sociais

Enquanto campo de pesquisa, a avaliação põe em evidência, de forma contundente, os grandes questionamentos metodológicos do campo das Ciências Sociais, da qual faz parte. O estudo dos fenômenos sociais coloca inevitavelmente uma problemática, de ordem epistemológica, constituída de três questões que geram intensos debates neste campo.

A primeira questão é sobre a possibilidade de abordar cientificamente os fenômenos sociais utilizando a mesma abordagem metodológica dos fenômenos naturais - físicos e biológicos, da qual resulta a questão da objetividade *versus* subjetividade na produção do conhecimento nas ciências sociais. A segunda questão, decorrente da primeira, refere-se à opção de utilizar metodologias quantitativas *versus* qualitativas na pesquisa em avaliação. A terceira, refere-se aos critérios de cientificidade ou validade do conhecimento produzido pelas ciências sociais, que diferem conforme a metodologia escolhida.

As diferentes respostas a estas questões definem as diferentes correntes de pensamento nas Ciências Sociais, na medida que agrupam as “convicções metodológicas maiores” que caracterizam os “paradigmas”, “quadros teóricos” ou “matrizes disciplinares” destas correntes ou escolas de pensamento (Bruyne et al.: 1982, 133-34).

Não iremos tratar das diferenças entre estas correntes de pensamento porque não estaria contido no âmbito deste trabalho. Além do que, são bastante conhecidas no meio acadêmico. Entretanto, caberia lembrar quais são, apontando como elas “respondem” às questões centrais do debate em Ciências Sociais, e como este debate reflete-se no campo da avaliação em saúde.

Segundo MINAYO, as três correntes de pensamento mais relevantes para o campo das Ciências Sociais aplicadas à saúde são: o Positivismo Sociológico, cujo Funcionalismo é uma variante; a Sociologia Compreensiva, que inclui a Fenomenologia, a Etnometodologia e o Interacionismo Simbólico; o Marxismo, que tem como princípios fundamentais o materialismo histórico e o materialismo dialético (Minayo: 1992, 27-88).

Para IANNI, as diversas “tendências, escolas, teorias, interpretações” da Sociologia¹⁷ se reduziriam a três polarizações fundamentais, com base em “três princípios explicativos”: causalção funcional, conexão de sentido e contradição, que caracterizam, respectivamente, as três correntes apontadas por Minayo (IANNI: 1989, 10).

¹⁷O autor destaca, entre outras, nove tendências: evolucionismo, organicismo, positivismo, formalismo, funcionalismo, estruturalismo, estrutural-funcionalismo, fenomenologia e historicismo. (Ianni: 1989, 10)

BRUYNE e colaboradores fazem um recorte diferente do campo das Ciências Sociais, mas que sobrepõe-se parcialmente aos anteriores, estabelecendo quatro “quadros de referência” na produção do conhecimento: positivista, compreensivo, funcionalista e estruturalista (Bruyne et al.: 1982, 134).

Seja qual for o recorte, estas correntes de pensamento tem origem na produção teórica de três autores fundamentais das Ciências Sociais: Durkheim (Positivismo), Weber (Sociologia Compreensiva) e Marx (Marxismo).

O positivismo é a corrente de pensamento que deu origem à ciência moderna. O longo processo de maturação histórica desta corrente, iniciado com o desenvolvimento das ciências físicas e, posteriormente, as biológicas, resultou em que grande parte do instrumental metodológico disponível no campo das Ciências Sociais esteja baseado, ainda hoje, na lógica ocidental positivista. Por esta razão, a noção de cientificidade dominante no campo da ciência em geral é fruto desta lógica.

Esta precessão histórica do positivismo fez com que as demais correntes de pensamento fossem construindo-se a partir das críticas ao mesmo, e por oposição a ele. Por isso, ao tratar das três questões da problemática, caracterizaremos em primeiro lugar a posição positivista, apontando em seguida as críticas feitas pelo autores das demais correntes de pensamento.

1.3.1.1. Objetividade *versus* Subjetividade

O positivismo responde que “sim” à primeira questão - é possível abordar cientificamente os fenômenos sociais utilizando a mesma lógica de abordagem dos fenômenos naturais.

Em suas diferentes variações teóricas, as correntes positivistas tomam o “modelo biológico como metáfora da sociedade” (Minayo: 1992, 16). O pressuposto básico deste modelo é a idéia de Durkheim de que o fato social é “coisa”. Em decorrência disto, a realidade social é considerada como “realidade natural”, portanto, pode e deve ser abordada pela mesma metodologia de abordagens das ciências naturais, oriundas especialmente da física e da biologia. Caracteriza-se, assim, uma concepção “física” da observação e do estudo dos fenômenos sociais (Thiollent: 1987, 17).

Vale a pena transcrever os princípios positivistas clássicos, que segundo HUGHES¹⁸, fundamentam a pesquisa quantitativa nas Ciências Sociais:

*“(a) o mundo social opera de acordo com leis causais; (b) o alicerce da ciência é a observação sensorial; (c) a realidade consiste em estruturas e instituições identificáveis enquanto dados brutos por um lado, crenças e valores por outro. Estas duas ordens são correlacionadas para fornecer generalizações e regularidades; (d) o que é real são os dados brutos considerados **dados objetivos**; valores e crenças são **realidades subjetivas** que só podem ser compreendidas através dos dados brutos [grifos nossos]” (Hughes apud Minayo: 1992, 30).*

Percebe-se claramente que é considerado “real” é o dado bruto, que pode ser percebido através da observação sensorial. Ou seja, há uma idéia de “materialidade” subjacente à noção do real. É desta noção de materialidade, de “coisa” que estaria disponível para ser observada, que surge a concepção de “dado objetivo”, isto é, o dado que pode ser medido ou contado. Em contrapartida, os elementos subjetivos - valores e crenças - não seriam reais, a não ser que pudessem ser transformados em “dados objetivos”, isto é, dados quantificáveis.

Esta concepção “física” da observação científica traz para dentro das Ciências Sociais, portanto, uma pretensão de “**objetividade**” na produção do conhecimento, que se realizaria através da quantificação.

Outro aspecto derivado desta concepção da observação científica é a postura metodológica conhecida como “empiricismo”, que propõe que o conhecimento “verdadeiro” pode ser baseado exclusivamente na experiência. Desta concepção deriva a idéia de que o cumprimento dos passos do método experimental é suficiente para garantir a cientificidade da investigação. Na radicalidade do pensamento positivista, o “dado objetivo” gerado na experimentação é tanto o ponto de partida como o eixo em torno do qual desenvolve-se todo o processo de pesquisa, cabendo às teorias um valor secundário, ou mesmo valor nenhum.

Segundo THIOLENT, “para o empirista puro, todas as teorias são vãs filosofias”, pois os dados revelariam por si mesmo a “verdade” sobre a realidade. O simples “arranjo” dos dados coletados, isto é, o processamento dos mesmos “sem a mediação de categorias ou de corpo teórico pré-existente” levaria à descoberta das “leis’ ou até ‘teorias’ relevantes do ponto de vista da sociologia.” Esta ênfase nos dados visaria dar suporte à noção de “objetividade” científica no positivismo. Porém, de fato estaria fundada na “ilusão da imaculada percepção” da realidade, gerando o “mito cognitivo” segundo o qual a **verdade sobre o real** poderia captada através dos dados, desde que fossem obtidos através da aplicação de técnicas rigorosas de coleta. Para o autor, este é o pressuposto equivocado da postura metodológica empiricista (Thiollent: 1987, 20).

A crítica de LADRIÈRE é elucidativa desta postura:

“A ciência não é simplesmente o prolongamento da visão espontânea do mundo, ou uma formulação um pouco sofisticada do que se oferece à percepção. Ela só consegue fazer com que capturemos aspectos inéditos da realidade na medida em que começa por substituir o campo perceptivo por um domínio de objetos que ela constrói por seus próprios meios. É através deste domínio, cujo funcionamento ela estuda em seguida, que ela vive o mundo real, isto é, este universo de coisas e significações que nos é revelado na percepção e no qual se inscreve nessa ação” (Bruyne et al.: 1982, 29, Prefácio de Jean Ladrière).

A sociologia compreensiva e o marxismo criticam ferozmente o empiricismo positivista, que pretende construir “verdades científicas” através da simples manipulação dos

¹⁸HUGHES, J. A *Filosofia da Pesquisa Social*. Rio de Janeiro, Zahar Ed., 1983, pp. 42-63)

dados, e cujo critério de validade científica é o cumprimento rigoroso do método. Para estas correntes, a construção do “domínio dos objetos da ciência” se faz a partir das teorias, que por sua vez são produzidas dentro dos paradigmas de cada corrente de pensamento.

A abordagem do social embasada nas metodologias de abordagem do biológico é chamada de “reducionismo”, pois implica em que se faça um “deslocamento do nível interpretativo que corresponde o fenômeno em estudo [nível social] para um nível inferior [nível biológico]”. Este deslocamento, criticável do ponto de vista analítico, seria “a principal fonte de graves erros científicos na interpretação dos fenômenos sociais” (Testa: 1992, 96; Bruyne et al.: 1982, 9-10, Prefácio de Jean Ladrière). Na medida em que “transforma a realidade social em realidade natural”, a concepção positivista “naturaliza” a vida social (Thiollent: 1987, 17).

Além disto, a concepção “física” da realidade social é intensamente criticada pelas demais correntes de pensamento, porque tem uma função ideológica de manutenção do **status quo** - a qual é dissimulada pela apologia à neutralidade da ciência -, função esta que consiste em induzir a pensar que os fenômenos sociais “são o que são”, e assim devem ser conhecidos, omitindo-se a idéia e a possibilidade de pensar na mudança ou transformação histórica da sociedade.

Entramos então no pilar do pensamento positivista - o princípio da neutralidade científica. Segundo este princípio, o campo da ciência pode e deve ser neutro em relação às influências ideológicas e sócio-políticas. O cumprimento rigoroso de procedimentos metodológicos bem definidos garantiria a neutralidade do pesquisador diante do objeto estudado - o fenômeno social. O autores vinculados às correntes compreensivas e marxistas, pelo contrário, apontam a vinculação que existe entre as linhas de pensamento e os aspectos ideológicos, por duas razões.

A primeira razão, porque as próprias correntes de pensamento, que fundamentam o desenvolvimento das ciências, tem caráter histórico e ideológico: “As correntes de pensamento têm sua história, veiculam uma visão de mundo e têm a ver com a realidade social complexa onde foram geradas e que elas tentam expressar.” Por isso, elas representam “não apenas diferentes possibilidades de análise, mas uma luta ideológica que, por sua vez, tem a ver com a luta política mais ampla na sociedade”. Ou seja, haveria um inevitável “imbricamento entre

conhecimento e interesse”, de forma que “toda a ciência é comprometida” (Minayo: 1992, 13; 19-21).

Para THIOLENT, no campo das Ciências Sociais “a neutralidade é falsa ou inexistente na medida que qualquer procedimento de investigação envolve pressupostos teóricos e práticos variáveis segundo os interesses sócio-políticas que estão em jogo no ato de conhecer” (Thiolent: 1987, 29). Por isso, a neutralidade não seria possível quando o objeto de pesquisa é de natureza social.

Neste sentido, para os pensadores de inspiração marxista, é fundamental que o pesquisador tome consciência das implicações políticas de cada corrente de pensamento, e que se posicione em relação a estas implicações, abandonando a pretensão de neutralidade.

A segunda razão pela qual entende-se que a neutralidade na produção do conhecimento não é possível, deriva da compreensão de que, enquanto nas ciências físicas e biológicas existe um certo distanciamento entre o sujeito e os objetos do estudo, no processo de pesquisa em Ciências Sociais há, necessariamente uma interação, um envolvimento entre pesquisador e pesquisado, ou seja, “uma identidade entre o sujeito e o objeto da investigação”, onde cada qual entra com sua inevitável carga ideológica (Minayo: 1992, 21).

Por tudo isto, voltando à pergunta inicial, as correntes de pensamento marxistas e compreensivas respondem que “não” é possível abordar os fenômenos sociais da mesma forma que os fenômenos naturais. Estas correntes estabelecem “nítidas diferenças” entre as ciências da natureza e as ciências sociais, considerando que estas diferenças são “essenciais e irreversíveis” (Ianni: 1989, 9).

O ponto fundamental da divergência é a concepção de que o fenômeno social não é “coisa”, no sentido de algo que está dado, pronto para ser observado em suas características “objetivas”. Para os marxistas e compreensivistas, é preciso considerar que a realidade social é histórica, ou seja, está em permanente processo de construção pelo homem, através dos conflitos ou lutas políticas. Logo, o objeto de estudo na esfera social - o fenômeno social -, é um objeto “complexo, contraditório, inacabado e em permanente transformação” (Minayo: 1992, 22). Ainda que se possa alegar que todo objeto científico tem estas características, seja ou não social, parece evidente que os fenômenos sociais respondem à determinações de múltiplas ordens de complexidade.

IANNI define desta forma o objeto de estudo das Ciências Sociais:

“O objeto da sociologia, bem como das outras ciências sociais, envolve o indivíduo, a coletividade, as relações de coexistência e seqüência, diversidades e antagonismos. Diz respeito a seres dotados de vontade, querer, devir, ideais, ilusões, consciência, inconsciente, racionalidade, irracionalidade. Os fatos e acontecimentos sociais são sempre materiais e espirituais, envolvendo relações, processos e estruturas de dominação, ou poder, e apropriação, ou distribuição. Implicam em indivíduos, famílias, grupos, classes, movimentos, instituições, padrões de comportamento, valores, fantasias. Esse é o mundo da liberdade e igualdade, trabalho e alienação, sofrimento e resignação, ideologia e utopia” (Ianni: 1989, 10).

Percebe-se que o autor acentua a característica subjetiva de todo o fenômeno social - vontade, ideais, ilusões, consciência de grupo ou de classe etc. Ou seja, o objeto social não é menos “real” por ser subjetivo, imaterial.

Sendo que o objeto das Ciências Sociais é complexo, contraditório, inacabado e imaterial, a objetividade na produção do conhecimento científico em Ciências Sociais será sempre relativa, será sempre apenas uma **tentativa de aproximação** do real (Minayo: 1992, 17; Thiollent: 1987, 29).

MINAYO coloca assim a questão da objetividade nas Ciências Sociais:

*“... não existe nenhuma evidência nos fenômenos sociais do campo da saúde: nada é dado, tudo é construído, inconcluso e superável; que a realidade aí encontrada como todo o social é infinitamente mais rica, mais dinâmica, mais complexa do que qualquer discurso científico sobre ela, e que a ciência que a aborda não a captura, ela apenas indica a direção e a organização intelectual segundo a qual se pode ter maior certeza da **aproximação do real**”*[grifo nosso] (Minayo: 1992, 249).

Todos os autores consultados concordam que a ação humana decorre de uma intencionalidade que se origina nas motivações, valores, crenças e objetivos vividos subjetivamente pelo sujeito da ação. A pergunta que se coloca é a seguinte: “pode-se tratar

esses componentes da ação à maneira das propriedades de um objeto físico, pode-se ‘objetivá-los’?’ (Bruyne et al.: 1982, 9-10, Prefácio de Jean Ladrière) Ou ainda, como operacionalizar os níveis mais profundos das relações sociais em “números e variáveis”? (Minayo: 1992, 16).

A resposta parece ter sido dada por LADRIÈRE. Segundo este autor, para estudar os fenômenos sociais, o pesquisador tem dois caminhos: o primeiro, seria colocar entre parênteses o “agente” da ação humana, ou seja, o homem ou sujeito da ação, e fazer aparecer “sistemas” que poderiam ser abordados por métodos análogos aos “que mostraram seu valor no estudo dos sistemas materiais”; o segundo caminho seria adotar a abordagem da compreensão, que assume o pressuposto de que **toda ação tem sentido** e de que **a realidade social é produto da ação do homem** (Bruyne et al.: 1982, 9-10, Prefácio de Jean Ladrière).

Neste segundo caminho, a questão da objetividade é colocada “em outro nível”. Abandona-se a pretensão de “objetividade” do positivismo - que parte do pressuposto que os fatos sociais são como as coisas da natureza, que estão dadas, prontas para serem conhecidas - e parte-se para uma “tentativa de objetivação do conhecimento”, que parte do pressuposto que a objetividade nas ciências sociais é sempre relativa (Minayo: 1992, 21-22, 35-36).

Ao invés de tentar garantir, inutilmente, através do cumprimento rigoroso de passo metodológicos, uma objetividade e neutralidade que serão sempre aparentes, propõe-se que o pesquisador reconheça e assuma que sua “visão de mundo” está implicada no processo de produção do conhecimento, da mesma forma que a visão dos demais atores sociais implicados neste processo. Por esta razão, é necessário que o pesquisador mantenha uma “crítica não só sobre as condições de compreensão do objeto como do próprio pesquisador” (Minayo: 1992, 21). Segundo LÉVY STRAUSS, “numa ciência onde o observador é da mesma natureza que o objeto, o observador é, ele mesmo, uma parte de sua observação” (Levy Strauss apud Minayo: 1992, 21).

BRUYNE e colaboradores falam que o pesquisador deve manter uma “vigilância epistemológica” sobre suas práticas, ou seja, uma reflexão interna e permanente sobre “seus procedimentos e seus resultados”. Esta vigilância deveria permear a produção do conhecimento em todos os seus níveis ou etapas, desde a formulação teórica, passando pelo controle metodológico mais geral e chegando até o controle na aplicação das técnicas de coleta de dados (Bruyne et al.: 1982, 41).

ALVES chama atenção que o “**objetivismo**” - isto é, a pretensão de objetividade do pensamento positivista - não deve ser confundido com a “**objetivação**”, que é a proposta de objetividade das demais correntes. Segundo ele, o objetivismo é uma resposta ao “problema da relação sujeito e objeto”, na qual o objeto é tomado como o elemento decisivo da relação. A “objetivação”, por sua vez, seria a “arrumação de uma rede conceitual que possa dar conta do real, explorá-lo, expressá-lo de forma fundamentada e adequada” (Alves: 1994, 66). Ou seja, o objetivismo está centrado no estudo das características do objeto, enquanto a objetivação está centrada na capacidade do sujeito de teorizar sobre o mesmo.

A tentativa dos autores em definir o conceito de “objetivação” resulta em definições muito amplas, que confundem-se com a própria noção de “metodologia” na produção do conhecimento em Ciências Sociais. Talvez por isto, MINAYO entende que a objetivação “define o próprio movimento investigativo que, embora não consiga reproduzir a realidade, está sempre em busca de uma maior aproximação” (Minayo: 1992, 239).

Para BRUYNE e colaboradores., o conceito de objetivação refere-se ao “conjunto dos métodos e das técnicas que elaboram o objeto de conhecimento ao qual se refere a investigação” (Bruyne et al.: 1982, 52). LADRIÈRE refere como “o caminho da análise sistemática” (Bruyne et al.: 1982, 11-12, Prefácio de Jean Ladrière), caminho este que MINAYO entende como sendo um processo “interminável e necessário de atingir a realidade” (Minayo: 1992, 35).

Seja qual for a definição do conceito, o princípio da objetivação implica “a crença na necessidade permanente de um diálogo crítico entre o investigador e seu objeto, sabendo que ambos compartilham a mesma condição histórica e os mesmos recursos teóricos” (Minayo: 1992, 239). Ouçamos a autora:

“É preciso aceitar que o sujeito das ciências sociais não é neutro ou então se elimina o sujeito do processo de conhecimento. Da mesma forma, o ‘objeto’ dentro dessas ciências é também sujeito e interage permanentemente com o investigador”
(Minayo: 1992, 35).

Na tentativa de melhorar a compreensão destes conceitos, recorreremos ao Dicionário Aurélio, onde verificamos que o objetivismo refere-se a “doutrina que afirma existirem normas

objetivas, de validade geral”, enquanto o termo “objetivação”, nas correntes dialéticas contemporâneas, refere-se “ao processo pelo qual a subjetividade ou consciência humana corporifica-se em produtos avaliáveis para ela e para os outros como elementos de um mundo comum”. Ou seja, enquanto o objetivismo busca encontrar as leis que regem os fenômenos sociais, através da quantificação de dados brutos, à semelhança das ciências naturais, a objetivação pretende alcançar o conhecimento da realidade social através da aceitação, explicitação e vigilância da **subjetividade** que permeia a própria construção dos objetos de conhecimento neste campo.

Embora a abordagem compreensiva adote o pressuposto de que a objetividade não é possível, isto não significa que proponha o abandono do rigor metodológico. Segundo MINAYO, a busca da objetivação implica no uso rigoroso de “instrumental teórico e técnico adequado”, com a finalidade de “reduzir a incursão excessiva dos juízos de valor na pesquisa”, mas também implica em “repudiar o discurso ingênuo ou malicioso da neutralidade” (Minayo: 1992, 22, 35-36).

Em resumo, o pensamento positivista considera que o cumprimento rigoroso de procedimentos metodológicos bem definidos, podem gerar dados “objetivos” sobre a realidade social, garantindo-se também a neutralidade deste conhecimento em relação às influências ideológicas e sócio-políticas. As correntes marxistas e compreensivas partem de pressupostos contrários. A partir do reconhecimento de que o fato social é de uma natureza diversa dos fenômenos físicos e biológicos, principalmente pela intencionalidade intrínseca da ação humana, propõe-se uma abordagem metodológica que, admitindo a subjetividade inerente à pesquisador e pesquisado, possa justamente captar o significado e a intencionalidade da ação social.

Como estas questões gerais das Ciências Sociais se colocam no campo específico da avaliação?

Como vimos anteriormente, o objeto privilegiado da avaliação são as intervenções sociais, ou seja, a ação humana visando a mudança social. Ao delimitar-se este objeto, localizado no campo subjetivo dos valores, fica caracterizado o contexto altamente conflitivo do processo avaliativo. Ou seja, se a investigação no campo social já é difícil, nos estudos

avaliativos é praticamente impossível evitar as “influências sócio-políticas” no processo de produção do conhecimento.

Além disto, a própria concepção da avaliação implica em um alto grau de subjetividade - formular um juízo de valor sobre uma intervenção -, ainda que o esforço avaliativo seja dirigido para buscar elementos da realidade objetiva, visando fundamentar este juízo, através da medição e comparação dos efeitos desta intervenção.

Enfim, tanto por seu objeto, quanto por seu contexto e sua finalidade, a pesquisa em avaliação põe em evidência, de forma crucial, o conflito entre a subjetividade e objetividade que permeia a discussão sobre a produção do conhecimento em Ciências Sociais.

A nosso ver, o aparente descompasso entre a "subjetividade" do conceito de avaliação e a "objetividade" de seus procedimentos reproduz, de forma sintética, o esforço positivista característico da ciência - objetivar o que é percebido sempre subjetivamente -, embora seja “continuamente desafiado pelo caráter dialético” da mesma (Nunes: 1995, 28).

A busca para corresponder à imagem hegemônica de cientificidade, oriunda do positivismo, tem levado o campo da avaliação ao privilegiamento do “objetivismo”. Ou seja, até agora, tem havido uma preferência pela primeira opção colocada por LADRIÈRE: colocar o agente da ação (o homem) entre parênteses, a fazer aparecer “sistemas”. A influência da Teoria de Sistemas no campo da avaliação implica, em parte, em abrir mão do sujeito desta ação, abrindo mão da discussão da subjetividade. Além disto, o caráter aplicado dos estudos em avaliação, aliado à uma certa deficiência teórica do campo, favorecem o empiricismo nos estudos avaliativos.

Porém, tem surgido uma tendência em voltar-se para os aspectos subjetivos envolvidos no processo avaliativo, como a proposta de negociar os resultados da avaliação com os atores envolvidos no processo, ou a proposta de avaliação emancipatória, na “quarta fase” da história da avaliação, que vimos anteriormente. Esta nova tendência reflete influências das correntes de pensamento compreensivas e marxistas, especialmente suas bases dialéticas.

Ao tratar das bases conceituais da dialética, MINAYO cita, entre outras, a capacidade de perceber que existe uma relação dialética entre o objetivo e o subjetivo, na medida em que se “considera parte da mesma totalidade o objeto e o sujeito.” (Minayo: 1992, 73). Ela acentua

a importância de evitar os exageros tanto num quanto noutro sentido, como vemos na citação seguinte:

“...combater tanto os exageros da supremacia da razão como dos subjetivismos, assumindo a relatividade dos fatores objetivo e subjetivo, material e espiritual e sua unidade dialética. Trata-se da relatividade intrínseca a todo o processo de conhecimento e não apenas da oposição entre subjetivo e objetivo. Tem a ver com a própria capacidade humana de apreensão da realidade...” (Minayo: 1992, 73)

Em princípio, poderíamos considerar que as diferentes correntes de pensamento podem aportar métodos e técnicas para captar diferentes aspectos da realidade: as compreensivas podem contribuir para captar os valores subjacentes às propostas de intervenção; as correntes positivistas para estabelecer relações de causa-efeito entre uma intervenção e os efeitos observados em uma determinada situação; as correntes marxistas para analisar os conflitos de interesses subjacentes às intervenções.

Veremos na seção seguinte como a pesquisa em avaliação tem recorrido à combinação de diferentes metodologias e técnicas de abordagem, como forma de apreender os aspectos objetivos e subjetivos da realidade social, superando a dicotomia quantitativo *versus* qualitativo.

1.3.1.2. Quantitativo *versus* Qualitativo

A segunda questão da problemática das Ciências Sociais refere-se ao debate sobre o uso das metodologias quantitativas *versus* qualitativas, que está vinculada à discussão sobre a subjetividade *versus* a objetividade da pesquisa neste campo.

A discussão entre as duas metodologias tem sua origem “nas diferentes formas de perceber a realidade social” (Minayo: 1992, 30). Como vimos na seção anterior, a visão “fisicista” da sociedade resulta no “objetivismo”, onde se busca transformar os aspectos subjetivos da vida social em “dados objetivos”, isto é, dados quantificáveis. Por outro lado, a visão histórica da sociedade leva à busca de dados que possam evidenciar o significado da ação humana, nem sempre traduzível por números e variáveis, através da tentativa de “objetivação” do conhecimento, que implica em assumir a subjetividade tanto do investigador quanto do investigado, ambos considerados “sujeito” na produção do conhecimento.

Ou seja, o debate objetivo *versus* subjetivo leva ao debate quantitativo *versus* qualitativo, e aponta para o problema fundamental que é o próprio “caráter específico do objeto de conhecimento” nas ciências sociais: o ser humano e a sociedade. Desta forma, segundo MINAYO, o dilema do estudioso é “contentar-se com a problematização do **produto humano objetivado** ou de ir em busca, também, do **significado da ação humana** que constrói a história [grifos nossos]” (Minayo: 1992, 36).

Veremos a seguir alguns aspectos envolvidos neste debate, que apresentamos em três seções: na primeira, discutiremos alguns aspectos históricos que determinaram a separação entre as metodologias quantitativas e qualitativas; na segunda seção, trataremos a tendência atual para integrar estas metodologias; e na terceira, caracterizaremos as principais diferenças entre estas duas abordagens.

1.3.1.2.1. Breve história de uma separação

Segundo MINAYO, os termos “metodologias quantitativas” e “metodologias qualitativas” refletem “uma luta teórica” entre as correntes de pensamento positivistas e compreensivistas em relação à questão da “apreensão dos significados” dos fenômenos sociais (Minayo: 1992, 13).

A influência do positivismo nas Ciências Sociais teria levado a enfatizar-se “a utilização dos termos de tipo matemático para a compreensão da realidade e a linguagem de variáveis para especificar atributos e qualidades do objeto de investigação” (Minayo: 1992; 30). Porém, neste campo de estudos, não bastaria demonstrar o "o quê" (aspecto objetivo do fenômeno), seria preciso também explicar "o porquê" (significado subjetivo do fenômeno), pois os fenômenos sociais são "vividos culturalmente". Os níveis mais profundos das relações sociais - valores, atitudes e crenças - não poderiam ser operacionalizados em “números e variáveis, critérios usualmente aceitos para emitir juízo de verdade no campo intelectual” (pp. 10, 16). Por isto, esta autora considera que o objeto das Ciências Sociais é “essencialmente qualitativo”, sendo até uma redundância falar em “pesquisa qualitativa” neste campo (pp. 21-22).

Evidentemente, este debate mais geral reflete-se no campo da Saúde Coletiva, onde as diferentes correntes fazem a apologia de uma ou outra metodologia, percebendo-se até mesmo um certo desprezo pela outra.

Porém, alguns autores lembram que, no passado, o uso integrado de abordagens qualitativas-quantitativas era “a norma” da pesquisa em Saúde Pública. YACH cita dois exemplos desta “abordagem integrada”: a investigação do surto de sarampo nas ilhas Faroe, realizada por Panum, em 1856, e a investigação da epidemia de tifo realizada por Rudolf Vischow, o qual utilizou “uma combinação de achados clínicos, patológicos, epidemiológicos e antropológicos, colocando-os em uma perspectiva sócio-econômica” (Yach: 1992, 603).

COSTA também chama atenção para o “espírito integrado” dos primórdios da Saúde Pública, no século XIX, tomando como exemplo o trabalho de John Snow sobre a investigação do surto de cólera em Londres. Para ela, Snow seria “um dos pais fundadores da reflexão totalizante em saúde”, pela maneira como integrava em suas investigações, todo o conhecimento disponível na época - processos físicos, químicos, biológicos, sociológicos e políticos (Costa: 1990, 168-69).

Se tudo começou assim, porque mudou?

A separação entre as metodologias qualitativas e quantitativas deve-se principalmente a um fenômeno que ocorreu no campo da ciência em geral - a fragmentação do conhecimento em várias disciplinas especializadas. No campo da Saúde Pública, alguns fatores específicos também contribuíram para esta divisão, conforme veremos a seguir.

Segundo YACH, a principal razão para que os investigadores tenham dividindo-se em “duas culturas separadas pelo uso de métodos quantitativos e qualitativos” foi o desenvolvimento de um grande número de disciplinas, depois da Segunda Guerra Mundial. Estas disciplinas passaram a erguer “paredes impenetráveis” entre si, mantendo-se isoladas em suas especificidades, criando culturas e jargões separados. Esta separação estendeu-se também ao uso de uma extensão limitada de métodos de pesquisa por parte de cada disciplina (Yach: 1992, 604).

A separação rígida dos campos disciplinares teria sido determinada menos pela necessidade de progresso da ciência e mais por “razões históricas e de conveniência administrativa (como a organização do ensino e do corpo docente)”(Karl Popper apud Costa: 1990, 194), que delimitam a história institucional das mesmas. Ou seja, esta separação visaria “primariamente a perpetuação das disciplinas” (Yach: 1992, 604).

Esta separação tem funcionado como uma “atitude intimidante e teoricamente restritiva” para o pesquisador, além do que cria uma confusão entre os “campos disciplinares” e os “objetos” de estudo para os quais estas disciplinas estão voltadas (Costa: 1990, 190). Entretanto, é preciso lembrar o que fala POPPER: “estudamos problemas, não matérias: problemas que podem ultrapassar as fronteiras de qualquer matéria ou disciplina” (Karl Popper apud Costa: 1990, 194-95).

No caso específico do campo da Saúde Pública, o fator primordial para o abandono da perspectiva integrada foi, segundo YACH, “o desenvolvimento das disciplinas de laboratório, e particularmente a bacteriologia e a virologia”, com as descobertas de Pasteur, Henle e Koch entre os anos de 1870 e 1920 (Yach: 1992, 603). Ou seja, o autor está referindo-se a um fator bastante conhecido no campo da Saúde Coletiva, que é a mudança do modelo da “Medicina Social” para o “modelo biomédico” na compreensão do fenômeno saúde-doença.

Além disto, ele identifica duas outras ordens de razão que contribuíram para a separação entre as metodologias quantitativas e qualitativas no campo da Saúde Pública (Yach: 1992, 604-605):

- a pressão das agências de financiamento para que os pesquisadores apresentassem “resultados rápidos”, o que favoreceu o uso dos “inquéritos quantitativos descritivos” e desfavoreceu a “pesquisa multidisciplinar qualitativa”, de duração mais prolongada. Esta pressão levou à ênfase na pesquisa epidemiológica, porém “despida” dos elementos qualitativos, tradicionalmente associados a ela;

- a dominância da profissão médica no campo da Saúde Pública, que cria uma tendência a priorizar o que as escolas médicas favorecem, ou seja, o clássico modelo biomédico.

Ele chama atenção que os maiores órgãos financiadores de pesquisas em Saúde Pública - OMS, UNICEF, USAID - mantêm programas que “claramente dicotomizam” as metodologias, colocando de um lado os estudos de caso controle, com medidas de exposição e efeito, e de outro lado, os inquéritos sobre conhecimentos, atitudes e práticas (Yach: 1992, 604).

Entretanto, observa-se indícios de mudança, de forma que a perspectiva para os anos 90 é um movimento em favor da integração entre as metodologias quantitativas e qualitativas, entendidas como sendo complementares (Yach: 1992, 603-604).

Também no campo específico da avaliação, o “aparente conflito” entre as dimensões quantitativas e qualitativas vinha sendo a “cruz dualística” do debate metodológico. Porém, refletindo a tendência geral das Ciências Sociais, observa-se uma inflexão no sentido de desenvolver desenhos de estudos onde estas duas dimensões sejam integradas (Yin: 1994, 287).

Veremos a seguir como vem manifestando-se a tendência para a integração das metodologias quantitativas e qualitativas.

1.3.1.2.2. A “indisciplinaridade” da avaliação

Como vimos na seção anterior, a principal causa da separação entre as metodologias quantitativas e qualitativas foi o surgimento de inúmeras disciplinas especializadas, que passaram a desenvolver-se de forma isolada, levando ao uso exclusivo de certas metodologias de pesquisa por parte de determinadas disciplinas. Logo, a integração entre estas metodologias deve necessariamente passar pela questão da interdisciplinaridade¹⁹.

A busca de uma abordagem integrada, que supere a “epistemologia da dissociação do saber”, manifesta-se tanto no campo mais amplo das Ciências Sociais em geral (Nunes: 1995, 96), como no campo específico da avaliação (Yin: 1994, 287).

A proposta da abordagem interdisciplinar nas Ciências Sociais está relacionada à própria compreensão da natureza subjetiva e objetiva dos fenômenos sociais, bem como à amplitude e complexidade dos objetos de estudo neste campo, o que demanda a utilização de diferentes métodos de pesquisa e de diferentes fontes de dados, para que se possa apreender suas múltiplas faces .

Segundo LADRIÈRE, o campo das ciências sociais é “excessivamente vasto para ser englobado ou reduzido a uma única disciplina; assim, de saída, esse campo é pluridisciplinar” (Bruyne et al.: 1982, 19, Prefácio de Jean Ladrière). Além de amplo, vimos anteriormente que o objeto das ciências sociais é especialmente complexo, contraditório, e está em permanente transformação, pois é histórico (Minayo: 1992, 22).

MINAYO considera que há uma relação de interdependência e inseparabilidade entre “os aspectos quantificáveis e a vivência significativa da realidade objetiva no cotidiano” (Minayo: 1992, 13). Ela cita Parga NINA²⁰, para quem “a utilização de métodos e técnicas não-quantitativas em uma pesquisa não é questão de escolha de alternativa ou de preferência pessoal: são procedimentos simplesmente necessários” (Parga Nina apud Minayo:1992, 133).

¹⁹Não desconhecemos as diferenças entre os termos “multidisciplinaridade”, “pluridisciplinaridade”, “interdisciplinaridade” e “transdisciplinaridade”, bem caracterizadas por Hilton Japiassu, conforme vemos em NUNES (Nunes: 1995, 19-22). Entretanto, no contexto deste trabalho, entendemos que podemos utilizá-los indistintamente, pois o fundamental é que todos refletem a “tentativa de sair da compartimentalização” do conhecimento, sendo abarcados pelo “campo genericamente denominado das relações interdisciplinares” (Nunes: 1995, 6). Para maiores detalhes, ver: JAPIASSU, Hilton. *Interdisciplinaridade e patologia do saber*. Rio de Janeiro, Imago, 1976.

²⁰Segundo Minayo, coordenador nacional, em 1974, do Estudo Nacional de Despesas Familiares (ENDEF), trabalho quantitativo que coletou informações de 53.000 famílias em todo o Brasil.(Minayo: 1992,132-33, Nota

Segundo QUEIROZ, as metodologias quantitativas e qualitativas não deveriam ser vistas como dois conjuntos de técnicas “opostos ou mutuamente exclusivos”, mas complementares. Os procedimentos quantitativos e qualitativos poderiam ser empregados tanto em tipos diferentes de pesquisa como em determinados momentos da mesma pesquisa (Queiroz: 1987, 278). As técnicas qualitativas, em geral, seriam muito úteis para um “primeiro levantamento de questões e de problemas” a respeito dos quais não existem conhecimentos. As técnicas quantitativas, por sua vez, serviriam principalmente para se conhecer “a intensidade de um fenômeno, o quanto se espalha por um grupo ou camada, como atinge grupos e camadas diferentes” (p. 283).

No caso específico da avaliação, a interdisciplinaridade é uma necessidade tanto pela natureza do seu objeto primordial - as intervenções sociais -, que colocam o problema da ação em nível coletivo, quanto pela natureza do processo avaliativo, cuja circularidade tem início no campo subjetivo e conflituoso dos valores sociais, onde são definidos os objetivos da ação, passa pelo esforço de buscar evidências objetivas sobre os efeitos da mesma, e volta à subjetividade para estabelecer um juízo de valor sobre a ação.

Parafraseando a colocação anterior de Minayo, no campo da avaliação não bastaria demonstrar o efeito concreto da ação - aspecto objetivo do fenômeno -, mas também explicar o significado desta ação e de seus efeitos - aspecto subjetivo do fenômeno.

A necessidade da troca interdisciplinar foi percebida desde os primórdios do desenvolvimento do campo da avaliação. Há trinta anos atrás, COTRELL apontava “a potencialmente produtiva influência recíproca entre uma pesquisa avaliativa bem desenhada e executada e o desenvolvimento teórico e metodológico das disciplinas científicas elas próprias” (Suchman: 1967, viii, Prefácio de Leonard S. Cottrell Jr.).

Recentemente, YIN analisa a gradual integração que vem ocorrendo entre as metodologias qualitativas e quantitativas, através do desenvolvimento de desenhos de estudo baseados nestas duas dimensões, especialmente os “estudos de caso” em avaliação. O autor aponta que, inicialmente mais separadas do que sobrepostas, as duas metodologias vem cada vez mais se sobrepondo, conforme vemos no Anexo III (Yin: 1994, 287-88).

de rodapé nº7). Apesar disto, este pesquisador “reconhece que pesquisas qualitativas podem ser de qualidade muito superior às que fazem análises quantitativas” (p. 29).

Voltando ao objeto da avaliação - as intervenções sociais - vemos que o **problema da ação** sobre a realidade social concreta parece ser o elemento determinante das relações interdisciplinares em geral.

Em trabalho recente, no qual analisa a interdisciplinaridade na Saúde Coletiva, NUNES aponta que a reflexão e a prática interdisciplinar são exigências das “situações concretas”, especialmente nas áreas da educação, saúde, tecnologia, planejamento (Nunes: 1995, 97-98; 106). Ora, estas são justamente as áreas privilegiadas dos estudos em avaliação.

Para este autor, a questão da interdisciplinaridade coloca-se tanto no plano do “conhecimento ou da teorização”, quanto no plano da “intervenção efetiva no campo da realidade social”, pois é necessário “dar respostas a problemas de ordem prática” (pp. 98, 101). Algumas citações demonstram o vínculo da interdisciplinaridade com o problema da ação concreta sobre a realidade social: “interdisciplina como prática do saber” (Benedito Nunes), “mobilização cada vez mais extensa dos saberes convergentes em vista da ação” (Hilton Japiassu) (Nunes: 1995, 101, 97). Voltando a Popper, estudamos “problemas” concretos, e não “matérias ou disciplinas” (Karl Popper apud Costa: 1990, 194-95).

É justamente neste ponto que a temática da avaliação apresenta um potencial específico para catalisar a questão das relações interdisciplinares. Com certeza, a complexidade dos problemas sociais e das intervenções planejadas para modificá-los é uma das razões pela qual o conceito de avaliação está enraizado na lógica da teoria de sistemas, como vimos na parte referente ao conceito de avaliação.

WEILENMANN colocava que a perspectiva cibernética do conceito de sistema permite abordar “a condição de circularidade e a característica de **rede** do fenômeno social” (Weilenmann: 1980, 13).

Segundo MORAES, a Teoria Geral de Sistemas foi desenvolvida pelo biólogo e filósofo alemão Ludwig von Bertalanffy, em trabalhos publicados entre 1950 e 1968 (Moraes: 1994, 23). Nas colocações da autora, percebemos que esta teoria originou-se da tentativa de superar a fragmentação do conhecimento em disciplinas estanques, sugerindo que seu idealizador estava orientado por uma perspectiva interdisciplinar, como vemos a seguir :

“Bertalanffy procurava refletir, a sua maneira, uma preocupação de parte da comunidade intelectual internacional da época acerca da necessidade de integrar

e aproximar as diversas correntes do pensamento científico de então, para as quais o estilhaçamento do conhecimento humano em dezenas de especialidades autolimitadas em alcance, totalmente separadas e estanques, sem uma linguagem ou conceituação comum, não era aceitável” (Moraes: 1994, 23).

Provavelmente, o esforço teórico de Bertalanffy buscava contrapor-se à fragmentação disciplinar que passou a manifestar-se mais fortemente após a Segunda Guerra Mundial, como vimos anteriormente. Por outro lado, o desenvolvimento de sua teoria coincidiu com o período de grande desenvolvimento do campo da avaliação, na década de 60, nos EUA. O impulso para o desenvolvimento da avaliação ocorreu devido à necessidade de avaliar as intervenções ou programas sociais, entre as quais os de Saúde Pública, como veremos melhor nos aspectos históricos.

Esta coincidência histórica entre o desenvolvimento da avaliação de programas sociais e o surgimento da Teoria de Sistemas, permite levantar a hipótese de que esta teoria ofereceu ao campo da avaliação um enfoque integrado, que possibilitou a abordagem interdisciplinar necessária para dar conta da amplitude e complexidade das intervenções públicas.

As colocações de KNOX e colaboradores favorecem esta hipótese. Segundo estes autores, o uso do conceito de sistemas no setor saúde provocou uma erosão na tradicional segregação entre as áreas da Epidemiologia e da Administração/Planejamento, que afastava os epidemiologistas do processo decisório, enfraquecendo a distinção, antes clara, entre os papéis profissionais de ambos os campos (Knox et al.: 1986, 2).

Entretanto, o crescimento da Teoria de Sistemas no meio científico não pode ser atribuído exclusivamente à sua proposta de integração disciplinar. MORAES aponta três forças que contribuíram para que esta teoria viesse a tomar corpo e desenvolver-se, que acreditamos estar relacionadas também com o desenvolvimento do campo da avaliação:

“...de um lado, o crescimento das organizações modernas e a necessidade de seu controle, de outro, a evolução da indústria eletrônica, que possibilitou a produção dos computadores e, finalmente, o pensamento científico e filosófico, reunido para melhor compreender o mundo em transformação acelerada...” (Moraes: 1994, 23)

Ela ressalta que o crescimento da teoria foi devido, provavelmente, à “associação do conceito de Sistema com o da Informática” (Moraes: 1994, 23). Confirmando esta idéia, KNOX e colaboradores apontam que crescimento da Teoria de Sistemas nas áreas da Epidemiologia e do Planejamento em saúde, pode ser atribuído, em parte, ao desenvolvimento tecnológico do manejo da informação (informática) e à necessidade de sistemas de informação em larga escala (Knox et al.: 1986, 2).

Podemos inferir que a necessidade de controlar a ação das “organizações modernas”, grandes e complexas, viabilizada pelo desenvolvimento de “sistemas de informação em larga escala”, através da informática, são forças que também contribuíram para o desenvolvimento da avaliação dos programas sociais.

Além destas razões de ordem geral, que colocam-se no campo das Ciências Sociais, a avaliação em Saúde Coletiva exige uma abordagem interdisciplinar devido à características específicas deste campo.

A primeira é a própria natureza do fenômeno saúde-doença. Segundo NUNES, etimologicamente, a palavra saúde, tanto em sua origem latina como grega, comporta, entre outros, o significado de “inteiro”, de integridade. Em conseqüência, como regra geral, os problemas de saúde “não são disciplinares” (Nunes: 1995, 107; 98). Ouçamos o autor:

“Vista sob este aspecto, a saúde como integridade não permite fragmentação em saúde física, mental e social, portanto, parte-se de uma visão holística que em tese supõe entendê-la na interface de uma grande diversidade de disciplinas” (Nunes: 1995, 107).

Além disto, esta complexidade aumenta quando a abordagem do fenômeno saúde-doença ocorre em nível coletivo. Esta segunda característica - a abordagem da saúde em nível coletivo - confere um “caráter pluridimensional” ao campo, cujo objeto “em sua dimensão real exige uma visão multifacética”, colocando a interdisciplinaridade como um “horizonte necessário” (Nunes: 1995, 107; 110-11).

MINAYO entende da mesma forma. Para ela, o campo da Saúde Coletiva tem uma “abrangência multidisciplinar e estratégica”: multidisciplinar porque refere-se a uma

“realidade complexa que demanda conhecimentos distintos e integrados”, e estratégica porque “coloca de forma imediata o problema da intervenção” (Minayo: 1992, 13).

Estes conhecimentos seriam “saberes teóricos e práticos” de três ordens: 1 - saúde e doença; 2 - institucionalização, organização, administração e avaliação dos serviços, 3 - clientela dos sistemas de Saúde (p.13). Evidentemente, a autora está referindo-se às três áreas básicas de conhecimento que, tradicionalmente, informam a Saúde Pública - a Epidemiologia, o Planejamento e as Ciências Sociais, respectivamente, que revelam a natureza interdisciplinar deste campo.

Por outro lado, ao mencionar as intervenções como um problema imediato da Saúde Pública, esta autora recoloca o **problema da ação** sobre uma realidade social concreta, que vimos anteriormente.

O intercâmbio que rege as relações interdisciplinares na conformação do campo da avaliação em Saúde Pública está bem caracterizado por diversos autores (Battista et al.: 1989; Novaes: 1995; Hartz&Camacho: 1996).

Na seção referente aos objetos da avaliação em saúde, vimos como as tradicionais disciplinas da Saúde Pública - Epidemiologia, Planejamento e Ciências Sociais - tem abordado estes objetos. Entretanto, em debate recente sobre a questão da qualidade na avaliação dos serviços de saúde, NOVAES²¹ coloca que a discussão da qualidade está fazendo com que estejam sendo revisados os “recortes” ou “propriedades” destas áreas, especialmente entre a Epidemiologia e o Planejamento. Nesta revisão, as duas áreas estariam cedendo informações para a discussão, as quais seriam “re-costuradas” de um novo jeito pela área da Avaliação.

Em outro momento, esta autora lembra que a articulação entre a Epidemiologia e a Avaliação levou, inclusive, à criação de uma nova área - “Pesquisas em Serviços de Saúde” (Novaes: 1995).

BATTISTA e cols. demonstram como esta nova área de pesquisa foi ampliando a abrangência de seus estudos, de forma que acaba incluindo todos os níveis das intervenções

²¹Conforme colocações da Profª. H. Maria Novaes, na abertura da Mesa Redonda "A Epidemiologia na Avaliação dos Serviços de Saúde: A discussão da Qualidade", durante o I Congresso Paulista de Epidemiologia, promovido pela Associação Paulista de Saúde Pública, e realizado de 9 a 11 de março de 1995, na Faculdade de Medicina da Universidade de São Paulo (USP), na cidade de São Paulo.

em saúde - técnicas, práticas, organizações, programas e políticas -, conforme veremos a seguir (Battista et al.: 1989, 1156-58).

Inicialmente, a área de “pesquisas em serviços de saúde” (health services research) surge da integração das pesquisas nas áreas de programas e de políticas de saúde, pois a complexidade da avaliação das intervenções de saúde em nível coletivo vinha exigido a contribuição de diferentes disciplinas. Depois, engloba também alguns aspectos da pesquisa em epidemiologia clínica, voltada para a avaliação das práticas de saúde. Por fim, amplia-se até o nível das técnicas, cuja avaliação, juntamente com a das práticas, também pode ser feita nos níveis comunitário e populacional, onde denomina-se “avaliação de tecnologias” (technology assessment). Neste ponto, os autores afirmam que, “de fato, todos os estudos sobre as intervenções em saúde podem ser considerados estudos do cuidado de saúde”, abarcados sob o nome genérico de “pesquisa sobre o cuidado à saúde” (health care research) (Battista et al.: 1989, 1156-58). Para visualizar a forma como foi ocorrendo esta integração, ver Anexo IV.

Neste mesmo sentido, HARTZ lembra que o “âmbito populacional” ao qual destinam-se os programas de saúde coloca-os “na perspectiva das intervenções coletivas cuja forma de avaliação se fundamenta na abordagem da Epidemiologia e das Ciências Sociais” (Hartz: 1995, 49), até porque a própria modelagem atual dos programas é feita com a “incorporação da contribuição das diferentes disciplinas” (Hartz&Camacho: 1996).

HARTZ&CAMACHO colocam que a pesquisa em avaliação de serviços de saúde não poderia ser feita baseada exclusivamente em uma única disciplina ou ciência. Embora a Epidemiologia “forneça satisfatoriamente a base quantitativa para os estudos de Saúde Coletiva”, os autores entendem que ela é “condição necessária mas não suficiente”. Seria preciso o uso combinado desta disciplina com técnicas estatísticas, elementos da teoria organizacional e das ciências comportamentais (Hartz&Camacho: 1996).

Por esta razão, os programas de capacitação na área de avaliação em saúde implicam na necessidade de “familiarizar-se com as metodologias de pesquisa quantitativas e qualitativas”, pois, sem esta integração, as equipes de avaliação jamais alcançariam “a INDISCIPLINARIDADE, que deve caracterizar um projeto de avaliação” (Hartz&Camacho: 1996).

O termo “indisciplinaridade” teria sido proposto por Edgar Morin, para superar a noção de “transdisciplinaridade”, na qual cada disciplina ainda continua isolada “com a sua própria cegueira e a sua própria arrogância”²². No campo da avaliação seria necessário despreocupar-se com os limites das áreas de origem dos conhecimentos, por isto os autores propõe a utilização deste conceito (Hartz: 1995, 49; Hartz&Camacho: 1996).

Por fim, caberia tratar da forma através da qual o campo da avaliação realiza o intercâmbio interdisciplinar.

Ao propor a “indisciplinaridade”, HARTZ&CAMACHO mencionam o enfoque da **triangulação** para integrar diversas disciplinas na avaliação dos serviços de saúde, onde são empregadas “múltiplas teorias, métodos e procedimentos de coleta e análise de dados para o estudo adequado do problema” (Hartz&Camacho: 1996).

Este termo vem sendo empregado no campo das Ciências Sociais, e segundo JICK, é uma metáfora das áreas de navegação e estratégia militar, as quais utilizam “múltiplos pontos de referência para localizar um objeto em sua exata posição”. Para ele, a definição mais amplamente utilizada seria a de Norman K. Denzin, que entende triangulação como “a combinação de metodologias no estudo do mesmo fenômeno”. A estratégia de utilizar “múltiplos métodos” seria também denominada de “metodologia convergente, multi-método/multi-traço, validação convergente” (Jick in Maanen: 1985; 135-36).

Segundo QUEIROZ, a necessidade de utilização de dados “colhidos de fontes as mais variadas” é uma especificidade das ciências sociais, para poder “abarcas de forma ampla a realidade que estuda ” (Queiroz: 1987, 278).

Para YACH, o que hoje está sendo rotulado e conhecido como triangulação nada mais é do que a abordagem integrada dos epidemiologistas sociais do século XIX (Yach: 1992, 603).

No campo específico da avaliação, YIN coloca que o uso de “múltiplas fontes de evidência, de uma maneira convergente” é uma das técnicas, entre as sete identificadas por ele, que tendem a ser mais valorizadas nos próximos anos (Yin: 1994, 185).

Em resumo, pretendemos ter demonstrado que a constituição do campo da avaliação vem ocorrendo fundamentalmente através das relações interdisciplinares, superando a

²²Lembramos que na classificação de Hilton Japiassu, conforme vemos em NUNES, a transdisciplinaridade

tradicional dicotomia entre as metodologias quantitativas e qualitativas. O processo avaliativo implica na necessidade de abordar tanto os aspectos subjetivos (qualitativos) quanto os aspectos objetivos (quantitativos) das intervenções sociais. Além disto, a amplitude e complexidade destas intervenções requer o concurso de várias disciplinas, para poder abordá-las em sua totalidade. No campo específico da avaliação em saúde, além destes fatores gerais, a interdisciplinaridade torna-se necessária devido à natureza integrada do fenômeno saúde-doença. Neste campo, as relações interdisciplinares tem propiciado, inclusive, o surgimento de novas áreas de estudo, como a “pesquisa em serviços de saúde”.

1.3.1.2.3. Qual é mesmo a diferença?

Como vimos inicialmente, a divergência fundamental entre as metodologias quantitativas e qualitativas é em relação à “apreensão dos significados” dos fenômenos sociais, ou seja, a apreensão de seus aspectos subjetivos. Enquanto as metodologias quantitativas buscam “objetivar” estes aspectos em números e variáveis, as metodologias qualitativas partem do pressuposto que os níveis mais profundos da realidade social - valores, atitudes e crenças - não podem ser operacionalizados desta forma.

Nesta seção, tentaremos caracterizar os aspectos mais gerais que diferenciam as metodologias quantitativas das qualitativas - pressupostos, princípios, finalidade dos estudos, conceito fundamental, modelos de estudo, técnicas de coleta e análise de dados -, bem como as vantagens e desvantagens de cada uma. Ao final, discutiremos alguns aspectos envolvidos nas abordagens de pesquisa que buscam integrar estas duas metodologias.

O princípio fundamental das **metodologias quantitativas** é o “objetivismo”, oriundo da visão “física” da realidade social, onde o que é considerado “real” é aquilo que pode ser apreendido como “dado bruto”, isto é, que pode ser contado ou medido. A idéia implícita nesta

corresponde ao grau mais integrado das relações interdisciplinares (Nunes: 1995, 19-22).

noção de “real” é a materialidade do fenômeno, que permite observá-lo diretamente, de onde deriva a noção de objetividade do positivismo e a ênfase na quantificação.

Por esta razão, nos estudos que utilizam a metodologia quantitativa, o conceito central é a “variável”. O conceito aplica-se aos “agentes, processos, objetivos etc. que são numéricos ou quantitativos”, isto é, que podem ser contados (enumerados) ou medidos (em uma escala de valores) (Ferrari: 1982, 145). Coerentemente com a noção de objetividade positivista, entende-se que “os dados quantitativos provindos de mensuração ou contagem fornecem informações mais concreta e objetiva do que as avaliações que incluam apenas elementos subjetivos” (Guedes&Guedes: 1988, 21).

De uma forma geral, a finalidade dos estudos quantitativos é a identificação de regularidades nos fenômenos observados, através das quais pode-se estabelecer leis gerais sobre os mesmos. Isto é feito pela busca de associação entre variáveis, identificada por testes estatísticos, através do teste de hipótese, que geralmente busca estabelecer uma relação de causa-e-efeito entre duas variáveis - uma considerada “independente” (a causa), a outra considerada “dependente” da primeira (o efeito).

Na perspectiva metodológica quantitativa, o modelo ideal de estudo é o experimental, geralmente referido como “desenho do estudo”.

A rigor, no **modelo experimental**, ou “experimento verdadeiro”, o pesquisador deveria ter controle da variável independente (intervenção), podendo decidir sobre o tipo de intervenção a ser feita (o “quê”), a forma de sua aplicação sobre a população em estudo (o “como”), o momento desta aplicação (o “quando”) e bem como poder fazer uma distribuição aleatória de quem (o “quem”) vai ser submetido à intervenção - o grupo experimental -, ou não vai ser - o grupo controle. Este tipo clássico de experimentação é chamado de “experimentação pura” ou “experimento randomizado”²³. Em função do controle que dispõe sobre as condições do estudo, o pesquisador pode realizar a medição das variáveis antes (pré-teste) e após a intervenção (pós-teste), permitindo avaliar o efeito da variável independente sobre a dependente, testando desta forma a hipótese de que há uma associação entre as duas (Contandriopoulos et al.: 1994, 36; Silva & Formigli: 1994, 86-87).

²³Do inglês “at random”, que significa “aleatoriamente”.

Ou seja, no modelo experimental ideal, é fundamental a existência duas coisas: a medição das variáveis “antes e depois” de uma intervenção provocada e a formação de dois grupos - experimental e controle - nos quais serão efetuadas as medidas (Suchman: 1967, 21; Weiss: 1972, 61). Para visualizar melhor este modelo, ver Anexo V.

Evidentemente, as situações da vida real (“real-life setting”) raras vezes oferecem condições para a experimentação (Yin: 1993, 78). Embora o modelo experimental represente o “tipo ideal” de avaliação, na prática é pouco utilizado por duas razões: primeiro, porque requer a distribuição aleatória das pessoas entre o grupo que se beneficiará do serviço, programa, enfim, do projeto que está sendo avaliado, o que quase nunca é possível; segundo, porque requer a medição de variáveis nos dois grupos “antes” e “depois” da implantação do projeto, o que também raramente é possível, pois na maioria das vezes o “avaliador entra em cena quando o projeto está em processo de execução” (Vergara: 1980, 97).

No caso específico da Saúde Coletiva, somente em alguns poucos casos admite-se a constituição de “grupo controle” para avaliar intervenções programáticas, visto que, por razões éticas, dificilmente pode haver controle da alocação da variável independente (intervenção) sobre as populações (Silva & Formigli: 1994, 86).

WEILENMANN chama atenção que o pressuposto deste modelo é que os problemas sociais são “bem estruturados”, sendo possível buscar a “verdade ou a lei” sobre os mesmos através do estabelecimento de uma “simples relação de causa e efeito”. Entretanto, os fenômenos sociais apresentam-se como uma “rede”, onde os problemas são inter-relacionados e as relações entre variáveis são confusas (Weilenmann: 1980, 12-13).

Em função das dificuldades de aplicar o modelo experimental no estudo das intervenções sociais, a alternativa que a literatura clássica apresenta é a utilização de modelos “quase-experimentais”, desenvolvidos originalmente por Campbell & Stanley²⁴ (Yin: 1993, 78; Silva & Formigli: 1994, 86). Nestes modelos, o pesquisador tem controle sobre “o que”, “o quando” e o “como”, mas não controla “o quem”, isto é, não pode distribuir aleatoriamente os indivíduos entre os grupos controle e experimental (Contandriopoulos et al: 1994, 38).

²⁴CAMPBELL, D.J. & STANLEY, J.C. *Experimental and Quasi-Experimental Designs for Research*. New York, Rand Mc Nally College Pub. Co., 1966. Edição em português. *Delineamentos Experimentais e Quasi-Experimentais de Pesquisa*. (Trad. Didio, Renato A.T.). São Paulo, Editora Pedagógica Universitária Ltda. - EPU/EDUSP, 1979.

Os **modelos quase-experimentais** tem a vantagem de serem práticos e factíveis, além do que garantem uma “margem aceitável de rigor científico, sempre e quando o avaliador tenha claro que elementos não poderão ser controlados e se previna de possíveis interpretações errôneas dos resultados” (Vergara: 1980, 101-102; 106).

Na clássica classificação de Campbell&Stanley, existem três sub-tipos básicos deste modelo: séries temporais; séries temporais múltiplas e grupo controle não-equivalente (Weiss: 1972; 68-73; Vergara: 1980, 101-106 Silva & Formigli: 1994, 85).

O modelo de séries temporais consiste em aplicar uma série de medições, a intervalos periódicos, antes que se comece o projeto, durante sua execução e depois que tenha terminado. A diferença entre estas medidas permite avaliar se houve uma “mudança decisiva” em função do projeto ou não (Weiss: 1972; 68; Vergara: 1980, 102).

O modelo de séries temporais múltiplas é uma variante do primeiro. Consiste em comparar as medidas realizadas antes e depois da implantação do projeto, com as medidas efetuadas no mesmo período em uma população que não sofreu a intervenção (Weiss: 1972; 68-69; Vergara: 1980, 104).

O modelo de grupo controle não-equivalente consiste em que, tanto o grupo experimental quanto o controle, não são formados aleatoriamente, mas por “seleção ou conveniência”, buscando-se garantir, entretanto, que a composição do grupo controle seja a mais parecida possível com a do grupo experimental. A diferença das medidas das variáveis entre o grupo experimental e o controle permite avaliar se houve ou não mudança em função do projeto (Weiss: 1972, 69; Vergara: 1980, 105).

Entretanto, na maioria dos caso não será possível aplicar nem os modelos quase-experimentais. O avaliador pode então recorrer aos **modelos não experimentais**, que embora apresentem menores possibilidades de concluir sobre os efeitos do programa, são melhores do que “estudo nenhum”, oferecendo uma base confiável para a retroalimentação do mesmo. Também neste caso são apresentados três sub-tipos, que dispensam explicação: avaliação antes-depois de um único programa; avaliação unicamente depois (também denominada *ex post facto*); avaliação depois com grupo de comparação (Weiss: 1972, 73-4; Vergara: 1980, 106).

Nas metodologias quantitativas, em qualquer um dos modelos apresentado, a **coleta dos dados** é feita através da medição ou contagem direta do fenômeno estudado, quando possível, ou através da coleta de informações dos indivíduos, aplicando questionários, geralmente com perguntas fechadas, isto é, que apresentam um elenco de respostas previamente definidas.

O modo de análise dos dados é através da estatística, que pode ser descritiva, isto é, apenas verificar as frequências relativas dos fenômenos (percentagens), ou analítica, isto é, verificar a associação ou correlação entre variáveis, através do teste de hipóteses.

MINAYO esclarece que a restrição ao “quantitativismo sociológico” não está relacionada com a técnica, pois “são poderosos e reconhecidos instrumentos, supondo-se a necessidade de dados aglomerados ou indicadores sobre coletividades”, caso da Saúde Coletiva. A crítica dirige-se à tentativa de “restringir a realidade social ao que pode ser observado e quantificado apenas” (Minayo: 1992, 30). Com esta restrição, aponta QUEIROZ, os valores e emoções permanecem “escondidos nos próprios dados estatísticos” (Queiroz: 1987, 273).

Além disto, é preciso considerar o risco de que um estudo competente do ponto de vista estatístico pode estar desprezando “aspectos essenciais da realidade”, mesmo quando alcança “uma manipulação sofisticada dos instrumentos de análise”. Ou seja, pode ocorrer que o estudo esteja fornecendo uma “resposta exata para perguntas erradas ou imprecisas” (Minayo: 1992, 30).

Até aqui, tentamos identificar os aspectos fundamentais das metodologias quantitativas. A seguir, discutiremos as metodologias qualitativas.

MINAYO define as **metodologias qualitativas** como "aquelas capazes de incorporar a questão do SIGNIFICADO e da INTENCIONALIDADE como inerentes aos atos, às relações, e às estruturas sociais". Para ela, mesmo as estruturas sociais devem ser entendidas como “construções humanas significativas” (Minayo:1992, 10).

O princípio fundamental das metodologias qualitativas é a “objetivação”, oriundo da visão histórica da sociedade, que implica em algumas idéias básicas: primeiro, que a objetividade é sempre relativa; segundo, que toda ação humana tem significado e a realidade social é uma construção do homem, logo, não basta medir ou contar o “dado objetivo”, é

preciso compreender o seu significado; terceiro, que tanto o investigador quanto o investigado são “sujeitos” na produção do conhecimento.

Por esta razão, ao invés da variável, o conceito fundamental na perspectiva qualitativa é a “palavra”, através da qual se pretende apreender o significado dos fenômenos sociais.

Segundo THIOLENT, as categorias e a sintaxe da linguagem “sempre envolvem pressupostos comprometidos com alguma representação de ordem cultural, ideológica ou teórica” (Thiolent: 1987, 20). SCHRAIBER entende que “o que cada pessoa relata, e o modo como relata” são construções teóricas determinadas na vida em sociedade (Schraiber: 1995, 65). Em suma, é através da palavra que cada um expressa a sociedade em que vive. Por esta razão, considera-se que a palavra é “a arena onde se confrontam os valores sociais contraditórios” (Bakhtin apud Minayo in Guareschi & Jovchelovitch: 1994, 103).

Obviamente, o que se pretende alcançar através da palavra não é o “dado objetivo” produzido por uma medição ou contagem, mas um conceito de difícil compreensão, que entretanto é fundamental na perspectiva qualitativa: as “representações sociais”.

MINAYO define este conceito da seguinte maneira:

*“Representações Sociais é um termo filosófico que significa a reprodução de uma percepção anterior ou do conteúdo do pensamento. Nas Ciências Sociais são definidas como categorias de **pensamento**, de **ação** e de **sentimentos** que expressam a realidade, explicam-na, justificando-a ou questionando-a [grifos nossos]”* (Minayo: 1992, 158)

Mais adiante, neste mesmo trabalho, a mesma autora caracteriza as representações sociais como “senso comum, idéias, imagens, concepções e visão de mundo que os atores sociais possuem sobre a realidade”, os quais não são necessariamente conscientes para os indivíduos que as expressam (p. 173). Esta última colocação, em nosso entendimento, tem uma grande similaridade com o conceito de ideologia.

O pressuposto fundamental do conceito de representação social é “a hipótese de que cada indivíduo é portador da cultura e das subculturas às quais pertence e que é representativo delas” (Guy Michelat in Thiolent : 1987, 194). A partir deste entendimento, torna-se possível propor o estudo da “realidade social e coletiva por meio de narrativas individuais e vividos

singulares” (Schraiber: 1995, 65). Há um paradoxo implícito neste pressuposto, na medida em que é a partir do que há de “mais individual e de mais afetivo” que se tenta alcançar o que é “sociológico” (Guy Michelat in Thiollent : 1987, 197).

SCHRAIBER apresenta assim esta questão crucial da metodologias qualitativas:

“Essa metodologia introduz relativamente à aproximação do objeto em estudo, pois, duas ordens articuladas de questões: o social e o coletivo por referência ao individual-singular que o apreende; e o real objetivo por referência à dimensão subjetiva que o evidencia” (Schraiber: 1995, 65).

Portanto, as principais técnicas de coleta de dados nas abordagens qualitativas estão baseadas na narrativa dos indivíduos sobre os fatos estudados. Ou seja, estão baseadas na palavra.

Enquanto nas metodologias quantitativas busca-se a “verificação das regularidades” dos fenômenos observados, as qualitativas visam analisar os “significados que os indivíduos dão às suas ações”. Ou seja, a finalidade destes estudos é a “**compreensão do sentido** dos atos e das decisões dos atores sociais ou, então, dos vínculos indissociáveis das ações particulares com o contexto social em que se dão [grifo nosso]” (Chizzotti: 1991, 78). Isto identifica claramente a influência das correntes de pensamento compreensivas nas metodologias qualitativas.

As metodologias qualitativas não apresentam uma classificação estruturada como as quantitativas. Segundo YIN, o conhecimento sobre a variedade de desenhos nesta linha é “mais primitivo e limitado” que o conhecimento sobre os modelos experimentais e quase-experimentais, os quais vem recebendo “praticamente toda a atenção metodológica” nos últimos tempos (Yin: 1993, 78). Para YACH, a pesquisa qualitativa funciona como um “termo guarda-chuva que inclui uma extensa gama de métodos” (Yach: 1992, 605).

CHIZZOTI entende que a pesquisa qualitativa pressupõe a utilização de várias técnicas, e “não deve construir um modelo único, exclusivo e estandardizado”. Para ele, a pesquisa qualitativa deve ser uma “criação” do pesquisador, que deve elaborar uma “metodologia adequada ao campo de pesquisa, aos problemas que ele enfrenta com as pessoas que participam da investigação” (Chizzoti: 1991, 85).

Talvez por esta razão, boa parte dos modelos de estudo mencionados pelos autores confundem-se com a própria técnica de coleta de dados. Além do modelo mais conhecido - o “estudo de caso” -, YIN menciona a etnografia, a observação participante, o trabalho de campo (“fieldwork”), as histórias de vida (Yin: 1994, 288). CHIZZOTI, por sua vez, menciona a observação participante, a história de vida autobiográfica, etnográfica ou relatos de vida, as projeções de situações de vida, o “teatro da espontaneidade”, o jogo dos papéis, a entrevista (individual ou coletiva, dirigida ou não-diretiva) e a análise de conteúdo, que ora ele refere como “técnicas” de estudo (Chizzoti: 1991, 85), ora como “instrumentos de coleta de dados” (pp.89-90).

No campo da avaliação, as técnicas qualitativas muitas vezes derivam de técnicas para a tomada de decisão: técnica de consenso (grupo consenso para definir normas), técnica de atribuição de valores, método Delphi, grupo nominal e outras técnicas como grupo focal, técnica de coleta rápida de informação, informante chave.

Seja como for, o objetivo é reunir um “volume qualitativo de dados originais e relevantes, não filtrados por conceitos operacionais, nem por índices quantitativos” (Chizzoti: 1991, 85), que permita captar “as representações subjetivas dos participantes, favoreça a intervenção dos agentes em sua realidade ou organize a ação coletiva para transformar as condições problemáticas” (pp. 89-90).

Evidentemente, nas metodologias qualitativas a coleta de dados não é uma maneira pela qual o investigador “extrai” das pessoas investigadas, dados “objetivos” a respeito dos fenômenos sociais, mas uma forma de todos tomarem consciência do significado dos mesmos.

A coleta de dados nas abordagens qualitativas não é uma fase única e bem delimitada dentro de uma seqüência rigorosa de passos metodológicos. Os dados são coletados “**iterativamente**”²⁵, num processo de idas e voltas, nas diversas etapas da pesquisa e na **interação** com seus sujeitos” (Chizzotti: 1991, 89). O produto final são “depoimentos”, “narrativas”, “discursos”, que evidentemente não podem ser analisados através de procedimentos estatísticos (Queiroz: 1987; Schraiber: 1995). Ou seja, enquanto nas metodologias quantitativas os dados são apresentados de forma **numérica**, nas qualitativas são apresentados de forma **verbal** (Contandriopoulos et al: 1994, 87).

A sofisticação alcançada pelas análises estatísticas induz muitas pessoas a pensarem que a análise dos dados qualitativos é simples. PITTA alerta que, ao contrário, “enquanto se manejam dados quantitativos, a análise é mais simples, ficando muito mais complexa à medida que se introduzem elementos qualitativos” (Pitta: 1992, 45).

Segundo MINAYO, a expressão mais comum para referir-se o tratamento dos dados nas metodologias qualitativas é a “análise de conteúdo”, que é “mais do que um procedimento técnico” (Minayo: 1992, 199). A autora refere a existência de várias técnicas, das quais descreve cinco - Análise de Expressão, Análise de Relações, Análise Temática e Análise da Enunciação (pp. 204-10).

Um ponto interessante a destacar é que boa parte do trabalho de análise nas metodologias qualitativas inicia na quantificação. Assim, na análise temática tradicional, inicialmente se estabelece uma “unidade de registro” (palavra, frase, tema, personagem, acontecimento), que são tomadas como “resultado bruto” sobre os quais se realiza “operações estatísticas simples (percentagens) ou complexas (análise fatorial)” (Minayo: 1992, 210).

Entretanto, o objetivo destes procedimentos quantitativos não é apresentar resultados estatísticos, mas evidenciar conteúdos, manifestos ou latentes, no discurso dos indivíduos. O essencial na análise de conteúdo é vincular o “texto”, escrito ou narrado, ao “contexto” em que é formulado, através de “procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens” (L. Bardin apud Minayo: 1992, 199).

Vale a pena reproduzir como MINAYO descreve o procedimento geral na análise de conteúdo:

“Do ponto de vista operacional, a análise de conteúdo parte de uma leitura de primeiro plano para atingir um nível mais aprofundado: aquele que ultrapassa os significados manifestos. Para isso a análise de conteúdo em termos gerais relaciona estruturas semânticas (significantes) com estruturas sociológicas (significados) dos enunciados. Articula a superfície dos textos descrita e analisada com os fatores que determinam suas características: variáveis psicossociais,

²⁵Cf. Dicionário Aurélio: “**iterativo**: 1. Que serve para iterar. 2. Repetido, reiterado”; “**iterar**. Tornar a fazer ou a dizer; repetir, reiterar.

contexto cultural, contexto e processo de produção da mensagem". (Minayo: 1992, 203)

CONTANDRIOPOULOS e colaboradores apresentam um seqüência de passos para a realizar a análise dos dados. Primeiro, fazer a "preparação e a descrição do material bruto", para organizar o grande volume de dados que geralmente são produzidos nas pesquisas qualitativas - transcrever gravações de entrevistas, elaborar lista cronológica de documentos, registrar fontes e proveniência do material. Segundo, reduzir o volume de dados, ressaltando os componentes do discurso associados aos temas de interesse, na qual podem ser utilizados três procedimentos - a redação de resumos, a codificação (de categorias definidas *a priori*) e a marcação de temas e a realização de agrupamentos, realizadas de maneira indutiva. Para estes autores, a análise de conteúdo corresponde à esta fase de redução dos dados. Terceiro, a escolha do modo de análise, para interpretar os dados - o emparelhamento de um modelo teórico, a elaboração de uma explicação e a análise histórica ou a análise de séries temporais - que discutiremos adiante, no final desta seção, ao tratar dos estudos de caso (Contandriopoulos et al: 1992, 87-90).

Como vemos, os procedimentos de coleta e análise de dados são bastante complexos nas metodologias qualitativas. Porém, o rigor de que se fala nas metodologias qualitativas não é uma "asséptica objetividade de uma ciência desprovida de sujeitos, mas a compreensão dos fundamentos e dos limites do método", pois o conhecimento científico é visto apenas como o "ato **possível** de sujeitos históricos em interação" (Jovchelovitch & Guareschi: 1994, 22).

O principal problema destas metodologias é o risco de reduzir a realidade "à concepção que os homens fazem dela" (Minayo in Guareschi & Jovchelovitch: 1994, 110). A questão fundamental é como fugir das "distorções trazidas pela subjetividade do narrador" (Queiroz: 1987, 283).

Veremos este assunto na seção referente à validade do conhecimento produzido pelas metodologias qualitativas. Por enquanto, voltamos a lembrar que a tentativa de objetivação, que implica no uso rigoroso de "instrumental teórico e técnico adequado", visa justamente "reduzir a incursão excessiva dos juízos de valor na pesquisa", mas também implica em "repudiar o discurso ingênuo ou malicioso da neutralidade" (Minayo: 1992, 22, 35-36).

Além disto, vale lembrar as colocações de PITTA:

“...mesmo as metodologias estruturadas envolvem um importante componente de subjetividade que tem repercussões muito diretas na escolha dos instrumentos que, se abertos, envolvem um componente subjetivo maior, e, se fechados, ‘mais objetivos’, estarão fatalmente mais limitados no seu alcance” (Pitta: 1992, 45).

Por fim, caberia lembrar algumas vantagens e desvantagens das metodologias qualitativas e quantitativas. Segundo QUEIROZ, os cientistas sociais são unânimes em reconhecer a “riqueza dos dados” que são captados pelas metodologias qualitativas. Para ela, as técnicas qualitativas apreendem “a vivacidade dos sons, a opulência dos detalhes, a quase totalidade dos ângulos que apresenta todo fato social”, enquanto as técnicas quantitativas reduzem a realidade social “à aridez dos números”, parecendo “amputá-la de seus significados”. Porém, as técnicas qualitativas implicam em um tempo mais longo na coleta e análise dos dados, além do que dependem muito do desempenho pessoal do pesquisador (Queiroz: 1987, 279, 272).

Em resumo, poderíamos afirmar que o ponto central do debate entre as metodologias quantitativas e qualitativas é a discussão sobre a "subjetividade" ou a "objetividade" dos procedimentos de coleta de dados. O esforço para objetivar o que é percebido sempre subjetivamente, constitui, historicamente, o esforço característico da ciência, embora seja “continuamente desafiado pelo caráter dialético” da mesma (Nunes: 1995, 28, referindo-se à idéias de Hilton Japiassu).

1.3.1.2.4. Uma abordagem integrada

Na seqüência, vamos tratar de alguns aspectos levantados por YIN, que desenvolveu uma nova perspectiva para os estudos de caso dentro do campo da avaliação em geral. Para este autor, no futuro, a maioria das avaliações deverá utilizar concomitantemente as

abordagens quantitativas e qualitativas, e o modelo de estudo de caso teria uma “força especial” para fazer esta integração (Yin: 1993; Yin: 1994).

No estudo de caso clássico, a unidade de análise é um único indivíduo, ou seja, o estudo está voltado para os fenômenos no âmbito individual. O autor propõe que se passe a ver os projetos, as organizações e os programas também como “casos” únicos, onde cada um corresponde a uma unidade de análise, embora o estudo esteja voltado para os fenômenos em âmbito coletivo. Ou seja, a “unidade de análise seria o programa em si mesmo” (Yin: 1993). Ele explica desta forma como o modelo de estudo de caso integra as metodologias quantitativas e qualitativas:

“Um exemplo desta integração é quando uma única avaliação, de um único programa, pode consistir, entretanto, de múltiplos sub-estudos. Cada sub-estudo pode ter seu próprio desenho de pesquisa e usar diferentes métodos, mas a agregação destes sub-estudos em uma única estrutura de avaliação ocorrerá de forma melhor quando esta estrutura singular for considerada como uma estrutura de estudo de caso” (Yin: 1994).

Evidentemente, esta concepção de estudo de caso - estrutura que integra múltiplos sub-estudos - é muito semelhante à noção de triangulação, que é a forma privilegiada através da qual as Ciências Sociais fazem a integração entre as metodologias qualitativas e quantitativas, conforme vimos na seção anterior.

O autor desenvolve suas idéias partindo da comparação entre os estudos experimentais, quase-experimentais e estudos de caso. Em primeiro lugar, ele apresenta os “requisitos básicos do desenho experimental” (Yin: 1993, 77-79):

1. Ter uma unidade de análise com um número de “sujeitos” (de análise) suficiente para produzir múltiplos pontos de dados (“data points”) para cada variável do estudo²⁶;
2. Ter um número limitado de variáveis de interesse, usualmente muito menor do que o número de pontos de dados disponíveis;

²⁶Segundo Contandriopoulos e colaboradores, é preciso que o número de unidades de análise seja muito maior (no mínimo dez vezes) do que o número de atributos (graus de liberdade do estudo), para aumentar o poder explicativo do estudo. Caso contrário, as análises estatísticas somente conseguirão evidenciar grandes diferenças, tornando-se irrelevantes (Contandriopoulos et al.: 1994, 41).

3. O investigador deve poder controlar e manipular as variáveis de interesse, e
4. Fazer uma distribuição aleatória dos sujeitos no grupo experimental (submetido à intervenção) ou no grupo controle (não submetido à intervenção).

Em seguida, ele aponta que somente o estudo experimental cumpre os quatro requisitos, enquanto que os estudos quase-experimentais preenchem apenas o primeiro e o segundo. A característica dos estudos de caso é que não cumpre nenhum destes requerimentos (Ver Anexo VI).

O autor chama atenção que, **por definição**, o estudo de caso contraria o primeiro e o segundo requerimentos, pois a característica dos estudos de caso é que o número de “sujeitos” ou “pontos de dados” é tão pequeno que não pode superar o número de variáveis de interesse. Ouçamos como ele descreve estes estudos:

“Um estudo de caso clama por uma quantidade intensiva de dados (isto é, um número grande de variáveis) sobre um pequeno número ou uma unidade única de análise (o caso). Os investigadores usam este método porque certos tópicos da pesquisa são tão complexos que o fenômeno de interesse não é facilmente distinguível das condições de seu contexto, e conseqüentemente são necessários dados sobre ambos” (Yin: 1994).

Mais adiante, ele coloca que os estudos de caso serão necessários sempre que “as questões do estudo levem à criação de mais variáveis do que pontos de dados”, e isto ocorre quando o estudo está voltado para “compreender as condições contextuais [pois] o contexto será tão rico que o número de variáveis sempre excederá largamente o número de pontos de dados” (Yin: 1994, 81). Este é justamente o caso dos estudos que avaliam projetos, programas e organizações.

Embora a taxonomia dos desenhos de estudo de caso ainda não esteja bem desenvolvida, ele entende que já são conhecidas duas características importantes destes desenhos.

A primeira delas é que existem pelo menos dois sub-tipos de estudo de casos - os desenhos de caso único e os desenhos de múltiplos casos. Quando o desenho é de casos múltiplos, a seleção dos casos não é feita pela “lógica da amostragem”, mas pela “lógica da

replicação”, isto é, o caso é selecionado, intencionalmente, para reproduzir as mesmas condições (ou diferentes condições) de intervenção, de modo que o estudo aproxima-se do terceiro requerimento - controle das variáveis de interesse - embora não haja uma manipulação direta das condições de intervenção. O autor reconhece que é necessário uma certa dose de sorte e de circunstâncias favoráveis para encontrar os “casos” candidatos à replicação²⁷ (Yin: 1993).

A segunda característica é, em nosso entendimento, a mais esclarecedora deste modelo:

*“...em ambos os tipos de estudo de caso, a generalização dos resultados é feita para a **teoria**, e não para **populações**. Em outras palavras, o desenho correto de estudo de caso identifica uma teoria que o estudo do caso está tentando testar, ao invés de considerar o estudo dos casos individuais como pontos de dados ou parte de uma amostra [grifos nossos]” (Yin: 1993).*

Ou seja, ao invés de fazer inferências para populações, conforme a “lógica da amostragem” da estatística, dos modelos experimentais e quase-experimentais, os estudos de caso estão desenhados para fazer inferências sobre teorias. No sub-tipo de estudos de casos múltiplos, isto é feito através da “lógica da replicação” do funcionamento destes modelos teóricos em cada contexto estudado.

Este novo enfoque do estudo de caso foi o referencial para o desenvolvimento de uma nova tipologia dos estudos em saúde, desenvolvida por autores canadenses, que integra em uma mesma estratégia de estudo as metodologias quantitativa e a qualitativa, denominada de “pesquisa sintética”.

CONTANDRIOPOULOS e colaboradores definem assim esta grande abordagem de pesquisa integrada:

*“A **pesquisa sintética** é aquela que, para explicar e prever comportamentos ou fenômenos complexos, examina o conjunto das relações onde intervêm, simultaneamente, diversas variáveis dependentes e independentes num modelo de*

²⁷Vemos em CONTANDRIOPOULOS e colaboradores que, nos estudos de casos múltiplos, os dados oriundos dos diversos casos são analisados através de uma “análise transversal”, que YIN denomina “cross-case replication logic”, para verificar se “os modelos observados se reproduzem” nos diferentes casos (Contandriopoulos et al.: 1994, 90; Yin: 1993).

relações interdependentes. Assim, a pesquisa sintética se diferencia da pesquisa experimental, onde se examina de maneira linear, freqüentemente bivariada, quase sempre unidirecional, a relação entre variáveis de interesse. Em outras palavras, a pesquisa experimental se baseia numa abordagem analítica, cartesiana, positivista, enquanto a pesquisa sintética pretende ser sistêmica (Contandriopoulos et al.: 1994, 40).

Observa-se que esta abordagem incorpora uma visão sistêmica da realidade social, superando as críticas à linearidade do modelo experimental, que vimos anteriormente.

Os autores classificam duas abordagens de pesquisa sintética: os “estudos de caso” (caso único ou casos múltiplos, ambos os tipos com um ou mais níveis de análise) e os “estudos comparativos”, onde colocam os “modelos estruturais” e a “meta-análise” (Contandriopoulos et al.: 1994, 37;40).

Nesta abordagem “sintética”, tanto a análise dos dados como a validade do estudo estão relacionadas à idéia de que a inferência será feita para o modelo teórico testado, e não para populações.

A análise dos dados é feita segundo três grandes “princípios de análise” dos dados qualitativos, a partir dos quais definem-se os modelos de análise (Contandriopoulos et al.: 1994, 89):

- o “emparelhamento de um modelo”, que se apóia na “comparação de uma configuração teórica predita com uma configuração empírica observada”;
- a “elaboração de uma explicação”, que se refere a um modo de análise iterativo, no qual, através do retorno constante aos dados, o pesquisador confronta a realidade empírica com diversas teorias explicativas, construindo, progressivamente e com um mínimo de formalização teórica, a melhor explicação possível do fenômeno estudado;
- a “análise histórica” ou a “análise de séries temporais”, que consiste em “formular predições sobre a evolução de um fenômeno com o tempo”.

Nesta perspectiva integradora, a validade do estudo está baseada essencialmente na qualidade de sua “articulação teórica” e na adequação do modo de análise dos dados ao modelo teórico que está sendo testado (Contandriopoulos et al.: 1994, 40).

Em suma, poderíamos destacar duas concepções fundamentais nesta proposta de integração entre as metodologias quantitativas e qualitativas, através do estudo de caso, originalmente formulada por Yin e desenvolvida por Contandriopoulos e colaboradores: primeiro, que a metodologia de integração está embasada, de certa forma, na triangulação; segundo, que a inferência é feita para teorias, e não para populações.

Estas duas concepções fazem com que a abordagem deste tipo de estudos seja muito adequada para avaliar as intervenções de saúde (projetos, programas, serviços), onde o modelo teórico que é testado no estudo é o próprio modelo teórico que embasa a intervenção (modelagem da intervenção), seja através da análise profunda do funcionamento de uma única intervenção (estudo de caso único), seja através da comparação entre diferentes modelos de intervenção, ou da comparação do funcionamento de um mesmo modelo em diferentes contextos (modelo de casos múltiplos).

1.3.1.3. A questão da validade

A terceira questão da problemática geral das Ciências Sociais é a da validade científica do conhecimento produzido neste campo, ou seja, a “cientificidade” deste conhecimento. Embora esta discussão seja, atualmente, o ponto central em qualquer área de conhecimento, nas Ciências Sociais o problema torna-se mais agudo pela subjetividade inerente ao sujeito e ao objeto da avaliação, ou seja, ao próprio homem e sua ação.

A reflexão sobre a questão da validade permite identificar, a nosso ver, os aspectos cruciais dos estudos em avaliação. A subjetividade inerente à tarefa de avaliar, que implica em emitir juízo de valor, pode dar margem a uma série de questionamentos sobre a validade dos estudos neste campo.

Além disto, o caráter essencialmente prático ou aplicado destes estudos, de uma certa forma, torna-os mais vulneráveis à crítica que é feita a partir dos critérios de cientificidade predominantes, fundamentados em procedimentos de validação oriundos de modelos de estudo experimentais, dificilmente aplicáveis no campo social.

No campo específico da pesquisa em avaliação, a questão da validade é sem dúvida o ponto crucial. Como vimos no conceito de avaliação, avaliar é emitir um “juízo de valor”, que significa “enunciar uma apreciação” sobre os fatos, opondo-se ao “juízo de realidade”, que significa “enunciar um fato ou uma relação entre fatos”. Ora, o propósito original da ciência está essencialmente vinculado ao juízo de realidade - qualquer ciência busca “apreender” o fato que é seu objeto de estudo -, no caso das Ciências Sociais, o fato social.

Por isto, o principal problema da pesquisa em avaliação, como de resto qualquer pesquisa, tem sido a tentativa de reduzir a subjetividade dos juízos de valor, buscando apreender a realidade através da “medição” mais exata possível do fenômeno que está sendo avaliado.

A questão da medida, ou melhor, da qualidade das medidas ou aferições feitas sobre o objeto de estudo, tem sido o problema o crucial no campo da avaliação, pois estas medições serão a base do juízo avaliativo. SUCHMAN chega a colocar que os conceitos de avaliação e de validade são similares, havendo mesmo um “paralelo verdadeiro” entre eles (Suchman: 1967, 48). DEMO coloca bem esta questão:

*“A avaliação é um dos desafios científicos que mais escancaram os limites da ciência, tanto na dificuldade de avaliar **isentamente**, pois no fundo é impraticável, como na impossibilidade concreta de encontrar critérios consensuais e definitivos. Em vista disso, a avaliação se refugia em critérios quantitativos, porque acredita que são menos manipuláveis e suspeitos” (Demo: 1992, 70-71).*

Porém, como vimos nas seções anteriores, os estudos em avaliação necessariamente devem incluir formas de apreender tanto a subjetividade, que norteia a intencionalidade da ação humana, quanto a objetividade dos efeitos desta ação sobre a realidade social, o que implica na necessidade de utilizar, integradamente, metodologias de pesquisa quantitativas e qualitativas, cujos critérios de validação são diferentes.

MINAYO coloca de forma clara a pergunta fundamental da validade:

“Como podemos garantir a desejada coincidência entre o pensamento sobre a realidade e a própria realidade? Tais perguntas são inevitáveis e fundamentais

dentro do campo científico, embora se aceite que a produção intelectual é sempre um ponto de vista a respeito do objeto” (Minayo: 1992, 238).

Veremos a seguir que o conceito de “validade” é muito elástico, pois aplica-se a diferentes aspectos e momentos do processo de produção do conhecimento, desde os níveis mais gerais até os mais específicos.

1.3.1.3.1. Validade nas metodologias quantitativas

As considerações sobre a validade de um estudo nas **metodologias quantitativas** são fruto da concepção de cientificidade do positivismo, cuja noção de objetividade é o “objetivismo”, embasada no empirismo e na quantificação. Em razão disto, o principal critério de validade científica é “a exigência colocada no método e nas técnicas de ‘captação’ da realidade”, que deve garantir a “objetivação” e a “verificabilidade” do conhecimento produzido (Minayo: 1992, 239). Coerentemente com o desenvolvimento histórico desta corrente, estes critérios estão embasados no padrão de observação das ciências físicas e biológicas, e na idéia do experimento como modelo de estudo.

Nas metodologias quantitativas, a questão da validade está referida essencialmente ao desenho do estudo, sempre tomando como referência o desenho experimental, onde as técnicas de amostragem desempenham um papel fundamental, pois podem aumentar o potencial do estudo para fazer as inferências estatísticas. A validade de estudo está essencialmente baseada na capacidade de evitar “vieses”, ou desvios, que possam interferir na relação estatística entre a variável independente e a variável dependente, distorcendo assim os resultados do estudo.

ALMEIDA FILHO&ROUQUAYROL colocam que a validade comporta dois tipos de componentes: um componente “conceitual” ou “teórico”, que é resultante de um modelo teórico sobre o fenômeno estudado, a partir do qual define-se o desenho do estudo - neste sentido, o conceito de validade constitui “um atributo do desenho da investigação”; e um

componente “pragmático” ou “operacional”, que está relacionada aos instrumentos de medida (Almeida Filho&Rouquayrol: 1992, 31;38). No primeiro caso, colocam-se as questões da validade interna, onde o esforço é dirigido no sentido de evitar possíveis “vieses” ou desvios no desenho do estudo; no segundo caso, coloca-se a questão da validade de medida, que refere-se ao controle da qualidade dos instrumentos de medida.

O conceito de validade está relacionado com diferentes aspectos e momentos dos estudos quantitativos, que correspondem a diferentes “níveis de validação”, conforme tentaremos demonstrar a seguir.

Do ponto de vista quantitativo, a validade de um estudo tem dois aspectos gerais, que correspondem ao primeiro nível:

- a **validade externa**, que é “o grau pelo qual os resultados de uma observação são aplicáveis em outras situações”, ou seja, é o que permite generalizar os resultados de um estudo, feito em uma população, para outras populações, que não participaram do estudo (capacidade de generalização);

- a **validade interna**, que é “grau pelo qual os resultados de uma aferição correspondem ao estado verdadeiro dos fenômenos que estão sendo medidos”, que é tanto maior quanto mais se consegue evitar os “vieses”, “bias” ou “vícios” do estudo, ou seja, a distorção nos resultados em função dos “erros aleatórios” (decorrentes do acaso) ou dos “erros sistemáticos” (decorrentes de problemas no desenho do estudo ou na aferição das medidas) (Fletcher; Fletcher & Wagner: 1989, 32; Duncan&Schmidt: 1986).

O erro aleatório não pode ser evitado, mas pode ser estimado, sendo do domínio da estatística. A tentativa de controlar os erros sistemáticos é a essência da questão da validade. Os esforços são dirigidos para controlar os “vícios”, através da cuidadosa modelagem do desenho do estudo. Este é o segundo nível de validade, e corresponde ao ser componente “teórico”.

Há inúmeros tipos de vieses que foram identificados nos desenhos de estudo quantitativo, que de uma forma geral são classificados em três tipos:

- 1 - **viés de seleção**, isto é, quando há problemas na seleção do grupo de estudo que distorcem os resultados. No caso dos estudos controlados, este viés afeta a comparabilidade dos resultados entre o grupo experimental e o grupo controle;

2 - **viés de confusão**, quando dois fatores, ou variáveis, estão associados e “o efeito de um é confundido ou distorcido pelo efeito do outro”. Quando não se considera a existência ou o efeito da variável de confusão ao formular o desenho do estudo, não se pode controlar seu efeito sobre os resultados, o que cria problemas para o teste de hipótese nos estudos analíticos;

3 - **viés de aferição**, quando há erro sistemático na medida de uma variável (Fletcher; Fletcher & Wagner: 1989, 26; Duncan&Schmidt: 1986).

É dentro do problema dos vícios de aferição que chegamos ao terceiro nível de validação nas metodologias quantitativas, dentro do qual se situam as questões de validade mais pertinentes aos estudos avaliativos. Estas questões dizem respeito à **qualidade das aferições** ou **medidas** que permitem dimensionar os efeitos de uma intervenção.

Como vimos anteriormente, a questão da medida é o ponto fundamental dos estudos avaliativos, pois ela permite apreciar a magnitude do efeito de uma intervenção - é a “prova” que fornece o subsídio para emitir um juízo sobre a mesma.

Segundo FLETCHER, FLETCHER& WAGNER, para “descrever a qualidade das aferições” são utilizados dois conceitos: o conceito de **validade de medida** (“validity”) e o conceito de **confiabilidade de medida** (“reliability”) (Fletcher; Fletcher & Wagner: 1989, 45).

Estes conceitos são de difícil compreensão, causando confusões muito freqüentemente. De uma forma bem geral, poderíamos dizer que as perguntas que estão por trás destes conceitos são as seguintes: até que ponto os instrumentos de coleta de dados estão medindo o que deve ser medido - validade de medida - e até que ponto estão medindo com precisão - confiabilidade.

Ou seja, ambos os conceitos referem-se aos **instrumentos de coleta de dados**, sendo que a validade de medida refere-se, de certa forma, à adequação do instrumento em relação à variável que se quer medir, enquanto que a confiabilidade de medida refere-se às próprias condições dos instrumentos para realizar esta medida com precisão.

Lembramos que o conceito de validade, neste nível, corresponde ao componente “pragmático” ou “operacional” da validade, a qual também comporta um nível superior - o componente “conceitual” ou “teórico”, que refere-se ao controle dos vieses. O conceito de confiabilidade, por sua vez, seria “uma qualidade exclusiva dos instrumentos de coleta de dados” (Almeida Filho&Rouquayrol: 1992, 38).

Vejamos inicialmente algumas definições do conceito de confiabilidade:

“A confiabilidade de um instrumento de medida é a sua capacidade de reproduzir um resultado de forma consistente no tempo e no espaço, ou com observadores diferentes, quando for utilizado corretamente” (Contandriopoulos et al.: 1994, 80).

“Confiabilidade é a extensão em que medidas repetidas de um fenômeno relativamente estável situam-se próximas umas das outras” (Fletcher, Fletcher & Wagner: 1989, 46).

Alguns autores utilizam os termos “reprodutibilidade” e “precisão”²⁸ como sinônimo de confiabilidade (Fletcher, Fletcher & Wagner: 1989, 46; Almeida Filho&Rouquayrol: 1992, 38), enquanto outros utilizam o termo “fidelidade” (Bruyne et al.: 1982, 83).

Segundo CONTANDRIOPOULOS et al., “por definição, a apreciação da confiabilidade se baseia na repetição da medida e na comparação dos resultados obtidos”. Quando o fenômeno que está sendo medido é estável, avalia-se a confiabilidade verificando a similaridade entre a diferentes medições ao longo do tempo; quando o fenômeno é instável ou “transitório (por exemplo, a satisfação, a ansiedade, etc.) os resultados poderão variar entre uma medição e outra, sem que isto se deva ao instrumento de medida” (Contandriopoulos et al.: 1994, 80).

A análise da confiabilidade dos instrumentos quantitativos de coleta de dados é feita através de testes estatísticos (uso de coeficientes em modelos estatísticos), cuja escolha depende do tipo de variável que está sendo medida: contínua, ordinal ou nominal” (Contandriopoulos et al.: 1994, 11-120).

A confiabilidade dos instrumentos de medida é uma característica “**independente** da questão da pesquisa [grifo nosso]”, mas pode “variar em função das populações estudadas” (Contandriopoulos et al.: 1994, 82). Para BRUYNE e colaboradores, a avaliação da confiabilidade, que ele denomina de fidelidade, não visa apenas verificar a fidelidade da

²⁸CONTANDRIOPOULOS e colaboradores, por sua vez, utilizam o termo “precisão” para se referir ao conceito de validade de medida, conforme vemos na seguinte afirmação: “A confiabilidade é a capacidade e um instrumento medir fielmente um fenômeno. A validade é a capacidade de um instrumento medir com precisão o

medida em si, mas “concerne igualmente a todo o procedimento da mensuração, as condições sob as quais ela é praticada e seu modo de aplicação” (Bruyne et al.: 1982, 83).

O conceito de validade de medida é menos preciso. Segundo CONTANDRIOPOULOS e colaboradores, este conceito ainda não teria alcançado o mesmo “grau de cristalização” que o de confiabilidade. Para estes autores, a noção de validade é “muito mais abstrata”, e esta característica dos instrumentos é muito “**dependente** do contexto (social, cultural, lingüístico) de utilização dos instrumentos [grifo nosso]”. Por isso, estimar a validade de um instrumento é “mais complexo que avaliar sua confiabilidade” (Contandriopoulos et al.: 1994, 82-83).

Vejamos algumas definições de validade:

“... a validade se define como a capacidade de um instrumento medir o fenômeno em estudo, isto é, a adequação existente entre as variáveis escolhidas e o conceito teórico a ser medido” (Contandriopoulos et al.: 1994, 83).

“... validade é o grau pelo qual os resultados de uma aferição correspondem ao estado verdadeiro dos fenômenos que estão sendo medidos” (Fletcher; Fletcher & Wagner: 1989, 46).

O termo “acurácia” tem sido empregado como sinônimo de validade (Fletcher; Fletcher & Wagner: 1989, 45).

Os conceitos de confiabilidade e validade das medidas ficam mais compreensíveis no esquema do “tiro ao alvo”, conforme vemos no Anexo VII.

Dentro do conceito de validade de medida desdobra-se o quarto nível da questão da validade. É neste nível que encontramos as questões que dizem respeito à **qualidade dos indicadores**, que são a base empírica dos estudos em avaliação. E é neste nível que a pretensão de objetividade começa a apresentar problemas mais difíceis de serem equacionados.

Referindo-se à prática clínica, FLETCHER, FLETCHER& WAGNER colocam que as observações que podem ser “aferidas por meios físicos são relativamente fáceis de serem validadas”, bastando para isto aplicar um teste laboratorial e comparar a medida encontrada

fenômeno estudado [grifo nosso]”. (Contandriopoulos et al.: 1994, 79) Acreditamos que parte da confusão ocorra por problema de tradução destes termos do inglês para o português.

com algum padrão já estabelecido. Entretanto, quando os fenômenos que se quer medir são “sentimentos, atitudes, conhecimentos, crenças”, é necessário utilizar entrevistas ou questionários (cujas perguntas os autores denominam de “constructo”), que fazem uma aferição indireta do fenômeno, através da informação fornecida pelos indivíduos. As respostas a essas perguntas ou “constructo” são então “convertidas em números e agrupadas de modo a formar uma escala”. Nestes casos, em que as medidas “não podem ser confirmadas diretamente através dos sentidos”, é necessário utilizar algumas “estratégias gerais” através das quais, indiretamente, se pode “armar evidências para confirmar ou refutar” a validade de uma escala de medidas (Fletcher; Fletcher & Wagner: 1989, 45-46).

É em relação à validade destes “constructos”, isto é, das perguntas através das quais pretende-se verificar, indiretamente, o fenômeno estudado, que colocam-se os três aspectos da validade de medida, ou melhor, as três “estratégias” de validação: validade de construção, validade de conteúdo e validade de critério.

Embora os autores consultados não tenham estabelecido uma ordem de hierarquia entre estes três aspectos da validade de medida, entendemos que estes aspectos contém, de forma concentrada, todo o esforço da ciência para transformar uma construção teórica - um conceito - em algo concretamente mensurável - uma variável. Ou seja, em nosso entendimento haveria uma hierarquia entre estas estratégias de validação, que inicia com uma estratégia mais geral (validade de construção), passa por um nível intermediário (validade de conteúdo) e chega ao nível mais operacional (validade de critério).

Para CONTANDRIOPOULOS e colaboradores, a **validade de construção**

“... trata da relação entre os conceitos teóricos e sua operacionalização. Ela concerne, pois, à relação epistêmica que deveria existir entre um conceito e sua medida” (Contandriopoulos et al.: 1994, 84).

A estratégia da validade de construção permite três formas de validação: a validação nomológica ou teórica; a validação de traço (que tem como variante a validação multi-traço ou multi-método) e a validação fatorial (Contandriopoulos et al.: 1994, 85-86).

A validação nomológica ou teórica consistiria em estabelecer algumas hipóteses *a priori* sobre “os resultados que deveriam ser observados com a ajuda do instrumento, em

certas condições e para certas populações”. Caso os resultados esperados sejam observados, confirma-se a hipótese teórica, sugerindo que o instrumento seja válido (p. 84). A validação de traço consistiria em “correlacionar os resultados obtidos pelo teste que se quer validar com os resultados obtidos por um outro teste medindo aproximadamente a mesma construção”. Caso observe-se correlação de resultados, tem-se um indício de validade (p. 86). A validação fatorial apóia-se em uma “análise fatorial dos elementos, constituindo-se o índice que se quer avaliar” (p. 86).

Segundo FLETCHER, FLETCHER& WAGNER, existe validade de construção quando os resultados de uma escala de medida “variam de acordo com a presença de outros indicadores estabelecidos”, que sabidamente estão associados ao fenômeno em estudo (Fletcher; Fletcher & Wagner: 1989, 46).

A **validade de conteúdo**, por sua vez, é definida como segue (grifos nossos):

*“A validade de conteúdo consiste em julgar em que proporção os **itens** relacionados para medir uma construção teórica representam bem todas as facetas importantes do conceito a ser medido”* (Contandriopoulos et al.: 1994, 83).

*“Validade de conteúdo é a medida pela qual um determinado método de aferição inclui todas as **dimensões** da construção que está sendo aferida e nada mais”* (Fletcher; Fletcher & Wagner: 1989, 45).

Na validade de conteúdo, verifica-se a coerência ou adequação aparente que existe entre o que se quer medir (ou “construção a ser medida”) e o instrumento de medida escolhido. Por isso, o aspecto crucial neste tipo de validade é a formulação das questões, que deve ser adequada ao nível de compreensão e à linguagem dos entrevistados. Para estabelecer a validade de conteúdo pode-se recorrer à opinião de especialistas ou de alguns membros da população alvo (neste caso, pré-teste do instrumento) (Contandriopoulos et al.: 1994, 83).

Em nosso entendimento, o conceito de validade de conteúdo refere-se ao segundo passo da operacionalização de um conceito (no caso da avaliação, de um indicador), que consiste na identificação do elenco de variáveis que permitem desdobrar o mesmo em dimensões mensuráveis.

Por fim, temos as seguintes colocações a respeito do conceito de **validade prática** ou de **critério**:

“Este tipo de validade se refere à capacidade do instrumento em medir alguma coisa relacionada com um critério de interesse, frequentemente um comportamento” (Contandriopoulos et al.: 1994, 83).

“A validade de critério é estabelecida mostrando-se que a medida prediz um fenômeno diretamente observável” (Fletcher; Fletcher & Wagner: 1989, 46).

Segundo CONTANDRIOPOULOS e colaboradores, a validade prática relaciona a medida obtida com um “critério empírico de resultado”. Quando este critério situa-se no futuro, fala-se de “validade preditora”; quando é contemporâneo, fala-se de “validade concorrente” ou “concomitante”. Nos estudos epidemiológicos, existem dois indicadores da validade de critério - a sensibilidade e a especificidade - bastante conhecidos especialmente na validação de testes diagnósticos (Contandriopoulos et al.: 1994, 84). O conceito de validade de critério corresponde ao nível mais operacionalizável, mais “prático” de um conceito teórico, e seu aspecto fundamental é a possibilidade de prever a ocorrência real do fenômeno observado.

HARTZ&CAMACHO resumem claramente os três aspectos da validade de medida:

“Espera-se que os métodos (e seus respectivos instrumentos”de medida) contemplem todas as dimensões do fenômeno ou atributo que se quer medir (validade de conteúdo), que as medidas concordem com critérios definidos por outros métodos tomados como referência (validade de critério) e que exista uma teoria estabelecida que permita correlacionar o conceito que está sendo medido com a medida empírica gerada pelo processo de mensuração (validade de construção” (Hartz&Camacho: 1996).

Nestes três últimos níveis de validação - validade de construção, validade de conteúdo e validade de critério - estão as questões mais relevantes para a avaliação, pois se referem aos indicadores através dos quais pretende-se “medir” o efeito da intervenção. De uma forma geral, vemos que a validade de construção refere-se a um plano conceitual, onde é definido o indicador; a validade de conteúdo refere-se ao elenco de variáveis que dimensionam o

indicador; e a validade de critério ou prática refere-se à questão do “**ponto de corte**”, dentro de uma escala de medidas, através do qual se estabelece uma norma ou padrão daquele indicador - o critério empírico de “bom” - que subsidiará o juízo avaliativo²⁹. Para visualizar a questão da validade nas metodologias quantitativas em seu conjunto, ver Anexo VIII.

É fundamental esclarecer que os resultados obtidos com estas diferentes estratégias de validação não permitem fazer afirmações categóricas sobre a validade de um instrumento de medida, pois apenas fornecem indícios da mesma (Contandriopoulos et al.: 1994, 85-86; Fletcher; Fletcher & Wagner: 1989, 45). Ou seja, todo o processo de mensuração será sempre uma “tentativa de aproximação do real de um atributo, cujos valores observados trazem necessariamente embutidos desvios que resultam das imperfeições dos métodos” (Hartz&Camacho: 1996).

Por outro lado, BRUYNE e colaboradores chamam atenção que a validade é função do “conjunto das operações metodológicas, e não unicamente da precisão das medidas”. A confiabilidade e a validade das medidas, ou seja, “a quantificação como tal” não pode ser a garantia da validade dos resultados sem cair na “ficção da verdadeira medida” ou na “falsa precisão”, que ocorre quando a medida é efetuada “sem conceitualização prévia suficiente ou quando o procedimento de quantificação é impelido para além dos limites permitidos” (Bruyne et al.: 1982, 82-83).

1.3.1.3.2. Validade nas metodologias qualitativas

Vimos até aqui que os procedimentos de validação nas metodologias quantitativas enfatizam dois aspectos: as questões referentes ao desenho do estudo, cujos critérios estão embasados no modelo experimental, dificilmente aplicáveis ao estudo da realidade social, e as questões referentes à precisão das medidas, cujo pressuposto é a possibilidade de medir

²⁹Este raciocínio está fundamentado nas colocações da Prof^a Zulmira M.A. Hartz, no curso “Epidemiologia na Avaliação dos Programas de Saúde”, durante o III Congresso de Epidemiologia da ABRASCO, realizado em Salvador, Bahia, de 24 a 28 de abril de 1995.

diretamente as variáveis em estudo, poucas vezes realizável quando estuda-se os fenômenos sociais, essencialmente subjetivos. Por estas razões, LADRIÈRE coloca que “o domínio dos fenômenos sociais pode dar lugar a uma imagem da cientificidade distinta da dos fenômenos físicos” (Bruyne et al.: 1982, 14, Prefácio de Jean Ladrière).

Nas metodologias qualitativas, que buscam apreender o significado da ação humana, a questão da validade tem outra conotação, embora, de certa forma, contemple os mesmos aspectos das metodologias quantitativas.

Para começar, vemos em CHIZZOTTI algumas colocações sobre validade que, a nosso ver, reproduzem as mesmas preocupações dos estudos quantitativos:

“Os dados qualitativos deverão ser validados segundo alguns critérios: fiabilidade (independência das análises meramente ideológicas do autor), credibilidade (garantia de qualidade relacionada à exatidão e quantidade das observações efetuadas), constância interna (independência dos dados em relação a acidentalidade, ocasionalidade etc.) e transferibilidade (possibilidade de estender as conclusões a outros contextos)” (Chizzoti: 1991, 90).

Uma rápida análise permite identificar uma certa correspondência entre estes critérios e aqueles das metodologias quantitativas: o de “transferibilidade” ao de “validade externa”, o de “constância interna” ao de “validade interna” e o de “credibilidade” ao de “confiabilidade” e “validade de medida”. Além disto, o conceito de “fiabilidade” - independência das análises meramente ideológicas - denota que, nas metodologias qualitativas, o descrédito na “neutralidade” não significa a apologia da tendenciosidade na produção do conhecimento.

Entretanto, os aspectos essenciais da validação nas metodologias qualitativas estão enquadrados no que DEMO classifica como critérios internos e critérios externos de cientificidade (Demo: 1981, 16-17).

Este autor define quatro “**critérios internos**” de cientificidade, que são (Demo: 1981, 16):

- a **objetivação**, que para ele “significa a tentativa de reproduzir a realidade assim como ela é, mais do que como gostaríamos que fosse”, e que, conforme vimos anteriormente, corresponde à noção de objetividade das correntes de pensamento compreensivas e marxistas, diferenciando-se do “objetivismo”, que corresponde à noção de objetividade do positivismo;

- a **coerência**, que é um critério “essencialmente formal”, referente à lógica interna do discurso, significando a “falta de contradição, argumentação estruturada, corpo não contraditório de enunciados, desdobramento do tema de forma estruturada, dedução lógica de conclusões”. Enfim, é um critério que não diz respeito ao conteúdo do discurso, mas a sua forma;

- a **consistência**, que “significa a capacidade de resistir a argumentações contrárias”, que se liga à “atualidade” da argumentação e à “capacidade de produzir argumentos sustentados pela comunidade científica”. Segundo Minayo, é a “própria qualidade argumentativa do discurso” (Minayo: 1992, 238);

- a **originalidade**, que significa que o conhecimento não é uma mera repetição do que já existe, mas que “representa uma real contribuição ao conhecimento”.

Os “**critérios externos**” são também quatro: a “intersubjetividade”, a “comparação crítica”, a “divulgação”, o “reconhecimento generalizado”. Para este autor, a intersubjetividade, definida como “a ingerência da opinião dominante dos cientistas de determinada época e lugar na demarcação científica”, seria o critério externo “propriamente dito”, enquanto os demais critérios seriam “facilmente dedutíveis ao da intersubjetividade” (Demo: 1981, 17).

MINAYO aponta outros autores que partilham desta opinião, pois colocam a “crítica intersubjetiva” ou “o julgamento intersubjetivo das descobertas” como o principal critério de validade, ou como o “procedimento fundamental de julgamento científico” (Minayo: 1992, 240-41). Para ela

“A submissão do produto do conhecimento à interface das discussões não significa que a verdade seja o resultado dos pontos de vista de vários estudiosos. Indica, entretanto, que a pluralidade de perspectivas permite lançar diferentes focos de luz dos conhecimentos a respeito do objeto em questão” (Minayo: 1992, 241).

Em resumo, poderíamos afirmar que os critérios de validação do conhecimento, em cada abordagem metodológica - quantitativa e qualitativa -, reproduzem a própria forma como o conhecimento é produzido nestas abordagens. Nas metodologias qualitativas, a validade está referida à aspectos subjetivos da produção do conhecimento, sendo a crítica intersubjetiva

sobre o produto científico o principal critério de validação. Nas metodologia quantitativas, os procedimentos de validação estão fundamentados na idéia de “objetividade” e “verificação”, enfatizando a modelagem do estudo, embasada no modelo experimental, e a questão da precisão das medidas.

1.3.1.3.3. Validade nas abordagens integradas

Para finalizar, caberia falar alguma coisa sobre a validação do conhecimento produzido através de uma abordagem de pesquisa que integre as duas metodologias. Vimos que a perspectiva integradora dos estudos de caso, dentro da abordagem de “pesquisa sintética”, está embasada em duas concepções fundamentais: a metodologia da triangulação, e a idéia de que a finalidade do estudo não é fazer inferências para populações, como nos estudos experimentais e quase-experimentais, mas para “teorias”. Logo, os critérios de validade destes estudos estão voltados para estas duas características.

Segundo MINAYO, do ponto de vista técnico, a própria triangulação é uma forma de fazer a “vigilância interna” da produção do conhecimento, funcionando como uma “prova eficiente de validação”. Além disto, o uso desta metodologia consagraria critérios externos de validade, como a crítica intersubjetiva e a comparação. Para relembrar este conceito, ouçamos a definição da autora:

“A ‘triangulação’ consiste na combinação e cruzamento de múltiplos pontos de vista através do trabalho conjunto de vários pesquisadores, de múltiplos informantes e múltiplas técnicas de coleta de dados” (Minayo: 1992, 241-42).

Ou seja, a abordagem da triangulação seria, em si mesma, uma forma de validação interna e externa do processo de produção do conhecimento.

Por outro lado, como a finalidade destes estudos é fazer inferências sobre teorias, e não sobre populações, a validade interna do estudo está fundamentada essencialmente na “qualidade, complexidade e exclusividade da articulação teórica sobre a qual se baseia o

estudo” e também “pela adequação entre o modo de análise escolhido e o modelo teórico a ser testado” (Contandriopoulos et al.: 1994, 40).

1.3.2. Questões metodológicas específicas: em busca de uma tipologia da avaliação em saúde.

“O fundamental não é o vocabulário de avaliação, nem as diversas metodologias, mas o compromisso do sistema de saúde em buscar, de forma permanente, aperfeiçoar sua contribuição à sociedade tanto no plano clínico como na perspectiva mais ampla da saúde pública. Tal busca precisa estar calcada não somente na luta política para a devida priorização da saúde e na análise das necessidades da população, como também na análise da atuação cotidiana do sistema hoje existente frente a essas necessidades” (Silver: 1992, 209).

Vimos nas seções anteriores que a avaliação, enquanto ciência social aplicada, põe em evidência as principais questões metodológicas das Ciências Sociais. Veremos agora alguns fatores específicos deste campo de estudos, que determinam a escolha da metodologia, e que referem-se essencialmente ao tipo de avaliação que se quer fazer, o qual por sua vez é determinado principalmente por três aspectos: o destinatário da avaliação, o objeto da avaliação e a finalidade da avaliação.

As pesquisas em avaliação geralmente tem um caráter aplicado, cumprindo um papel subsidiário ao processo administrativo (Suchman: 1967, 21, 29). A avaliação é útil “como meio para corrigir erros e poder modificar cursos de ação, ou seja, avaliar é “conhecer para melhorar as formas de agir”. Em outras palavras, “a avaliação só tem sentido na medida em que serve para tomar decisões concretas” (Aguilar&Ander-Egg: 1994, 26;32).

Isto coloca de imediato a questão do destinatário da avaliação como um fator condicionante do processo avaliativo. Veremos alguns aspectos referentes a esta questão ao discutir o contexto da pesquisa em avaliação.

O primeiro aspecto a destacar é que no campo da avaliação, a relação método-objeto coloca-se de forma muito direta. A avaliação “não pode existir no vácuo”, requer sempre que se pergunte: avaliação “de quê?” (Suchman: 1967, 38).

Todos os autores consultados concordam que a metodologia a ser utilizada depende do objeto da avaliação. É necessário utilizar “desenhos distintos de investigação” conforme a complexidade do objeto avaliado (Pitta: 1992, 45). Como vimos antes, as intervenções em saúde comportam vários níveis de complexidade, logo, “a depender do nível da realidade que é tomado como objeto do processo de avaliação, os problemas metodológicos e teóricos podem ser distintos” (Silva&Formigli: 1994, 81)

Da mesma forma que no campo geral das Ciências Sociais, há uma relação dialética, de “mútuo condicionamento” entre objeto da avaliação é o método de pesquisa escolhido para abordá-lo, de forma que não há “um caminho preestabelecido” (Bruyne et al.: 1982, 49, 66,69).

Ou seja, é a partir do objeto que vai ser avaliado que define-se o modelo de avaliação. SUCHMAN chega a colocar que a pesquisa em avaliação é uma “ferramenta, e como qualquer ferramenta, para ser mais efetiva, deve ser desenhada para uma função específica” (Suchman: 1967, 32).

Em relação ao primeiro objeto da avaliação em Saúde Coletiva - a avaliação dos estados de saúde - vimos anteriormente que é tratado dentro da questão do “diagnóstico” ou “avaliação das necessidades de saúde”, que historicamente tem ficado ao cargo dos estudos epidemiológicos. As principais questões referentes ao diagnóstico de saúde serão tratadas na seção correspondente ao sujeito da avaliação. De qualquer forma, diagnóstico é visto apenas como o passo inicial do processo avaliativo, através do qual se busca justificar a intervenção, esta sim objeto primordial da avaliação. Por esta razão, ao tratar da pesquisa em avaliação estaremos enfocando o estudo das intervenções em saúde.

No campo dos estudos em avaliação, a metodologia de abordagem do objeto - no sentido amplo de "referencial teórico que informa a maneira pela qual o autor vai descrever e buscar apreender o real" - é chamada geralmente de "abordagem" (“approach”) (Silva & Formigli: 1994, 83).

Há diferentes abordagens para fazer uma avaliação. Seja qual for a escolhida, o trabalho em avaliação deve partir de duas perguntas: para quê e para quem (Silver: 1992, 200; Weiss apud Cohen&Franco: 1994, 281).

O “para quem” remete ao problema do destinatário da avaliação, cujo papel é fundamental no processo avaliativo, visto a importância da questão dos valores neste processo. Além disto, a escolha da metodologia apropriada para a avaliação está também diretamente vinculada à “perspectiva do avaliador” (Silver: 1992, 208).

O “para quê” coloca a questão da finalidade ou dos objetivos da avaliação. Segundo SUCHMAN, a avaliação em si mesma é uma “meta” que pode ser alcançada de diferentes maneiras, de modo que os estudos em avaliação podem ser classificados conforme seu propósito (Suchman: 1967, 31).

CONTANDRIOPOULOS e cols. classificam os “objetivos oficiais” de uma avaliação em quatro tipos:

- “1. Ajudar ao planejamento e à elaboração de uma intervenção (objetivo estratégico).*
- 2. Fornecer informação para melhorar uma intervenção no seu decorrer (objetivo formativo).*
- 3. Determinar os efeitos de uma intervenção para decidir se ela deve ser mantida, transformada de forma importante ou interrompida (objetivo somativo).*
- 4. Contribuir ao progresso dos conhecimentos, à elaboração teórica (objetivo fundamental)” (Contandriopoulos et al.: 1992).*

Observa-se que os três primeiros objetivos são voltados para uma **intervenção**, denotando o caráter essencialmente aplicado da avaliação.

A distinção entre os objetivos formativo e somativo remete a dois tipos clássicos de avaliação: a **Avaliação Formativa** e a **Avaliação Somativa**. Estes conceitos são originários do campo da Educação, no contexto da avaliação de currículos. A “avaliação formativa” ocorreria enquanto a atividade a ser avaliada ainda estivesse em andamento, com a finalidade de melhorá-la, redirecioná-la. Em contraste, a “avaliação somativa” seria dirigida a um

produto final, buscando verificar a efetividade da intervenção, bem como o potencial deste “produto” em relação a futuras aplicações (Weilenmann: 1980, 15).

Dito de outra forma, a Avaliação Formativa refere-se “ao seguimento que se realiza durante o processo de execução de um programa ou projeto”, com o objetivo de fornecer uma retroalimentação (“feedback”) aos responsáveis pela intervenção que está sendo avaliada, podendo realizar-se durante o desenvolvimento da intervenção, nas fases de diagnóstico, formulação, implementação da execução da mesma. A Avaliação Somativa, por sua vez, refere-se aos “resultados ou efeitos” da intervenção, sendo realizada ao final da mesma, daí também ser denominada de “Avaliação Final”, cujos resultados servem para “determinar futuras ações” sobre a intervenção, no sentido de mantê-la, modificá-la ou suspendê-la (Aguilar&Ander-Egg: 1994, 43; Vergara: 1980, 18).

Quanto ao objetivo estratégico, entendemos que seja relativamente uma novidade na literatura, pois não vimos menção ao mesmo nos demais trabalhos consultados. Este objetivo remete a avaliação aos problemas que justificam a intervenção e aos objetivos que ela se propõe alcançar para resolvê-los. Ou seja, trata-se de uma avaliação que coloca sob questionamento a própria concepção da intervenção.

Enfim, podemos afirmar que os três primeiros objetivos da avaliação voltam-se para os três momentos de uma intervenção: **antes** de implementá-la, na fase de planejamento e elaboração (objetivo estratégico); **durante** a implementação da mesma (objetivo formativo) e **depois** de sua execução (objetivo somativo).

Os resultados da avaliação nestes três momentos propiciam conhecimentos que permitem alcançar o quarto objetivo fundamental da avaliação - “contribuir para o progresso do conhecimento e para a elaboração teórica”. Este objetivo coloca os estudos de avaliação em um papel de ciência básica. Embora a avaliação ainda hoje seja vista essencialmente como uma forma de pesquisa aplicada, vemos que este papel já vinha sendo atribuído desde algum tempo à pesquisa em avaliação

É considerando este objetivo que WEILENMANN entende a avaliação como “um meio para desenvolver, testar ou refinar teorias de mudança social”, somando-se ao seu propósito mais tradicional que diz respeito à determinação da efetividade de programas específicos (Weilenmann: 1989, 29).

Entretanto, no início da avaliação de programas sociais, havia uma negação deste papel, conforme vemos em SUCHMAN. Para ele, o objetivo principal da pesquisa em avaliação não seria “a produção de novos conhecimentos básicos, mas, ao invés disto, o estudo da efetividade da aplicação de tal conhecimento”. Ou ainda, a pesquisa avaliativa teria por finalidade “testar o grau no qual um conhecimento científico está sendo posto em prática com sucesso” (Suchman: 1967, 29; 8).

Seja como for, os objetivos da avaliação podem ser resumidos em dois: a) contribuir para o processo de decisão em relação à uma intervenção; b) contribuir para o progresso do conhecimento em relação à ação humana.

Em resumo, devido ao seu caráter de pesquisa aplicada, a escolha do método de pesquisa em avaliação está, acima de tudo, diretamente relacionada com o propósito ou objetivo da avaliação, bem como ao destinatário da avaliação e à natureza do objeto avaliado. Além disto, pela natureza subjetiva do processo avaliativo, a perspectiva do sujeito da avaliação é fundamental na escolha do tipo de avaliação a ser feita.

A seguir, veremos alguns aspectos referentes à tipologia da avaliação.

Ao tentar sistematizar as diferentes abordagens da avaliação, ficamos inicialmente muito confusos com o que estávamos lendo. Havia várias tipologias, que em parte se sobrepunham, em parte se sucediam, em diferentes formas de classificação. Ora, falava-se simplesmente de **avaliação**, classificada de diferentes maneiras: formativa/somativa, normativa ou administrativa/estratégica; interna ou externa; *ex ante*, durante ou *ex post*, entre outras maneiras. Ora, falava-se em avaliação como **abordagem** da estrutura, dos processos ou dos resultados. Ora, se fazia referência à **pesquisa** avaliativa, grafada com iniciais em letras minúsculas, ora Pesquisa Avaliativa grafada com iniciais em maiúsculas. Ora falava-se em **análise** de diversos tipos: estratégica, de eficiência, de eficácia, de produtividade, de efetividade, de eficiência, de implantação etc.

A classificação de AGUILAR&ANDER-EGG, que utiliza quatro critérios de interesse eminentemente prático, esclareceu a confusão que mencionamos acima (Aguilar&Ander-Egg: 1994, 41-51).

O primeiro critério é o “momento em que se avalia”, e que distingue três tipos de avaliação: “avaliação antes (ou *ex ante*), também chamada inicial ou predecisão”, onde se

avalia a “pertinência, viabilidade e eficácia potencial” da intervenção que está sendo proposta; a “avaliação **durante** a execução, também denominada avaliação da gestão, avaliação contínua, avaliação concorrente, monitoramento ou avaliação concomitante”; a “avaliação *expost*”, que se realiza quando a intervenção “alcançou seu pleno desenvolvimento”, e não necessariamente ao final da mesma.

O segundo critério refere-se às “funções que a avaliação deve cumprir”, e que distingue dois tipos principais - a avaliação formativa e a avaliação somativa - que vimos no início desta seção.

O terceiro critério refere-se à “procedência dos avaliadores”, e distingue quatro tipos de avaliação, que serão vistos na seção referente aos sujeitos da avaliação: a avaliação externa, a avaliação interna, a avaliação mista e a auto-avaliação.

O quarto critério refere-se aos “aspectos” ou “componentes” de um programa, objeto de avaliação, que segundo os autores é uma das classificações mais úteis, embora os avaliadores e teóricos não concordem totalmente sobre os tipos de avaliação definidos conforme este critério. É com base neste critério que discutiremos as diferentes tipologias de avaliação, na seção referente ao discurso integrado da avaliação. Para visualizar esta classificação, ver Anexo IX.

O importante é destacar que “não existe uma única metodologia de avaliação (e sobretudo, nenhuma técnica única simples e barata)” (Eicken apud Cohen&Franco: 1994, 281). Seja qual for o tipo de avaliação e o método escolhido para avaliar, o fundamental é respeitar as diretrizes e princípios que devem nortear o processo avaliativo: a objetividade, a validade, a confiabilidade ou fiabilidade, a praticidade e a oportunidade (Vergara: 1980, 20-21; Ander-Egg: 1984; 22-24;). Ou seja, o importante é que a avaliação proporcione “os resultados mais confiáveis e válidos dentro das restrições políticas e éticas e dos limites impostos pelo tempo e os recursos monetários e humanos” (OPS/OMS: 1985, 2).

Por fim, fomos distinguindo, aos poucos, duas linhas discursivas que predominavam no campo da saúde - a avaliação de programa e a avaliação de qualidade -, sendo a primeira oriunda do campo das Ciências Sociais e a segunda, no que concerne à avaliação em saúde, oriunda do olhar da Epidemiologia sobre a clínica médica.

Ao invés de entrar na discussão dos métodos e técnicas de avaliação, optamos, neste trabalho, por discutir estas duas grandes linhas discursivas que, em nosso entendimento, delimitam as abordagens da avaliação no campo da Saúde Coletiva. Pretendemos demonstrar, ao longo desta seção, que a linha que separa o discurso da avaliação em saúde é a tradicional diferenciação entre o nível individual e coletivo, que tem fundamentado a distinção entre a Saúde Pública da Clínica. Ao final, apresentamos uma proposta de abordagem metodológica que busca integrar estas duas linhas discursivas. Por último, discutiremos algumas dificuldades específicas da pesquisa em avaliação, derivados do contexto onde ocorre o processo avaliativo.

1.3.2.1. As grandes linhas discursivas da avaliação em saúde

1.3.2.1.1. A Avaliação de Programas

O entendimento da avaliação como uma forma de pesquisa científica está vinculado ao desenvolvimento da avaliação de programas. Como veremos no histórico, o grande impulso do campo mais geral da avaliação ocorreu nos EUA, na década de 1960, em virtude da necessidade de avaliar os programas sociais do Estado norte-americano, em expansão naquele período

Nesta época, surgiu a chamada “Pesquisa Avaliativa”, com a finalidade de avaliar os programas sociais. Ou seja, desde seu surgimento, a avaliação enquanto campo de pesquisas estava vinculada à avaliação de programas públicos, sendo voltada para a prestação de serviços por parte do Estado.

Nos primórdios da Pesquisa Avaliativa, SUCHMAN afirmava que a avaliação pressupunha “a existência de um programa ou atividade para ser avaliada”, o que a colocava na “área da **ação programática** ou atividade orientada para um objetivo [grifo nosso]”. Para ele, “se uma atividade qualquer tem um objetivo, então ela pode tornar-se objeto de uma pesquisa avaliativa” (Suchman: 1967, 28, 37).

Ou seja, existe uma relação “quase intrínseca” entre a avaliação, o planejamento e o desenvolvimento de programas (Suchman: 1967, 30). Como vimos no conceito de avaliação, a mesma imagem de circularidade que habitualmente é utilizada para descrever o processo de planejamento é também utilizada para descrever o processo de avaliação, gerando uma confusão muito freqüente entre a avaliação e o planejamento ou programação. Vimos também que esta confusão teve origem na própria história da avaliação, inicialmente considerada apenas a última etapa do planejamento e da programação, que acabou desprendendo-se e tornando-se, de certa forma, “uma função em si e para si mesma” (Weilenmann: 1980, 24).

WEILENMANN coloca que o rótulo de “Pesquisa Avaliativa”, grafado com letras maiúsculas, refere-se à “uma metodologia de alguma maneira formalizada, institucionalizada” em contraste com a tarefa intelectual concebida genericamente como avaliação. Entretanto, por ocasião de seu trabalho, ele colocava que a distinção entre as duas formas de avaliação não era rígida, pois a Pesquisa Avaliativa ainda estaria “se desenvolvendo” (Weilenmann: 1980, 15).

A diferenciação entre a avaliação, como um campo geral, e a Pesquisa Avaliativa como uma forma específica de avaliar programas, é colocada desta forma por SUCHMAN:

“Primeiramente, nós fazemos a distinção entre a avaliação como o processo social mais geral de fazer julgamentos de valor, sem levar em conta as bases de tais julgamentos, e a pesquisa avaliativa como referindo-se ao uso do método científico para coletar dados concernentes ao grau através do qual alguma atividade específica atinge algum efeito desejado” (Suchman In Caro: 1977, 49).

Ou seja, havia dois tipos de avaliação: uma que era pesquisa, a outra que não era, ambas englobadas sob a denominação geral de “processo avaliativo”.

Ainda hoje, AGUILAR&ANDER-EGG endossam esta distinção. Para eles, a avaliação tomada em sentido lato, denominada de “avaliação informal”, distingue-se da avaliação em sentido estrito, denominada de “avaliação sistemática”, não em relação aos seus propósitos, mas em relação aos seus métodos. Segundo estes autores, quando se fala em avaliação em sentido estrito está se fazendo referência à “pesquisa avaliativa”, que não possui métodos e técnicas próprias, mas “utiliza os mesmos métodos e técnicas da pesquisa social” (Aguilar&Ander-Egg: 1994, 23).

Além do uso do método científico, a idéia central da definição de Pesquisa Avaliativa, conforme vemos em SUCHMAN, é a medição do “grau” de alcance de um objetivo. Segundo o autor, “a característica que mais identifica a pesquisa em avaliação é a presença de alguma meta ou objetivo cuja medida do alcance constitui o principal foco do problema da pesquisa”. Por isto, “a natureza da avaliação dependerá largamente do tipo de medida que temos disponível para determinar o alcance de nossos objetivos”. (Suchman: 1967, 33-34; 37).

Ou seja, enquanto pesquisa, a avaliação envolveria “mais do que julgamento”, pois também abrangeria “a pesquisa sobre as condições que afetam o sucesso ou o fracasso” de um programa ou de um objetivo administrativo (Suchman In Caro: 1977, 48). Por outro lado, considerando que a finalidade da pesquisa avaliativa é “avaliar”, é preciso lembrar que o processo avaliativo não se esgota na “mensuração” do alcance dos objetivos, que pode ser feita cientificamente ou não. Esta mensuração apenas serve de base para a formação de um “juízo” sobre o valor da ação orientada para um objetivo (Suchman: 1967, 33-34; 37).

Em trabalho anterior, o mesmo autor afirmava que a utilização do termo “pesquisa” procurava enfatizar “aqueles procedimentos para a coleta e análise de dados que aumenta a possibilidade de **provar** ao invés de **afirmar** o valor de alguma atividade social” (Suchman: 1967, 8). Neste sentido, o autor colocava:

“Existem apenas dois caminhos através dos quais alguém pode se mover dentro da escala de objetivos de uma avaliação: (a) provando os pressupostos inter-relacionados através da pesquisa, isto é, transformando um pressuposto em um fato, ou (b) assumindo a validade dos pressupostos sem esta prova” (Suchman: 1967, 57).

WEILENMANN distingue duas correntes de “Pesquisa Avaliativa”, que se conformaram nos EUA, na década de 60, para avaliar os programas sociais do governo norte-americano:

- a primeira corrente seria a “pesquisa avaliativa clássica”, cujo enfoque é o **impacto**. Neste modelo, que ele denomina de “classical impact model”, “goal achievement”, “program goal attainment”, o esforço maior é dirigido à padronização das medidas para avaliar a **efetividade do programa**. É enfatizado o uso de métodos e técnicas “válidas” e “confiáveis”

para “medir” os efeitos dos programas (intervenções) sobre as metas ou objetivos estabelecidos. Este modelo está baseado nos desenhos de estudos experimentais, onde se compara *ex post facto* o grupo que é objeto da intervenção com outro grupo, similar ao primeiro, que não é submetido à mesma. A intervenção, em qualquer nível de abrangência, é considerada a variável independente, enquanto o efeito esperado é a variável dependente, entre as quais se busca estabelecer uma relação de causa-efeito. Este modelo apresentaria duas dificuldades: a principal delas é a freqüente (im)possibilidade de designar aleatoriamente a população que irá ou não ser objeto do programa; a segunda, é a dificuldade de aplicar os procedimentos estatísticos necessários para proteger o desenho de possíveis vieses que ameacem sua validade;

- a segunda corrente seria a pesquisa avaliativa “formativa”, cujo enfoque são os **processos**. Nesta corrente, denominada “avaliação formativa”, a finalidade da avaliação seria “guiar as atividades em andamento e prover as informações necessárias para mudar o foco ou redirecionar as metas do programa, em resposta ou em antecipação às mudanças no alvo do programa”, decorrentes de alterações no “ambiente”. Esta corrente teria surgido a partir da constatação de que alguns programas tinham um efeito pequeno, ou mesmo negativo, em relação ao esperado, o que originou duas ordens de interpretação por parte dos pesquisadores. Os pesquisadores da corrente “clássica” tinham a tendência de atribuir este fato aos erros ou vieses do desenho avaliativo; ou seja, o fracasso era atribuído à avaliação, e não ao programa. Porém, ao contrário dos primeiros, um grupo de pesquisadores passou a cogitar a possibilidade de que os programas não estivessem sendo adequadamente implementados, na condução das suas atividades. Daí a idéia de avaliar também os “processos” na implantação dos programas, e não apenas os resultados finais dos mesmos. Nesta corrente, “o rigor científico e a aderência estrita ao desenho experimental são considerados de menor importância, se não por outro motivo, por obstáculos políticos e práticos” (Weilenmann: 1980, 25-27).

Ele entendia que a aguda distinção entre as duas correntes estaria ficando menos precisa, com o surgimento, na época (1980), da chamada “avaliação compreensiva” que preconizava a combinação de ambas as correntes da Pesquisa Avaliativa - a que enfatizava o impacto e a que enfatizava os processos (p. 28).

A concepção da Pesquisa Avaliativa “clássica”, que enfoca o impacto, fica bem clara nesta colocação de SUCHMAN:

“Um projeto de avaliação pode ser visto como um estudo da mudança - o programa a ser avaliado constitui efeito estímulo ou “a causa” ou a variável independente, enquanto a mudança desejada é similar ao “efeito” ou a variável dependente. Caracterizado desta maneira, pode-se formular um projeto de avaliação em termos de uma série de hipóteses que afirmam que ‘Atividades A, B e C produzirão resultados X, Y e Z’. Note-se que esta formulação requer tanto a colocação dos resultados finais, ou objetivos de um programa, como a especificação do quê, no programa, se espera que produza estes resultados.”
(Suchman: 1967, 38)

Segundo o autor, esta analogia “estímulo-mudança” busca colocar o problema de “estabelecer conexões causais entre as atividades e os resultados”. Ou seja, a pesquisa avaliativa pode ser usada para “estabelecer uma **conexão causal** entre o que foi feito e os resultados que foram obtidos [grifo nosso]” (Suchman: 1967, 38, 66). Ao mesmo tempo, este autor também enfatizava a importância do enfoque nos “processos”, pois seria através do estudo dos mesmos que se explicariam o “como” e o “porquê” um programa funciona ou não funciona (Suchman: 1967, 66).

WEILENMANN entende que a corrente “clássica” tem por pressuposto que a situação alvo dos programas sociais são problemas bem estruturados (“well-structured problem situation”), quando na realidade, devido à complexidade do processo social, os problemas quase sempre são mal estruturados (“ill-structured”) (Weilenmann: 1980, 25, 17).

A noção de problema mal ou bem-estruturado foi bem desenvolvida por TESTA (1992: 93-94) e MATUS (1989: 233-42), como um dos pontos fundamentais do Planejamento Estratégico em Saúde.

Segundo Testa, um fenômeno qualquer pode ser caracterizado em função de dois eixos: simplicidade-complexidade e definido-indefinido. O eixo simplicidade-complexidade é definido pelo “número de variáveis que intervêm no problema”, enquanto o eixo definido-indefinido o é pela “precisão das relações que ligam essas variáveis”. O cruzamento destes

eixos define quatro tipos de problemas: 1) simples definido; 2) simples indefinido; 3) complexo definido; 4) complexo indefinido. Nos dois primeiros casos, é possível definir as variáveis que compõem o problema e fazer previsões do comportamento do mesmo com bastante segurança; no terceiro caso, a definição das variáveis é complexa, mas possível, permitindo também fazer previsões; no quarto caso, devido ao grande número de variáveis envolvidas no problema, nem sempre é possível enumerar todas, além do que, não é possível conhecer suas relações, logo, não é possível prever o comportamento do problema. Para o autor, o processo saúde-doença em uma população corresponde a este quarto caso. Ouçamos o que ele diz:

“Em termos do enfoque de sistemas, pode-se dizer que os modelos das ciências naturais aproximam-se dos casos 1, 2 e 3, que correspondem a sistemas relativamente fechados, enquanto que os das ciências sociais são do tipo 4, o que significa, além disto, que são tão abertos que é impossível definir seu contorno. Isto é, dissolvem-se as condições de contorno dos problemas sociais, de maneira que passam a fazer parte do problema de fundo” (Testa: 1992, 93-4).

Na formulação de MATUS, a primeira situação corresponde aos problemas bem-estruturados, a segunda e a terceira aos problemas quase-estruturados, e a quarta aos problemas mal-estruturados (Matus: 1989, 233-42). Para visualizar este esquema, ver Anexo X.

Ou seja, é muito remota a possibilidade de estabelecer relações de conexão causal quando se está voltado para problemas mal-estruturados, como é o caso da maioria dos programas sociais.

Vimos até aqui, que a concepção clássica da avaliação de programas implica em duas idéias básicas: primeiro, a medida do grau de alcance dos objetivos do programa; segundo, a busca para estabelecer uma relação causal entre o programa e os efeitos observados.

Segundo AGUILAR&ANDER-EGG, os estudos baseados nos objetivos, que foram os precursores dos trabalhos avaliativos, e os estudos baseados em experimentação, não são avaliações “verdadeiras”, mas “quase-avaliações”. No primeiro caso, porque “especificar até que ponto os objetivos propostos foram alcançados não permite emitir um juízo de valor sobre um programa, plano, atividade ou serviço” - os objetivos podem ser alcançados sem que o que se esteja fazendo “tenha algum valor e seja significativo”. No segundo caso, porque embora os

estudos experimentais permitam estabelecer relações causais, nem sempre incluem questões relacionadas com os juízos de valor, que é o aspecto fundamental da avaliação. Para eles, as “verdadeiras avaliações” seriam as avaliações “científicas” ou “pesquisa avaliativa propriamente dita”(Aguilar&Ander-Egg: 1994, 39-40).

Para COHEN&FRANCO, a concepção da avaliação “centrada na apreciação do grau em que foram alcançadas as metas é resultante da translação acrítica ao campo social das técnicas de avaliação de projetos econômicos”. No campo econômico, a simples escolha da tecnologia que vai ser utilizada já define “as etapas lógicas e inevitáveis” da implementação do projeto, enquanto que no campo social, um dos objetivos mais importantes da avaliação consiste justamente em “descobrir alternativas que otimizem o alcance dos objetivos procurados” (Cohen&Franco: 1994, 76).

Em literatura recente, vemos que de fato esta distinção entre a avaliação clássica, que enfatiza o impacto, e a avaliação de processos, já não existe mais, conforme previa Weilenmann. Em documento da OPS/OMS sobre avaliação de programas de saúde, vemos que a proposta busca afastar-se do “ponto de vista clássico segundo o qual a avaliação consistiria simplesmente em apreciar ou medir o grau de consecução de objetivos determinados ou a conformidade dos resultados de um programa com os critérios previamente estabelecidos” (OPS/OMS: 1985, 6). A proposta considera não só a avaliação do impacto, entendido como alcance de objetivos, mas também a avaliação dos processos, como vemos na seguinte colocação:

*“A avaliação não tem simplesmente por objeto determinar se foram ou não alcançados os objetivos de um programa, mas deve tratar de saber **como** funciona o programa e em que contexto, quais são os problemas ou questões que coloca, quais são os efeitos indesejáveis que produz e quais são os elementos que favorecem ou entorpecem seu êxito“ (OPS/OMS: 1985, 6).*

Até aqui, parece ter ficado bem demonstrado o surgimento da abordagem da “Pesquisa Avaliativa” dentro do campo mais geral da avaliação de programas sociais, entre os quais se incluem os programas da área da saúde.

Ainda hoje, a distinção entre a Pesquisa Avaliativa e a avaliação enquanto “processo social mais geral de fazer julgamentos de valor”, ou como “tarefa intelectual concebida genericamente”, ainda está presente na literatura sobre avaliação. Para AGUILAR&ANDER-EGG, a avaliação científica ou pesquisa avaliativa propriamente dita são as verdadeiras avaliações, enquanto as demais formas são classificadas como “pseudo-avaliações” ou “quase-avaliações”.

De uma forma geral, os objetivos da avaliação de programas podem ser assim resumidos (Ander Egg: 1984, 29-31; OPS/OMS: 1985, 8):

1. Medir o grau de consecução de objetivos determinados;
2. Estabelecer em que grau foram produzidas outras conseqüências imprevistas;
3. Medir o grau de pertinência, idoneidade, efetividade e eficiência de um projeto;
4. Emitir um juízo bem fundado sobre a natureza e a qualidade de um programa;
5. Facilitar o processo de tomada de decisões para melhorar e/ou modificar um programa ou projeto;
6. Obter lições da experiência e aproveitá-las para melhorar as atividades em curso e fomentar o planejamento mais eficaz mediante uma seleção judiciosa ente as possíveis ações ulteriores.

Em suma, a idéia central da avaliação de programas é “**aprender da experiência**, tanto exitosa ou fracassada, ajudando a racionalizar a tomada de decisões” (Cohen&Franco: 1994, 279). Por fim, vale a pena transcrever os objetivos da avaliação de programas, conforme vemos em documento da OPS:

“As avaliações são levadas a cabo por diversas razões: para julgar o valor dos programas em marcha e determinar a utilidade dos esforços por melhorá-los; para apreciar a utilidade de programas inovadores ou de iniciativas concretas; para incrementar a eficácia da administração ou da gestão de programas, e também para satisfazer as exigências fiscais que se impõe. Assim, as avaliações podem aportar conhecimentos substanciais e metodológicos sobre as ciências sociais” (OPS/OMS: 1985, 1).

1.3.2.1.2. A Avaliação de Qualidade

A segunda linha discursiva que distinguimos no campo da avaliação em saúde é a Avaliação de Qualidade. Observamos, na literatura, que alguns autores diferenciavam a “Avaliação de Programas” da “Avaliação de Qualidade”, bem como esta da “Garantia de Qualidade” (Brook&Lohr: 1985; Clemenhagen & Champagne: 1986; Silver: 1992). Além disto, havia referências ao uso, no campo da saúde, da tão propalada “Teoria da Qualidade Total” ou “Gestão de Qualidade Total” (Berwick: 1989; Nogueira: 1994).

Ao tentar caracterizar as diferenças entre estas abordagens da questão da qualidade em saúde, identificamos dois enfoques principais: a abordagem da qualidade voltada para o controle das **práticas individuais** de trabalho, especialmente a clínica médica, através das auditorias; e a abordagem da qualidade voltada para os **aspectos organizacionais** do trabalho coletivo, através do enfoque da gestão de qualidade. Em qualquer dos enfoques, a maior parte da literatura consultada refere-se à estudos feitos em ambiente hospitalar.

SILVER faz uma diferenciação entre a “avaliação de qualidade” e a “avaliação de programas”, ao discutir o conceito de cobertura. Esta autora vincula a primeira à noção de assistência curativa (“assistência à morbidade”) e a segunda à assistência preventiva, dirigida a “atividades específicas, com populações-alvo bem definidas, como por exemplo, a vacinação” ou dirigida ao atendimento de “doenças de prevalência bem estabelecida e tratamento padronizado (tuberculose, por exemplo)”. Para ela existiriam “poucos instrumentos que combinem adequadamente os aspectos clínicos com os aspectos organizacionais” (Silver: 1992, 208-209). Ora, esta distinção nada mais representa do que a tradicional divisão entre a “clínica” e a “Saúde Pública”, reforçando a idéia de que as atividades de “garantia de qualidade” estão mais voltadas para as práticas clínicas.

CLEMENHAGEN & CHAMPAGNE, por sua vez, distinguem as atividades de “garantia de qualidade” e as de “avaliação de programas”, que teriam se desenvolvido de “raízes diferentes e tendo enfatizando o uso de diferentes fontes de informação e o tomada de diferentes tipos de ação” (Clemenhagen & Champagne: 1986, 383). Vejamos a colocação dos autores:

“A garantia de qualidade tradicionalmente tem referido-se quase que exclusivamente à conduta dos médicos auditores, utilizando critérios objetivos para áreas clínicas específicas; a avaliação de programas refere-se a um conjunto de abordagens que podem ser utilizadas para fazer um julgamento de valor sobre um programa ou qualquer um de seus componentes, com vistas a modificar o programa” (Clemenhausen&Champagne: 1986, 383).

Ou seja, enquanto que as atividades de garantia de qualidade ficavam ao cargo das auditorias médicas, voltadas para o exercício das práticas clínicas, envolvendo a aplicação de critérios e normas específicas de cada área, as atividades de avaliação de programas eram vistas como uma das responsabilidades dos gerentes, voltada sobre o desempenho global dos serviços (de âmbito hospitalar). Entretanto, os autores consideram que estaria havendo, mais recentemente, uma tendência para eliminar esta separação, com o surgimento de propostas de integração entre as duas modalidades de avaliação, a partir da visão de que ambas são “sobrepostas e convergentes” e fazem parte da “função de controle” dos administradores. A partir deste ponto de vista, eles traçaram as linhas iniciais de um modelo teórico que propõe a integração destas duas abordagens, envolvendo os “aspectos clínicos e não clínicos” dos programas de saúde em um modelo de “garantia de qualidade de programas” (Clemenhausen & Champagne: 1986; 383-84). A formulação final deste modelo teórico integrado será vista mais adiante, como uma proposta de abordagem integrada da avaliação em saúde.

BROOK&LOHR estabeleceram a diferença entre os conceitos de “avaliação de qualidade” e “garantia de qualidade”, ambas referidas especificamente ao cuidado médico. Segundo eles, a qualidade do cuidado médico “é o componente da diferença entre a eficácia e a efetividade que pode ser atribuída aos provedores de cuidado”, e a avaliação de qualidade seria o “ato de detectar e medir esta diferença, incluindo as variações através das regiões e pessoas”. Como a diferença entre a eficácia - efeito máximo possível em condições ideais - e a efetividade - efeito em condições reais - seria uma decorrência dos aspectos técnicos e interpessoais do cuidado, a avaliação de qualidade, em termos práticos, seria a “medição” destes aspectos (Brook&Lohr: 1985; 711). Veremos, logo a seguir, que estes autores estão baseando-se no mesmo conceito de qualidade de Avedis Donabedian, que é essencialmente centrado no desempenho do profissional médico.

A “garantia de qualidade”, por sua vez, é definida como “o exercício formal e sistemático de identificar problemas na prestação de cuidados médicos, definir atividades para superar os problemas, e levar adiante os passos seguintes para garantir que nenhum problema novo tenha sido introduzido e que as ações corretivas tenham sido efetivas” (Brook&Lohr: 1985; 711).

Percebe-se que a noção de garantia de qualidade está mais próxima aos conceito de monitoramento (ou acompanhamento) e de controle de **processos**, na linha do que BERWICK denomina de “qualidade por inspeção” (Berwick: 1989; 53); enquanto que a noção de avaliação de qualidade, através da determinação de eficácia e efetividade, isto é, está voltada para os **resultados**.

Na linha da avaliação de qualidade em saúde, o quadro conceitual desenvolvido por Avedis Donabedian é considerado paradigmático. Segundo SILVA&FORMIGLI “o esforço mais bem-sucedido de classificação das abordagens possíveis para a avaliação em saúde foi desenvolvido por Donabedian, que concebeu a tríade ‘estrutura-processo-resultados’, a partir do referencial sistêmico” (Silva&Formigli: 1994, 83-84). Em nosso entendimento, esta “tríade” corresponde às noções da Teoria de Sistemas, qual seja, “input-process-output”. Quando tratarmos dos conceitos analítico-operacionais do campo da avaliação, veremos melhor o significado destes termos; por ora, destacamos apenas suas principais características.

A “estrutura” corresponderia ao conjunto de recursos utilizados na intervenção (físicos, materiais, humanos, financeiros e administrativos), os quais determinam “o potencial do sistema”. O “processo” corresponderia ao uso destes recursos na produção de atividades (ou serviços), e sua análise indica “como o sistema realmente funciona” (Silver: 1992, 202). Os “resultados” corresponderiam às mudanças ocorridas, em nível individual ou populacional, em decorrência da intervenção sobre o problema (Silver: 1992, 203; Pitta: 1992, 45; Silva&Formigli: 1994, 83).

Segundo PITTA, a abordagem da estrutura tem como pressuposto que “boas condições têm mais probabilidade de resultar em um processo adequado de cuidados e em um resultado favorável do que condições básicas precárias”. Na abordagem de processos, parte-se do pressuposto que “se todos fazem o que, à luz do conhecimento corrente, é o mais correto, é mais provável que o resultado seja melhor do que se o processo certo não for

seguido”. Por esta razão, nesta abordagem a noção de “**boa prática**” é um elemento fundamental, tendo encontrado sua “melhor aplicação” nas auditorias médicas. Entretanto, a autora aponta que esta concepção da abordagem de processos, restrita às auditorias, evoluiu para tornar-se “um exercício de resolução de problemas que tende a culminar com a identificação de problemas de qualidade”. Finalmente, a abordagem dos resultados teria como ponto essencial medir a realização dos objetivos propostos (Pitta: 1992, 46).

Embora entendendo que a abordagem de processos fosse “o caminho mais direto para o exame da qualidade do cuidado”, Donabedian recomendava que a melhor estratégia para a avaliação da qualidade seria “a seleção de um conjunto de indicadores representativos das três abordagens”, visto que “isoladamente cada uma dessas alternativas apresenta problemas e insuficiências” (Cf. Silva&Formigli: 1994, 83).

O fundamental a destacar na abordagem proposta por Donabedian é que ele desenvolveu seus trabalhos tendo como objeto principal de estudo o cuidado médico (“medical care”) ou o cuidado de saúde (“health care”) prestado individualmente, ou seja, referindo-se ao nível individual das intervenções em saúde e dentro da perspectiva de “garantia de qualidade” (“health care assurance”), ou seja, de “monitoramento do **desempenho clínico** com vistas a manter e melhorar sua Qualidade [grifo nosso]” (Donabedian apud Nogueira: 1994, 99; 101).

A ênfase no desempenho do médico como ponto central da qualidade fica muito evidente na distinção que DONABEDIAN et al. fazem entre “eficiência clínica”, referida à prática médica e “eficiência de produção”, referida ao conjunto dos processos de trabalho do hospital (taxa de ocupação, existência de pessoal treinado, desempenho do laboratório etc.). Vejamos como se posicionam os autores:

*“Enquanto a eficiência clínica é um componente fundamental da nossa definição de qualidade, nós preferimos **excluir** a eficiência de produção de nossa definição de qualidade porque não envolve o uso do julgamento clínico (Donabedian et al.: 1982, 985-86).*

Embora reconhecendo a importância da “eficiência de produção”, visto que é um “componente da qualidade do sistema”, os autores justificam excluí-la da definição de qualidade porque “não é um componente da qualidade do cuidado em si mesmo” (p. 986).

Segundo SILVA&FORMIGLI, inicialmente Donabedian manifestava “não estar claro se a qualidade era um atributo único, uma classe de atributos funcionalmente relacionados ou um conjunto heterogêneo de fenômenos reunidos pelo uso, razões administrativas ou preferências pessoais” (Silva&Formigli: 1994, 81).

Em suas primeiras formulações, ele desenvolveu o conceito de qualidade em torno do que considerava ser as três “dimensões” ou “componentes” do cuidado médico individual, quais sejam: (a) o conhecimento técnico-científico; (b) as relações inter-pessoais entre o médico (ou outro profissional) e o paciente; (c) as “amenidades”, isto é, as condições de conforto e estética das instalações e equipamentos no local onde ocorre a prestação de cuidados (Silva&Formigli: 1994, 81, Pitta: 1992, 45).

Nesta perspectiva, a qualidade do cuidado médico é definida como um “**produto**, medido como a expectativa de melhora no estado de saúde que pode ser atribuída ao cuidado [grifo nosso]”. Esta noção de qualidade envolve dois componentes básicos: primeiro, a definição e a medida do estado de saúde (inicial), bem como das mudanças ou diferenças neste estado (após o tratamento); segundo, a especificação do tipo de cuidado médico que pode estar associado às mudanças no estado de saúde, uma vez que estas as mudanças são o “produto” deste cuidado (Donabedian et al.: 1982, 975-76). Ou seja, é a mesma noção de causa-efeito que embasa a avaliação de programas, embora neste caso aplicada ao nível individual do cuidado.

Este raciocínio tem como pressuposto uma idéia de causa-efeito, em nível individual, da mesma forma como vimos anteriormente, na Pesquisa Avaliativa voltada para a avaliação de programas, em nível coletivo.

No mesmo artigo, DONABEDIAN e cols.. colocam que a escolha da “melhor estratégia de cuidado” deve levar em conta as três formas como se define a qualidade em saúde: (a) a “absolutista”, na qual a qualidade é definida exclusivamente em função das condições ideais de cuidado (o médico ideal, dispondo dos melhores recursos tecnológicos, podendo utilizá-los sem restrições de custo); (b) a “individualizada”, na qual a qualidade é definida em função das expectativas de melhora de cada paciente, bem como da sua disponibilidade financeira para custear o tratamento, o que obviamente varia de pessoa para pessoa; (c) a “social”, ou “valoração externa” da qualidade, na qual a mesma é definida

considerando-se o interesse de outras pessoas, além do paciente, como o médico, a família e amigos do paciente e a sociedade em geral (Donabedian et al.: 1982).

Os autores apontam que sempre existem diferenças, às vezes incompatibilidades, entre a definição individualizada e a definição social da qualidade, sendo que o principal problema do médico consiste em fazer um balanço entre as duas. A definição da qualidade do cuidado ficaria centrada na capacidade do médico em ponderar as expectativas individuais, de curto prazo, com os interesses sociais, de longo-prazo, cotejando-as com o padrão técnico-científico de qualidade (a qualidade ideal ou absolutista) (Donabedian et al.: 1982; 968).

Evidentemente que estas três definições - absolutista, individualizada e social - demonstram uma compreensão da natureza “contingencial” do conceito de qualidade, conforme os próprios autores apontam. Entretanto, todo o esforço é para definir, através de curvas estatísticas, o ponto ótimo do balanço entre custos, benefícios e riscos do cuidado à saúde. Ou seja, é um esforço para “objetivar” o conceito de qualidade em saúde, porém restringindo a abordagem a uma questão de eficiência. Os próprios autores reconhecem, ao tratar das diferenças do impacto no estado de saúde, entre o cuidado prestado pelo “médico ideal” e o prestado pelo “médico não ideal” (real), que estas noções são semelhantes às de “eficácia” e “efetividade” do cuidado (Donabedian et al.: 1982, 979; 981-82).

Segundo SILVA&FORMIGLI, Donabedian entendia que a ampliação do conceito de qualidade para o âmbito do social implicaria no risco da perda da “utilidade analítica” do mesmo, razão pela qual justificava aplicá-lo apenas ao cuidado médico. Para ele, a avaliação de programas somente se aproximaria da aferição da qualidade “quando o mesmo envolvesse a prestação de cuidado individual” (Silva&Formigli: 1994, 82). Talvez seja esta visão que tenha originado a tradicional distinção entre as atividades de garantia de qualidade e as de avaliação de programas na área da saúde, que apontamos anteriormente.

Posteriormente, DONABEDIAN ampliou o conceito de qualidade, desdobrando-o em dimensões que ele denominou de “sete pilares da qualidade” - eficácia, efetividade, eficiência, otimização, aceitabilidade, legitimidade e equidade (Donabedian: 1990). Ou seja, o autor passa a definir a qualidade do cuidado à saúde através dos conceitos tradicionais da avaliação de programas, embora entendendo-os como atributos do cuidado médico.

Em artigo de 1991 o autor retorna à formulação inicial, qual seja, a qualidade como fruto de componentes técnicos e interpessoais. Porém, ele passa a reforçar a subjetividade inerente ao conceito de qualidade, definindo-a como “um julgamento tanto sobre o componente técnico quanto sobre as relações interpessoais entre o cliente e o profissional, naquilo que estas características tem de adequação (*goodness*) ”, ressaltando novamente que nem sempre há concordância entre os julgamentos feitos do ponto de vista de cada um dos componentes (Silva&Formigli: 1994, 82; Pitta: 1992, 45).

A ampliação do conceito de qualidade através da inclusão de uma dimensão “social”, ainda que voltada para o cuidado médico individual e sob restrições de ordem analítica, na abordagem de Donabedian, permite deslocar a questão para outro patamar, voltado para as intervenções de saúde em nível coletivo. Isto abre perspectivas para tratar a questão da qualidade em níveis mais complexos de análise, como o desempenho global dos programas de saúde e até mesmo os sistemas de atenção e as políticas de saúde.

Nesta perspectiva, caberia passar para outro enfoque que vem ganhando corpo dentro do setor saúde, qual seja, a teoria da “Gestão de Qualidade Total” (GCT) ou “Controle de Qualidade Total” (CQT), em inglês “Total Quality Management” (TQM) ou “Total Quality Control” (TQC), respectivamente, a qual é voltada para a dimensão organizacional da qualidade, como o próprio nome indica. Esta abordagem, originalmente voltada para a gerência de indústrias, vem sendo aplicada também ao setor de serviços, tanto públicos como privados, como veremos a seguir. Daqui para diante, estaremos sempre nos referindo ao trabalho de NOGUEIRA, que discute as perspectivas desta abordagem no campo da saúde (Nogueira: 1994).

Em primeiro lugar, cabe destacar que o termo “total” não se refere a uma noção de qualidade absoluta, idealizada, mas à extensão deste processo de gerenciamento por todos os setores e níveis das organizações (p. 9).

Para compreender melhor as implicações desta teoria, o autor historia a evolução da questão da qualidade nas empresas, que teria se desenvolvido em quatro estágios, a saber: 1) a qualidade como pressuposto da excelência dos agentes; 2) o controle departamental da qualidade; 3) a garantia da qualidade e 4) a gestão ou controle da qualidade total.

No primeiro estágio, a qualidade era considerada um atributo de agentes (trabalhadores), ou seja, o pressuposto era que a “capacitação científica e tecnológica dos agentes, concomitantemente com processos informais de avaliação pelos pares é suficiente para definir a Qualidade dos bens e serviços que produzem”. Para o autor, embora os manuais de gerência de Qualidade geralmente omitam este estágio, “não há dúvidas de que esta é a forma corrente de avaliação de Qualidade em inúmeras áreas de produção de serviços” (p. 6-7).

No segundo estágio, este processo informal de avaliação de qualidade foi substituído pela criação de “departamentos” dentro das organizações, com a função específica de controlar a qualidade dos produtos (ou serviços), através do estabelecimento de “critérios e normas estritas de Qualidade” e do desenvolvimento de técnicas estatísticas de controle, as quais eram aplicadas sobre o produto final (p. 7).

No terceiro estágio, em função dos “grandes custos resultantes da detecção *a posteriori* da má-Qualidade”, foram desenvolvidos os programas de garantia de Qualidade para a “detecção preventiva de falhas ou defeitos nos produtos” (p. 99).

Segundo o autor, no campo da saúde o conceito de garantia de Qualidade (“health care assurance”) teria “conotações um pouco diferentes daquelas associadas às atividades industriais [pois] refere-se à elaboração de estratégias tanto para avaliação de Qualidade quanto para implementação de normas e padrões de conduta clínica através de programas locais ou nacionais” (p. 8). Nos EUA, as atividades de garantia de qualidade tornaram-se inclusive uma especialidade médica, tendo havido o estabelecimento de uma “ampla rede de especialistas e unidades de auditoria médica (*peer review*) dedicados ao monitoramento da atenção prestada por programas” de saúde, cujos altos custos vinham sendo criticados. Estes comitês de auditoria atuam “*post factum*, avaliando a adequação da conduta clínica dentro de uma amostra de tipos de casos atendidos” (pp. 8; 99).

Finalmente, o quarto estágio corresponderia à Gestão ou Controle de Qualidade Total, que se caracteriza como

“um duplo movimento de difusão de responsabilidades pela Qualidade para o conjunto de dirigentes e trabalhadores da unidade, simultaneamente com a adoção

de métodos participativos que permitam, aos distintos departamentos ou setores organizacionais, planejar, acompanhar e corrigir os rumos das ações pró-Qualidade bem como solucionar os problemas associados com o mau desempenho e a baixa produtividade”. (Nogueira: 1994, 8-9)

Ou seja, a garantia de qualidade deixa de ser responsabilidade de um departamento ou de especialistas no assunto, e passa a ser responsabilidade de todos. Isto quebraria dois pilares da administração taylorista:

- “a separação, em termos sociais e técnicos, entre as funções intelectuais de planejamento e as funções operacionais da produção direta” (p.10);

- o trabalho alienado, ao repor de certa forma “os elementos psicológicos que o capitalismo subtraiu aos trabalhadores” (p. 5).

A teoria de Gerência de Qualidade Total foi desenvolvida principalmente por dois autores - W. Edwards Deming e Joseph M. Juran - que tiveram carreiras profissionais semelhantes, tornando-se grandes consultores internacionais, especialmente para a indústria japonesa (p. 20), cujo sucesso no desempenho econômico é bastante conhecido.

Os aspectos teóricos essenciais da abordagem de Deming são “a centralidade do papel do consumidor [ou cliente]; e a preocupação em problematizar pedagogicamente o sentido e o uso da própria noção de Qualidade” (p. 20).

Neste segundo ponto, localizam-se as divergências entre Deming e Juran: para o primeiro, o conceito de qualidade é sempre relativo, por isso, o que importa é introduzir a discussão sobre o assunto dentro das organizações; para o segundo, é importante “buscar definir conceitos e métodos de Qualidade em abstrato”, estabelecendo-se assim as “possibilidades de mensuração” da mesma. Ou seja, enquanto Deming assume a subjetividade do conceito de qualidade, Juran busca a “cientificidade” através da mensuração (p. 21, 23).

NOGUEIRA adota as proposições de Deming, que ficam mais claras em algumas conclusões que elabora, ao analisar a questão da Qualidades:

? “o juiz da Qualidade é primariamente o consumidor, tendo os técnicos e especialistas uma função coadjuvante;

- ? *não há qualquer possibilidade de avaliar por supostos critérios ‘científicos’ a Qualidade de um produto quando se elide à subjetividade dos consumidores;*
- ? *o mundo de subjetividade em que se movem os problemas da Qualidade tem sua contrapartida objetiva nos métodos e tecnologias usados para estender e garantir a permanência de certas propriedades úteis de um produto (“garantia de Qualidade”);” (Nogueira: 1994, 32).*

Ora, a centralização do juízo de qualidade no “consumidor” (ou cliente) significa uma inversão da base da qual partiu Donabedian, que colocava o médico como juiz da qualidade do cuidado à saúde. Esta postura, evidentemente, corresponde ao primeiro estágio da discussão da qualidade, que vimos anteriormente, qual seja, a qualidade pressuposta como um atributo do agente produtor (no caso, o médico prestador de serviços). Portanto, a abordagem da Gestão de Qualidade Total contraria frontalmente o “axioma” herdado da medicina flexneriana, segundo o qual “o preparo científico do corpo de profissionais e o atributo *high-tech* dos equipamentos usados” garantiriam *per si* a qualidade do cuidado à saúde (p. 17).

BERWICK aponta que, apesar de ter cinquenta anos de existência (desde 1930), as técnicas de qualidade total tem sido pouco aplicadas aos sistemas de saúde. Ele atribui este fato, em parte, à “barreiras culturais” colocadas pelos médicos, que parecem ter dificuldades em ver-se como “participantes do processo”, percebendo-se, ao invés disto, como os “agentes isolados do sucesso ou fracasso” da atenção à saúde (Berwick: 1989, 55). Este é também o principal problema identificado por NOGUEIRA na adoção da GQT no campo da saúde. Para ele, os médicos tendem a atuar como “agentes organizacionais autônomos”, em uma “prática solo”, cujos interesses subordinam a estrutura geral dos serviços de saúde (Nogueira: 1994, 86; 94).

A conduta autônoma dos médicos dentro das organizações de saúde, justificada na visão flexneriana de que uma sólida formação científica garante uma boa prática, levou à imposição de uma “forma de ajuizar a Qualidade em saúde” essencialmente corporativa, através das auditorias médicas. Segundo NOGUEIRA, a principal crítica a este modelo de controle de qualidade é justamente que ele “está fundado num esquema de monopólio do saber que reforça e amplia o poder do médico”. Além disto, implica em altos custos e pouco retorno em termos de melhora efetiva da Qualidade (Nogueira: 1994, 97; 100).

A Gestão de Qualidade Total desloca a abordagem de qualidade do âmbito corporativo das auditorias para o âmbito mais geral da gerência dos serviços e programas, onde se alcança um “estatuto inédito de *accountability*”, isto é, da “necessidade de prestar contas a alguém”³⁰, no qual a categoria médica perde em parte sua autonomia como “juiz da qualidade”, passando a subordinar-se à “controles e normas de conduta estatuídos fora da profissão” (Nogueira: 95-98).

Além disto, o trabalho das auditorias médicas está na linha da garantia de qualidade, ou seja, a busca da qualidade através da “abordagem disciplinar” ou “inspeção” do desempenho das pessoas, baseado no esquema tradicional da indústria, que segundo BERWICK é “na melhor das hipóteses, ineficiente, e na pior, uma fórmula para o fracasso”, pois gera um medo nas pessoas, levando-as a jogar um papel defensivo em relação à avaliação, perdendo-se a “chance de aprender” com os resultados. O que ele denomina de “Teoria do Melhoramento Contínuo” seria um “modo novo, mais convincente e mais válido para enfocar a questão da qualidade” (Berwick: 1989, 53-54).

Ele destaca que o pressuposto desta teoria é de que a maior parte dos problemas de qualidade, e conseqüentemente as oportunidades para melhorá-la, são “construídos diretamente dentro do complexo processo de produção”, e somente em poucas ocasiões os defeitos, isto é, os problemas de qualidade poderiam ser atribuídos “à ausência de vontade, de habilidade ou de boas intenções” por parte das pessoas envolvidas no processo. Por isto, a melhoria de qualidade deveria ser buscada na “compreensão e revisão do processo de produção com base nos dados do processo em si mesmo”. O enfoque passa a ser “no produtor médio, e não no desviante, e na aprendizagem, e não na defesa”, ou seja, mais na cooperação e menos na censura àqueles que apresentam alguma deficiência. Pelo contrário, neste enfoque, “todo o defeito é um tesouro”, porque é na descoberta da imperfeição que está a chance de melhorar a qualidade (Berwick: 1989, 53-54).

Em resumo, no campo da saúde, o principal referencial teórico para a avaliação de qualidade é o modelo de Donabedian, concebido à partir da avaliação das práticas da clínica médica. Neste sentido, entendemos que a linha da “Avaliação de Qualidade” desenvolveu-se no campo da saúde como uma questão vinculada à prestação de cuidados em nível individual,

³⁰Ver adiante, no histórico da avaliação em Ciências Sociais, a discussão

enquanto a “Avaliação de Programas” voltava-se para as intervenções em nível coletivo. Em função de estar voltada para a avaliação de práticas clínicas, geralmente no âmbito hospitalar, a avaliação de qualidade ficava restrita à atuação das auditorias médicas.

Entretanto, fora do âmbito específico da saúde, a questão da qualidade vinha sendo tratada como uma questão mais ampla de gerência das organizações, especialmente no setor industrial. Considerando a história da questão da qualidade na indústria, a auditoria médica corresponderia ao terceiro estágio do desenvolvimento da avaliação de qualidade, qual seja, o dos programas de garantia de qualidade.

Em nosso entendimento, o enfoque da qualidade em saúde como atributo quase exclusivo da prática médica pode ter retardado a adoção da perspectiva da qualidade como uma questão global dos sistemas de atenção à saúde, reforçando uma cultura que entende a qualidade como um produto do desempenho individual do médico e da disponibilidade de equipamentos de alta tecnologia. O enfoque atual da “Gestão de Qualidade Total” desloca radicalmente a questão da qualidade em dois sentidos: primeiro, que o juízo da qualidade da atenção à saúde deixa de ser exclusivamente dos médicos, e passa a ser fundamentalmente um papel do usuário dos serviços de saúde, sendo, portanto, mais adequado à visão da saúde como um direito de cidadania; segundo, que a avaliação de qualidade deixa de ser uma questão de estudos baseados na análise estatística de padrões de desempenho individual, e passa a ser uma questão global da gestão, onde o papel do avaliador está mais vinculado à capacidade de negociação do gerente dos serviços e programas.

1.3.2.1.3. Um discurso integrado

Vimos até aqui, que as atividades de garantia de qualidade e as de avaliação de programas desenvolveram-se de “raízes diferentes”, demandando diferentes tipos de informação e ação, e caracterizando duas linhas discursivas no campo da saúde (Clemenhagen&Champagne: 1986, 383).

sobre o significado do conceito de “accountability”.

Entretanto, CLEMENHAGEN&CHAMPAGNE detectaram uma tendência para integrar estas abordagens, o que levou-os a formular uma proposta de “estrutura conceitual para a avaliação de programas em cuidados de saúde”, que implica em uma concepção abrangente de programa, englobando tanto as atividades dirigidas para o nível individual - aspectos clínicos -, como para o nível agregado ou coletivo - aspectos não clínicos - da prestação de cuidados em saúde (Clemenhagen&Champagne: 1986).

A partir de sua experiência em administração hospitalar, estes autores procuraram demonstrar que existem relações de sobreposição e convergência entre os enfoques da “avaliação de qualidade” e da “avaliação de programas”, tradicionalmente separados. Eles adotaram a perspectiva de que a avaliação é uma das “funções de controle” do processo de administração, ou seja, é uma das “responsabilidades de controle” do gerente, que pode exercê-la de duas maneiras:

- aplicando “critérios” de avaliação a cada um dos **componentes** do programa (Problemas, Objetivos, Recursos, Serviços ou Atividades, Efeitos), que geralmente é identificada com os métodos de “avaliação de qualidade” ou de “garantia de qualidade”;

- avaliando as **relações entre os componentes** do programa, visando responder como cada um destes componentes deriva do outro, que tende a ser identificada com a “avaliação de programa” e necessita do instrumental da Pesquisa Avaliativa (Clemenhagen & Champagne: 1986, 383-84).

Estas duas abordagens são entendidas como uma questão de “foco” da avaliação: ou a avaliação enfoca cada um dos componente do programa, ou enfoca as relações entre estes componentes. Conforme o foco da avaliação - nos componentes ou nas relações entre eles - corresponde um conjunto de análises a serem feitas.

Quando o foco da avaliação são os componentes do programa, temos o seguinte: se é dirigido para os objetivos do programa, procede-se uma “apreciação das necessidades” (“needs assessment”); se é dirigido para os recursos do programa, procede-se uma “avaliação da estrutura”; se é dirigido para os serviços prestados através do programa, procede-se uma “avaliação dos processos”; e finalmente, se o foco é dirigido para os efeitos do programa, procede-se uma “avaliação dos resultados” (“evaluation of outcome”), tudo isto mediante a aplicação de critérios e normas (Clemenhagen & Champagne: 1986, 385).

Quando o foco da avaliação são as relações entre os componentes do programa, temos os seguintes tipos de análise: a “análise estratégica”, entre os problemas que justificam o programa e os objetivos estabelecidos; a “análise do programa” ou “revisão do desenho do programa”, entre os objetivos e os recursos disponíveis; a “análise de produtividade”, entre os recursos utilizados e os serviços produzidos; a “análise dos efeitos”, entre os serviços produzidos e os efeitos do programa, que desdobra-se em “análise de eficácia” e “análise de efetividade”; e finalmente a “análise de eficiência”, entre os recursos utilizados e os efeitos observados (Clemenhagen & Champagne: 1986, 385). A proposta destes autores fica mais clara no esquema que está no Anexo XI.

No primeiro enfoque, predomina o tradicional referencial de avaliação de qualidade - estrutura, processos, resultados. A novidade fica por conta de que os autores propõe a aplicação de critérios e normas aos objetivos do programa, o que coloca sob foco da avaliação a própria finalidade do mesmo. No segundo enfoque, predomina o referencial da avaliação de programas - eficácia, efetividade e eficiência do programa.

A avaliação baseada na aplicação de critérios aos componentes dos programas poderia ser realizada pelos próprios gerentes do mesmos, visto que está fundamentada em padrões de desempenho (“standards”) já definidos (Clemenhagen & Champagne: 1986, 384).

Porém, quando o foco da avaliação são as relações entre os componentes do programa, seria necessário “mais do que o tempo e a experiência do administrador permitem”, pois, mais do que a simples aplicação de critérios e normas, seria necessário determinar “como um componente do programa deriva do outro”. Este tipo de análise requer uma “pesquisa científica”, para assegurar que “são válidas as conclusões alcançadas acerca das relações entre os componentes do programa” (p. 384). Embora os autores não coloquem explicitamente, estão obviamente sugerindo que esta tarefa demandaria o concurso de um avaliador, munido do instrumental da Pesquisa Avaliativa.

Ao unir a abordagem da garantia de qualidade com a abordagem mais ampla da avaliação de programas, este modelo une a avaliação da prática clínica - nível individual - e a tradicional avaliação de programas de Saúde Pública - nível coletivo ou agregado, superando a tradicional separação entre as duas. Além disto, esta integração desloca a questão da qualidade em saúde do âmbito de controle corporativo das práticas clínicas, através das auditorias

médicas, para o âmbito da gerência dos serviços e programas de saúde. Para facilitar a compreensão de como o eixo individual/coletivo é superado nesta abordagem integrada, ver Anexo XII.

É justamente neste ponto que o modelo converge para a estratégia fundamental que especialistas na “Teoria de Qualidade Total” propõe para o campo da saúde. Um dos relatórios do Comitê para pesquisar a aplicação desta teoria ao cuidado de saúde, publicado em 1992 nos EUA³¹, ressalta que “os conceitos de **totalidade** e **integração** emergem aqui como referências decisivas na abordagem da Qualidade do trabalho em saúde”. O principal aspecto é justamente a necessidade de fazer um “esforço de integração entre vários tipos de serviços”, que o relatório divide em duas categorias: os chamados “serviços operacionais” e os “serviços clínicos” (ou cuidado clínico) (Nogueira: 1994, 84-5).

Acreditamos que ao adotar a perspectiva da avaliação de qualidade e de programas como “função gerencial de controle”, os autores estejam provavelmente refletindo uma influência da teoria de “Gestão de Qualidade Total”. Esta suspeita origina-se não somente do tratamento “gerencial” dado à questão da qualidade, mas também de um detalhe que nos chamou atenção no roteiro de avaliação sugerido pelos autores, onde eles mencionam, como critério de avaliação da estrutura, a existência de “círculos de qualidade”, termo que refere-se a uma das estratégias típicas daquele tipo de gestão (Clemenhagen & Champagne: 1986, 386).

Tomando por base o modelo teórico proposto por CLEMENHAGEN&CHAMPAGNE, em 1986, que acabamos de ver, CONTANDRIOPOULOS e colaboradores estabeleceram, em 1992, uma nova tipologia para as abordagens da avaliação no campo da saúde.

Estes últimos autores demonstram que a intersecção de dois campos gerais - o da pesquisa e o da avaliação - delimitam três áreas de conhecimento e de práticas. A primeira refere-se à intersecção destes dois campos, isto é, à área em que a pesquisa e a avaliação se sobrepõem, que caracteriza a Pesquisa Avaliativa. A segunda, refere-se ao que, dentro do campo geral da pesquisa, está fora desta intersecção, que é a pesquisa não-avaliativa. A terceira refere-se ao que, no campo geral da avaliação, está fora da intersecção, e denomina-se

³¹GOAL/QPC. Putting the T in Health Care TQM, A Model for Integrated TQM: Clinical Care and Operations. A GOAL/QPC Health Care Application Research Committee Report, Methuen, 1992, p.7.

“Avaliação Normativa” (Contandriopoulos et al.: 1992). O esquema dos autores está no Anexo XIII.

Ou seja, olhando-se do território da avaliação, a Pesquisa Avaliativa é um tipo de avaliação; olhando do território da pesquisa, o termo “avaliativa” torna-se “um adjetivo que especifica um tipo de pesquisa”, conforme afirmava SUCHMAN anos antes, ao referir-se à Pesquisa Avaliativa dentro do campo geral das Ciências Sociais (Suchman: 1967, 7).

Vejamos então as duas abordagens da avaliação em saúde - a Avaliação Normativa e a Pesquisa Avaliativa - conforme caracterizam CONTANDRIOPOULOS e colaboradores (1992).

A Avaliação Normativa é definida da seguinte forma:

“A avaliação normativa é a atividade que consiste em fazer um julgamento sobre uma intervenção, comparando os recursos empregados e sua organização (estrutura), os serviços ou bens produzidos (processo), e os resultados obtidos, com critérios e normas” (Contandriopoulos et al.: 1992).

A Avaliação Normativa busca “estudar cada um dos componentes da intervenção em relação a normas e critérios” (Contandriopoulos et al.: 1992). Ou seja, nesta abordagem, compara-se a estrutura, os processos e os resultados com determinadas normas ou padrões (daí o termo “normativa”), identificando-se as diferenças entre o que foi planejado - o desejado, a norma - com o que de fato está acontecendo na intervenção.

Este tipo de avaliação é “normalmente feita por aqueles que são responsáveis pelo funcionamento e pela gestão da intervenção”, fazendo parte da atividade “natural” de gerência. Estas atividades corresponderiam “às funções de controle e de acompanhamento, assim como aos programas de garantia de qualidade”. Por esta razão, é também denominada de “avaliação administrativa” (Contandriopoulos et al.: 1992).

A aplicação de critérios e normas a cada elemento da intervenção - estrutura (recursos), processos (serviços) e resultados (efeitos) - é denominada de “apreciação” (“appraisal”), ao invés de avaliação (“evaluation”), termo utilizado na proposta anterior, de Clemenhagen&Champagne (1986).

Na apreciação da estrutura, busca-se essencialmente verificar se os recursos (humanos e materiais) são suficiente e adequados (em termos de quantidade e qualidade), sendo este o tipo de avaliação no qual se apoiam os organismos de **acreditação**. A apreciação dos processos decompõe-se em três dimensões: técnica, relações interpessoais e organizacional; sendo objeto dos programas de garantia de qualidade. Quanto à apreciação dos resultados, os autores entendem que o enfoque da Avaliação Normativa é “muitas vezes insuficiente para se fazer um julgamento válido sobre os resultados de uma intervenção”, sugerindo que seja empregada a abordagem da Pesquisa Avaliativa.

É evidente que o que a “apreciação dos processos” corresponde à linha da “avaliação de qualidade”, apoiada no referencial de Donabedian, porém dentro de uma visão administrativa/programática mais global.

A abordagem da **Pesquisa Avaliativa**, por sua vez, é definida da seguinte maneira:

“Podemos definir a pesquisa avaliativa como o procedimento que consiste em fazer um julgamento ex-post de uma intervenção, usando métodos científicos. Mais precisamente, trata-se de analisar a pertinência, os fundamentos teóricos, a produtividade, os efeitos e o rendimento de uma intervenção, assim como as relações existentes entre a intervenção e contexto no qual ela se situa, geralmente com o objetivo de ajudar na tomada de decisões” (Contandriopoulos et al.: 1992).

Ou seja, a Pesquisa Avaliativa busca “examinar por um procedimento científico as relações entre os diferentes componentes de uma intervenção” (Contandriopoulos et al.: 1992). Há muita semelhança entre esta definição de “Pesquisa Avaliativa” e a que vimos anteriormente de Suchman (1967) e Weilenmann (1980), na linha da avaliação de programas.

Ao mesmo tempo, em trabalho posterior, CONTANDRIOPOULOS e cols..retomam a idéia central de Battista et al. (1989), em sua proposta de estrutura integrada para a pesquisa em saúde, qual seja, a posição intermediária da pesquisa avaliativa dentro dos estudos em saúde, colocando que “a principal característica desta pesquisa consiste no fato dela ser o eixo entre a pesquisa sobre os estados de saúde e a pesquisa sobre as intervenções em saúde” (Contandriopoulos et al.: 1994, 21).

O produto final deste desenvolvimento teórico sobre a Pesquisa Avaliativa em saúde, oferece a possibilidade de seis (6) enfoques analíticos, que apresentamos a seguir (Contandriopoulos et al.: 1992, 1994; Clemenhagen&Champagne: 1986):

1. **Análise Estratégica** (ou Revisão da Racionalidade): relaciona a situação problema que deu origem à intervenção com a própria intervenção, buscando analisar a adequação estratégica da mesma, ou seja, a **pertinência** da intervenção. A pergunta ser respondida é: “Se justifica a intervenção em relação aos problemas da população e é pertinente que seja este interventor que a assuma?”. Para respondê-la, busca-se verificar “a forma como a situação problemática foi identificada”, isto é, os próprios métodos de apreciação da necessidades; “o grau de prioridade do problema escolhido em relação aos demais problemas identificados”; a “pertinência da intervenção escolhida em relação a todas as intervenções possíveis”. A análise estratégica poderia ser feita a partir da análises de necessidades, métodos de determinação de prioridades, análises de mercado, análise de carteira etc.

2. **Análise da Intervenção** (ou Análise do Programa ou Revisão do Desenho): relaciona os objetivos da intervenção e os meios empregados para atingi-los, buscando analisar a “capacidade dos recursos que foram mobilizados e dos serviços que foram produzidos para atingir os objetivos definidos”. A pergunta é: “A intervenção se propõe a alcançar os objetivos de maneira satisfatória?”. Para responder, é necessário verificar “se a teoria na qual a intervenção foi construída é adequada e, por outro lado, se os recursos e as atividades são suficientes em quantidade e qualidade”. Em suma, busca-se saber da **suficiência** da intervenção proposta. Os métodos para fazer este tipo de análise seriam “extrapolações dos métodos desenvolvidos para apreciar a qualidade de um instrumento de medida”, ou seja, são métodos que permitem interrogar sobre a validade e a confiabilidade de uma intervenção, através da apreciação da qualidade de seu modelo teórico, isto é, da veracidade e generalidade deste modelo.

3. **Análise da Produtividade**: relaciona os serviços produzidos com os recursos utilizados, buscando estudar “o modo como os recursos são usados para produzir serviços”. A

pergunta é: “Os recursos são colocados de maneira que otimizem os serviços produzidos?”. Esta pergunta desdobra-se em duas outras questões: “Poderíamos produzir mais serviços com os mesmos recursos? Poderíamos produzir a mesma quantidade de serviços com menos recursos?” A produtividade poderia ser medida de duas formas: em unidades físicas (produtividade física) ou em unidades monetárias (produtividade econômica). Os autores chamam atenção para a dificuldade da área da saúde em “definir e medir sua produção”, por isto, sugerem que a análise de produtividade seja decomposta em diferentes níveis de resultado dos serviços ou atividades (resultados primários, intermediários ou finais). Os métodos para análise de produtividade seriam derivados dos métodos econômicos e dos métodos de contabilidade analítica.

4. **Análise dos Efeitos** (Análise de Eficácia e Análise de Efetividade): relaciona os efeitos com os serviços produzidos, buscando “determinar a eficácia dos serviços para modificar os estados de saúde”. A pergunta é: “Quais os efeitos que podem ser atribuídos à intervenção?”. Os autores definem quatro tipos de eficácia, dependendo do contexto no qual se realiza a intervenção, e que determinam diferentes tipos de pesquisa para avaliá-las: **eficácia teórica** (contexto de laboratório), onde se pode realizar uma pesquisa experimental; **eficácia dos ensaios** (contexto dos ensaios clínicos randomizados), que permite avaliar os “efeitos puros” da intervenção sobre o homem; **eficácia de utilização** (contexto natural das “práticas normais”), voltada para o grupo de usuários que se beneficiou da intervenção; e **eficácia populacional** (contexto comunitário-populacional), onde se considera não apenas a população beneficiada, mas “toda a população a quem a intervenção era destinada”, cuja análise deve considerar o grau de cobertura, a aceitabilidade e a acessibilidade da intervenção na sua população alvo. Além disto, nesta análise, deve-se considerar tanto os efeitos desejados quanto não desejados, os efeitos de curto, médio e longo prazo, bem como os efeitos diretos, sobre a população alvo, e os indiretos, sobre outras populações. As eficácias de utilização e populacional são geralmente denominadas de “efetividade”. Os métodos para fazer a análise de eficácia podem ser quantitativos e qualitativos, utilizando-se modelos da pesquisa experimental ou da pesquisa sintética (ver adiante, na parte referente aos modelos de pesquisa).

5. **Análise do Rendimento** (ou Análise de Eficiência): consiste em relacionar “a análise dos recursos empregados com os efeitos obtidos”, combinando duas das análises anteriores (de produtividade econômica e dos efeitos). A pergunta é: “Com que custo são obtidos os efeitos da intervenção?”. Este tipo de análise é desdobrado em três sub-tipos: custo/benefício, custo/eficácia e custo/utilidade. Nas análises de **custo/benefício**, o numerador e o denominador são expressos em termos monetários, o que torna este tipo de análise “muito difícil de se conduzir na área da saúde, particularmente por causa da dificuldade que se tem de dar um valor monetário para variações nos estados de saúde”. Por esta razão, as outras duas análises são mais frequentes. Na análise de **custo/eficácia**, os custos são expressos em termos monetários, porém o denominador é expresso “por índices reais de resultado” ou de eficácia, como anos de vida ganhos, índice de satisfação, redução da dor etc. Na análise de **custo/utilidade**, da mesma forma, o denominador é expresso “pela utilidade que traz a intervenção para aqueles aos quais ela se destina”, como os indicadores de qualidade QALY (Quality Adjusted Life Years). Os autores colocam que o conceito de “utilidade” para o usuário permite comparar intervenções com diferentes objetivos (isto é, com diferentes propostas de resultados), o que não se pode fazer nas análises de custo/eficácia (que comparam intervenções com os mesmos objetivos). Entretanto, haveria sérios problemas metodológicos para determinar a utilidade associada a diferentes tipos de efeitos, bem como para medir as vantagens de uma intervenção e definir os critérios de comparabilidade do rendimento de diferentes opções de intervenção.

6. **Análise da Implantação**: este tipo de análise consiste “por um lado, em medir a influência que pode ter a variação no **grau de implantação** de uma intervenção nos seus efeitos e, por outro lado, em apreciar a influência do ambiente, do **contexto**, no qual a intervenção está implantada nos efeitos da intervenção [grifos nossos]”. No primeiro caso, considera-se que às vezes “uma ausência de efeitos pode ser devida ao fato de que a intervenção avaliada não foi realmente implantada”; no segundo caso, procura-se avaliar os sinergismos e antagonismos existentes entre o contexto e a intervenção. Este tipo de análise é pertinente especialmente quando se observa “uma grande variabilidade nos resultados obtidos por intervenções semelhantes”, porém implantadas em contextos diferentes, e também nos

casos em que a intervenção é “complexa e composta de elementos seqüenciais sobre os quais o contexto pode interagir de diferentes modos”. A pergunta é: “Como e por quê os efeitos variam entre os meios nos quais a intervenção é introduzida?”. Os métodos apropriados para fazer este tipo de análise são os estudos de caso.

No primeiro modelo proposto por Clemenhagen&Champagne, em 1986, não mencionava a análise de implantação, que entretanto já havia sido considerada na proposta de Battista e cols., em 1989. Segundo estes últimos, a análise de implementação permite verificar as relações entre os componentes de programas complexos, revelando a “dinâmica interna das intervenções” e permitindo determinar “qual a seqüência de eventos que produz o efeito desejado ao custo mais baixo” (Battista et al: 1989, 1159).

Evidentemente, há uma ordem hierárquica nas análises da Pesquisa Avaliativa, que correspondem à seqüência temporal, em etapas ou fases, dos componentes do programa. Assim, a rigor, a análise dos efeitos de um programa somente poderia ser realizada depois que as análises correspondentes à fase dos pré-requisitos e da implantação do programa já estivessem feitas. Ou seja, seria preciso conhecer, em primeiro lugar, se existe adequação entre os objetivos propostos para a intervenção e os problemas sobre os quais se quer intervir (análise estratégica); em segundo lugar, se existe adequação entre os objetivos e os recursos disponíveis (análise da intervenção); em terceiro lugar, qual a relação entre os recursos e os serviços produzidos (análise de produtividade). Somente após o conhecimento dos resultados destas análises é que se poderia pretender analisar os efeitos de um programa (análise dos efeitos), isto é, verificar a relação entre os serviços produzidos e os efeitos alcançados, em termos de eficácia e efetividade.

A análise da relação entre os recursos utilizados e os efeitos alcançados (análise da eficiência), seria a penúltima etapa desta seqüência de análises. A última buscaria relacionar o desempenho global do programa, verificado nas análises anteriores, com o contexto de sua implantação (análise de implantação). Isto permitiria avaliar porque um mesmo modelo de programa apresenta diferentes efeitos em contextos diferenciados. Para visualizar as diferentes análises da Pesquisa Avaliativa, ver o Anexo XVI.

HARTZ&CAMACHO destacam a importância da análise de implantação, porque colocam em evidência os processos internos dos programas, permitindo a melhoria de qualidade dos mesmos. Os autores chama atenção que, por influência da Epidemiologia, geralmente há uma ênfase nas “análises de impacto”. Em função disto

“...omitem-se os problemas metodológicos e operacionais implicados na monitoração dos serviços e análise dos efeitos intermediários que deveriam ser abordados previamente para se conhecer o grau de implantação destas intervenções, condição indispensável para a validade do próprio julgamento sobre o Impacto observado” (Hartz&Camacho: 1996).

Além disto, considera-se que a eficácia das intervenções coletivas já deveria “ter sido demonstrada quando da implementação do programa”, de forma que o foco principal da avaliação deve ser a própria “capacidade de implementar a intervenção” (Hartz&Camacho: 1996).

De uma certa forma, portanto, a análise de implantação traz para dentro da Pesquisa Avaliativa as preocupações típicas da abordagem da Avaliação Normativa, voltando-se para os processos que ocorrem durante a implantação/execução dos programas, ao mesmo tempo que busca estabelecer uma relação de causalidade entre os eventos que ocorrem nesta fase com os resultados finais do programa, apontando também a variabilidade destes resultados em função do contexto.

Enquanto a Avaliação Normativa pode ser conduzida pelos próprios gerentes do programa, a Pesquisa Avaliativa, como já vimos, requer uma perícia metodológica e teórica que dificilmente poderia ser feita pelos responsáveis pela intervenção, ficando geralmente ao cargo de avaliadores externos à mesma. Em termos estritamente metodológicos, a aplicação de critérios e normas previamente estabelecidos aos componentes do programa possibilitaria estudos descritivos, enquanto que a análise das relações entre estes componentes ensejaria estudos analíticos. Talvez por isto, na Avaliação Normativa os termos utilizados são simplesmente “apreciação” ou “avaliação” do comportamento dos componentes do programa, enquanto na Pesquisa Avaliativa o termo utilizada é “análise” das relações entre estes componentes, o que possibilita conhecer como um deriva do outro e permite estabelecer um

raciocínio de causa-efeito entre o funcionamento interno da intervenção e seus efeitos em um contexto específico.

Vimos até aqui, que as abordagens da avaliação em saúde tiveram duas raízes principais, que caracterizaram duas linhas discursivas e metodológicas - a avaliação de programas, voltada para as intervenções coletivas, e a avaliação de qualidade ou garantia de qualidade, voltada para o cuidado à saúde prestado individualmente (prática clínica). A partir do trabalho de Clemenhagen&Champagne, de 1986, que considera ambas abordagens como uma das responsabilidades dos gerentes, Contandriopoulos e cols. desenvolveram, em 1992, um modelo integrado de avaliação em saúde, através do entrecruzamento da tradicional abordagem de “avaliação de qualidade” - caracterizada como a aplicação de critérios e normas aos componentes do programa - e a abordagem de “avaliação de programas” - caracterizada como a análise das relações entre estes componentes -, renomeando-as como “Avaliação Normativa” e “Pesquisa Avaliativa”. Além disto, o modelo incorpora a idéia de Battista e cols., de que a pesquisa avaliativa seria o eixo dos estudos no campo da saúde, relacionando a pesquisa sobre os estados de saúde e a pesquisa sobre as intervenções de saúde.

A novidade da proposta final parece ficar por conta de dois aspectos: a denominação do campo da Avaliação Normativa, como área complementar à Pesquisa Avaliativa, dentro do campo maior da avaliação em saúde; e, fundamentalmente, a incorporação da noção de “contexto” como parte intrínseca da avaliação das intervenções em saúde, através da análise de implantação.

A relação entre as abordagens de avaliação de qualidade/avaliação de programas e de Avaliação Normativa/Pesquisa Avaliativa pode ser apresentada da seguinte maneira:

- a Avaliação Normativa (ou administrativa) é identificada com as funções de acompanhamento e controle, e consiste na aplicação de normas e critérios técnicos aos componentes de um programa - estrutura, processos e resultados (avaliação de qualidade ou garantia de qualidade), durante a execução do mesmo, visando acompanhar e, se necessário, reorientar sua execução (avaliação formativa), podendo ser realizada pelos próprios responsáveis pelo programa;

- a Pesquisa Avaliativa consiste essencialmente na aplicação de procedimentos científicos para analisar as relações entre os componentes de um programa, enfocando

principalmente o impacto final do mesmo e buscando estabelecer relações de causa-efeito entre a intervenção e seus resultados (avaliação de programas), visando emitir um juízo final sobre a mesma (avaliação somativa), requerendo, geralmente, o trabalho de um profissional de avaliação, tanto pelo rigor dos procedimentos quanto pelo tipo de análise necessária, além do que implica em um maior distanciamento entre o processo de avaliação e os responsáveis pelo programa, o que em princípio, levaria a um juízo menos parcial.

Historicamente, localizamos esta distinção nos primórdios do desenvolvimento da própria Pesquisa Avaliativa aplicada aos programas sociais, nos EUA, com o surgimento da “corrente clássica”, que enfatizava o impacto, e que evidentemente corresponde a uma avaliação com finalidade somativa, e a corrente denominada de “avaliação formativa”, que enfatizava os processos e obviamente tinha finalidade formativa (Weillenman: 1980).

Neste ponto, caberia lembrar as colocações de ANDER EGG, que vimos no conceito de avaliação, onde destaca os dois modos como se pode considerar processo avaliativo: como um processo permanente de controle operacional e como a última etapa ou fase de uma intervenção (Ander Egg: 1984, 24). É na primeira forma que a avaliação se confunde com as noções de “controle” e de “acompanhamento”, “seguimento” ou “monitoramento”, e evidentemente corresponde à Avaliação Normativa, enquanto que é na segunda forma que a avaliação caracteriza-se como “estudo”, “pesquisa” ou “investigação”, correspondendo à Pesquisa Avaliativa.

Estes dois modos de ver a avaliação estão orientados por duas racionalidade, conforme apontam COHEN&FRANCO: a “ação racional funcional”, que refere-se ao **desenvolvimento** da ação, e a “ação racional substancial”, que refere-se ao **resultado** da ação (Cohen&Franco: 1994, 65), correspondendo, novamente, à Avaliação Normativa e à Pesquisa Avaliativa, respectivamente.

Estas duas racionalidades de ação dão origem a uma outra classificação da avaliação, formulada por AGUILAR&ANDER-EGG, que corresponde as duas formas básicas que vimos até aqui: a **avaliação formal**, que se volta para os componentes internos do programa, buscando responder a pergunta: “O que fazemos corresponde ao programado?”; e a **avaliação substantiva**, voltada para a pertinência do programa em relação aos problemas ou necessidades da população destinatária, que busca responder a pergunta: “As mudanças e

modificações que apareceram na situação-problema são resultantes (total ou parcialmente) do programa ou projeto?” (Aguilar&Ander-Egg: 1994, 37, 49). A relação entre estas racionalidades e as dois tipos básicos de avaliação pode ser vista no Anexo XIV.

Para finalizar, tentaremos dar uma visão geral das abordagens que vimos até agora, buscando relacioná-las com os objetivos da avaliação, que vimos na seção inicial, os aspectos avaliados nas intervenções em saúde, as fases ou momentos dos programas, o enfoque da avaliação e os tipos de análise da Avaliação Normativa e da Pesquisa Avaliativa.

Em primeiro lugar, vemos que cada um dos três primeiros objetivos da avaliação corresponde a um aspecto que é avaliado na intervenção: quando o objetivo é estratégico, avalia-se a **necessidade** ou pertinência da intervenção; quando o objetivo é formativo, avalia-se a **adequação** ou suficiência da intervenção; quando o objetivo é somativo, avalia-se a **efetividade** da intervenção, que inclui a eficácia³². Cada um destes aspectos correspondem à diferentes fases ou momentos do programa, com seus respectivos componentes: a fase de pré-requisitos inclui os problemas e os objetivos; a fase de implementação inclui os recursos e os serviços ou atividade; a fase dos efeitos, obviamente, inclui os efeitos do programa.

Em segundo lugar, a escolha da abordagem vai depender do enfoque da avaliação - se nos componentes do programa - Avaliação Normativa - ou nas relações entre estes componentes - Pesquisa Avaliativa, sendo que a primeira implica em procedimentos normativos/administrativos e a segunda em procedimentos de maior rigor científico. A escolha do caráter da avaliação - administrativo ou pesquisa - define o tipos de avaliação ou análise.

No caso da Pesquisa Avaliativa, estas análises visam responder questões específicas, que implicam em diferentes tipos de estudo adequados para respondê-las.

Em suma, a escolha da melhor abordagem metodológica para fazer uma avaliação obedece a uma “hierarquia” de opções, poderia ser resumida da seguinte maneira: uma vez definida a finalidade da avaliação (estratégica, formativa ou somativa), define-se o tipo de abordagem (avaliação normativa ou pesquisa avaliativa), do qual deriva o foco da avaliação (nos componentes do programa ou nas relações entre os componentes) e o tipo de análise (no primeiro caso, análise das necessidades, da estrutura, dos processos ou dos resultados; no

³²A sinonímia entre “necessidade” e “pertinência” e entre “adequação” e “suficiência” também fica por nossa conta. Estes três aspectos gerais da

segundo caso, análise estratégica, do desenho ou da intervenção, da produtividade, dos efeitos, do rendimento ou eficiência, da implantação). Estas relações podem ser visualizadas nos Anexos XV, XVI e XVII.

Para finalizar, identificamos alguns “eixos” que facilitam a compreensão das diferentes classificações de avaliação que vimos até aqui, entendendo que os autores estabeleceram diferentes cruzamentos destes eixos para definir suas tipologias:

- o eixo finalístico/teleológico, que enfatiza a finalidade ou o porquê da avaliação: fins estratégicos, fins formativos, fins somativos;

- o eixo funcional, que enfatiza a utilização da avaliação: função administrativa, isto é, subsídio ao planejamento, à gerência, à tomada de decisões (avaliação normativa) ou função científica, isto é, produzir conhecimento sobre as intervenções (pesquisa avaliativa);

- o eixo temporal, que enfatiza o momento em que é realizada a avaliação: durante a intervenção (avaliação de processo ou normativa) ou após sua finalização (avaliação clássica de resultado ou impacto);

- o eixo estrutural, que enfatiza a avaliação dos componentes da intervenção (avaliação normativa) ou as relações entre estes componentes (pesquisa avaliativa);

- o eixo individual/coletivo, que enfatiza o nível de agregação ou abrangência da intervenção: aspectos da prática clínica ou do cuidado individual prestado pelos profissionais de saúde em geral (avaliação de qualidade) ou aspectos organizacionais (avaliação de programas).

1.3.2.2. Algumas dificuldades específicas: o contexto do processo avaliativo

Vimos nas seções anteriores, que a avaliação expõe de forma crucial a problemática que permeia a investigação em Ciências Sociais: a busca da objetividade na produção do

avaliação são por demais conhecidos. Apenas para mencionar uma fonte, ver: ABRASCO: 1995b, 30; AGUILAR&ANDER-EGG: 1994, 62.

conhecimento sobre a ação do homem, a disputa entre as metodologias qualitativas e quantitativas, a validade do conhecimento produzido nos estudos avaliativos. Além disso, por ter essencialmente caráter aplicado, os estudos em avaliação enfrentam dificuldades adicionais: as decorrentes do próprio contexto da avaliação, e as decorrentes da relação entre o avaliador e o destinatário da avaliação, ou seja, o “para quem” é feita a avaliação.

A maior parte das dificuldades encontradas por aqueles que pretendem fazer uma avaliação decorre do fato de que o processo de avaliação situa-se “num contexto ideológico, organizacional e sóciopolítico definido”. Em consequência, o papel do avaliador “nunca pode ser neutro”, pelo contrário, existe sempre a possibilidade de que o processo avaliativo venha a gerar conflitos entre os interesses dos avaliadores e dos demais atores sociais dentro de uma organização, especialmente dos níveis políticos de decisão (Aguilar&Ander-Egg: 1994, 79).

O papel do avaliador é emitir um juízo bem fundado, “utilizando critérios racionais e tratando de compreender o que ocorre com o programa”. ou qualquer forma de intervenção (OPS/OMS: 1985, 6). Daí que uma avaliação conduzida com seriedade e bem feita pode evidenciar problemas de “má implementação e execução” da intervenção, ou mesmo vir a questionar seus próprios objetivos, apresentando resultados que podem ter consequências negativas sobre a intenção, sempre presente nos gestores/administradores, de obter maiores recursos (Aguilar&Ander-Egg: 1994, 78-79).

Por estas razões, todo processo avaliativo pode “ameaçar alguma parcela daqueles que estão relacionados com o projeto avaliado”, transformando-se em algo “potencialmente perigoso”, daí o motivo pelo qual “a tarefa avaliativa não costuma ser tão freqüente quanto devia ser” (Cohen&Franco: 1994, 281; Aguilar&Ander-Egg: 1994, 78).

O contexto político dentro do qual desenvolvem-se as avaliações implica, pois, em uma diversidade de problemas que dificultam a operacionalização do processo avaliativo, aos quais somam-se problemas de caráter estritamente metodológico, típicos da investigação em Ciências Sociais. Dentre as dificuldades apontadas pelos autores, selecionamos algumas, de caráter mais geral, que nos pareceram mais importantes: a questão da “avaliabilidade” dos programas; a tensão entre as demandas administrativas e as exigências de rigor científico na avaliação; a confusão entre o fracasso do programa e o fracasso da avaliação.

Segundo WEILENMANN, o primeiro problema com que se depara o avaliador é a falta de clareza dos objetivos e da definição das metas, que dificulta a “avaliabilidade” (“evaluability”) dos programas. A própria condição de “avaliabilidade” dos programas seria o principal problema com o qual se embate a literatura em Pesquisa Avaliativa. Para ele, vêm de longa data as queixas dos avaliadores de que os objetivos dos programas não são claros, ou são formulados em termos muito vagos, o que dificulta extremamente a tarefa de avaliar. Esta questão teria inclusive originado uma demanda para incluir profissionais em avaliação nas etapas preliminares do planejamento e do desenho dos programas (Weilenmann: 1980, 29).

AGUILAR&ANDER-EGG mencionam “a característica às vezes vaga e vaporosa de certos programas”, especialmente na área do bem-estar social (Aguilar&Ander-Egg: 1994, 78). Por esta razão, não seria uma tarefa fácil para o avaliador encontrar “a definição exata dos objetivos de um programa”, bem como identificar os responsáveis pelas atividades para alcançá-los (Suchman: 1967, 38).

Há um certo consenso de que o setor sanitário “não se presta facilmente à medição dos resultados em função de objetivos numéricos previamente determinados”. Por isto, freqüentemente o avaliador se vê compelido a “emitir um juízo qualitativo apoiado, sempre que seja possível, por dados seguros e numéricos” (OPS/OMS: 1985, 14). Embora entenda-se que os objetivos dos programas devam ser definidos em termos “relativamente exatos”, eles geralmente são definidos por termos “carregados de valores e com magnitude imprecisa”, necessitando então serem traduzidos em termos “programáticos e quantitativos” para permitir a avaliação (Knox et al.: 1986, 12). É preciso considerar, entretanto, que não é tarefa fácil “especificar objetivos realistas num ambiente de mudança constante”, o que leva muitas vezes a que os objetivos estabelecidos sejam “irreais ou francamente errados ou freqüentemente subjetivos, sendo uma resposta a alguma demanda imposta externamente” (Silver: 1992, 205). Ou seja, são objetivos “para inglês ver”....

Além disto, HARTZ&CAMACHO apontam dificuldades na própria concepção teórica e na implantação dos programas, que dificultam sua avaliação:

“O desconhecimento da estrutura teórica da rede de causalidade de grande parte dos programas já em funcionamento e o reconhecimento da existência de

adaptações contextuais, em programas cujo grau de implantação é ignorado, tornam a avaliação extremamente complexa” (Hartz&Camacho: 1996).

Em função desta realidade, os estudos avaliativos deveriam concentrar-se na análise do “modelo teórico” ou “constructo” da intervenção” (Análise das Intervenções), bem como na análise do grau de implantação dos programas (Análise da Implantação). Um programa concebido a partir de um modelo teórico da realidade falho ou equivocado evidentemente não poderá alcançar os efeitos esperados; da mesma forma, um programa bem concebido teoricamente, porém com problemas de ordem prática na sua implementação, também não poderá alcançar os resultados desejados (Hartz&Camacho: 1996).

Neste mesmo sentido, WEILENMANN coloca que “o componente teórico da avaliação propõe que o problema social atacado pelo programa deveria estar bem fundamentado em uma teoria” (Weilenmann: 1980, 29). Ele lembra que os problemas na elaboração e implementação dos programas na área social são muito complexos, devido à natureza essencialmente mal-estruturada dos problemas sociais.

O modelo clássico de avaliação parte do pressuposto de que os problemas sociais comportam-se em uma relação “uni-direcional de causa e efeito”, por isto espera-se que objetivos dos programas e os critérios de avaliação possam ser “descobertos e definidos sem ambigüidade”. Entretanto, os fenômenos sociais são essencialmente mal estruturados, isto é, comportam-se como uma “rede”, onde há relações de mútua causalidade. Isto dificulta muito a definição e a escolha de objetivos de ação (Weilenmann: 1980, 13, 29-30).

Em função disso, no modelo clássico, o problema da “avaliabilidade” dos programas passa a ser encarado como um “problema técnico-administrativo”, no qual a “vaguidade dos objetivos do programa e a ambigüidade das expectativas são deploradas e atribuídas à gerência do programa e à situação política”. Porém, o que se espera do avaliador é que possa estabelecer as “condições de avaliabilidade” do programa, re-estruturando, se for possível, a situação problema mal-estruturada, porém sem reduzir, indevidamente, a sua complexidade. Além do que é preciso reconhecer que nem tudo é passível de ser objeto de uma Pesquisa Avaliativa (Weilenmann: 1980, 29; 30).

A dificuldade de definir os objetivos dos programas não está vinculada somente à complexidade dos fenômenos sociais, mas também à dificuldade de definir o tipo de mudança

desejada com a ação programática. WEILENMANN coloca que se os projetos pretendem afetar a transformação social, surge a questão: “que mudança e que tipo de transformação social?”. Para ele, há “muitas teorias e diversos princípios” sobre o quê deveria consistir esta mudança, sendo este o problema do desenvolvimento de políticas e do planejamento de projetos, bem como da avaliação dos projetos e seu impacto (Weilenmann: 1980, 12). Da mesma forma, SUCHMAN coloca:

“Em relação ao tipo de mudança, é preciso especificar a mudança do que para o que. É preciso que se seja capaz de determinar o estado existente das coisas antes que a atividade programática seja iniciada, e então especificar em que consiste a mudança desejada” (Suchman: 1967, 39).

Para este autor, a resposta a muitas das questões da avaliação requer o exame do “pressupostos subjacentes aos objetivos estabelecidos”, o que seria esperado, visto que a atitude crítica, o “desafio ao *status quo*” é inerente à idéia de avaliação (Suchman: 1967, 41). O papel da avaliação, neste aspecto, seria forçar a estrutura administrativa a “especificar os valores que a administração está visando, ao menos até o ponto no qual algumas questões visíveis de pesquisa possam ser especificadas (Gruenberg apud Suchman: 1967, 41). Com isto, o administrador seria forçado a “definir mais aguçadamente sua própria imagem do que está tentando fazer”(p. 41).

A idéia de que o processo avaliativo deve alcançar o questionamento dos valores subjacentes aos objetivos não é, entretanto, partilhada pela maioria dos autores consultados. ANDER-EGG colocam que a avaliação deve verificar o que foi feito, e não “as intenções ou propósitos” (Aguilar-Egg: 1984, 19). COHEN&FRANCO lembram que esta questão é essencialmente de caráter político, implicando em conflitos de valores cuja solução extrapola em muito o campo de ação do avaliador. Vale transcrever a posição destes autores:

“Em primeiro lugar, convém lembrar que o estabelecimento dos objetivos não corresponde ao âmbito da avaliação; é anterior a ela, está circunscrito à competência dos tomadores de decisões e são, na realidade, um pré-requisito para a avaliação” (Cohen&Franco: 1994, 280).

Porém, admitem que o avaliador teria o papel de mostrar a “incoerência” ou “inadequação” dos objetivos, visando contribuir para a melhoria do projeto (Cohen&Franco: 1994, 280). Para isto, seria necessário que ele fosse capaz de planejar a avaliação de forma que ficasse “sob medida” às características do programa avaliado, tanto no sentido de apontar suas qualidades como seus defeitos (Eicken apud Cohen&Franco: 1994, 280).

A segunda questão que destacamos são as frequentes disparidades entre os requisitos da Pesquisa Avaliativa e as demandas concretas de avaliação. WEILENMANN define o problema como uma tensão entre “a busca de precisão e validade científica”, de um lado, e os “constrangimentos das demandas políticas e práticas”, de outro lado (Weilenmann: 1980, 28). COHEN&FRANCO mencionam a existência de uma “forte tendência a julgamentos intuitivos”, por parte dos administradores e legisladores, e, por outro lado, “um gosto exagerado pelos níveis puramente científicos por parte dos acadêmicos” (Cain e Hollister apud Cohen&Franco: 1994, 280). COTTRELL JR. coloca a questão em termos de “balanço entre o desenho rigoroso e o método e as realidades situacionais” onde o avaliador deve trabalhar (Suchman: 1967, viii, Prefácio de L.S. Cottrell Jr.).

Ou seja, como lembra SUCHMAN, o principal problema do avaliador não é fazer o desenho da avaliação, mas aderir aos princípios deste desenho diante das questões administrativas, entre as quais ele destaca (Suchman: 1967, 21):

- os objetivos do programa, que são as “hipóteses de estudo” para o avaliador, são definidos pelo administrador, que dificilmente aceita que sejam questionadas as suposições ou pressupostos que embasam estes objetivos;

- os estudos avaliativos interferem nas atividades cotidianas dos programas ou serviços, o que cria dificuldades para sua aceitação;

- a avaliação geralmente requer coleta de dados que não são necessários para a operação do programa, mas que são necessários para a avaliação, o que gera uma sobrecarga para o pessoal da linha operacional.

Segundo autor, o administrador percebe a avaliação como um processo “simples e prático”, enquanto o avaliador a percebe como “complexa e especializada”, gerando um conflito que deve ser solucionada pelo balanceamento entre estas duas exigências (Suchman: 1967, 21-22, 35).

O problema da tensão entre as expectativas do administrador e as do avaliador tem como corolário a questão da pertinência e da utilização dos resultados da avaliação.

Para COHEN&FRANCO, a “exagerada preocupação” que os avaliadores têm pela pesquisa científica vem em prejuízo da “produção de resultados úteis para a tomada de decisões”, especialmente quando os avaliadores provêm das universidades, pois tendem a dar preferência às “orientações próprias de seu âmbito de inserção cotidiano”, prejudicando a instauração da avaliação como “uma atividade necessária” (Cohen&Franco: 1994, 280). WEILENMANN aponta que o apego aos critérios científicos de avaliação pode levar à “sub-utilização” dos seus resultados, pois são os critérios administrativos que determinam a utilização dos resultados de uma avaliação, sendo este um dos sérios problemas da Pesquisa Avaliativa (Weilenmann: 1980, 25).

Como a finalidade principal da avaliação é contribuir para o processo decisório, é necessário que o processo avaliativo possa compatibilizar as expectativas do avaliador e do administrador, superando a inevitável tensão entre os dois, para que os resultados da avaliação sejam, de fato, aproveitados. É preciso considerar que “o gestor é o contraponto do avaliador, porque ele é o usuário da avaliação”³³.

Entretanto, isto não deve significar um subordinação ou atrelamento acrítico do avaliador às demandas da administração. AGUILAR&ANDER-EGG lembram que em certas situações ocorrem “alianças” entre o avaliador e o político, nas quais o trabalho avaliativo visa apenas legitimar decisões políticas já tomadas, que ele denomina de “pseudo-avaliações” (Aguilar&Ander-Egg: 1994, 78). Ou, como colocam COHEN&FRANCO, são avaliações que visam apenas “outorgar justificção científica às decisões já adotadas pela cúpula da organização agente” (Cohen&Franco: 1994, 281).

Neste ponto, saímos das restrições de ordem política e entramos na discussão da ética na avaliação. A pergunta “para que” avaliar pode ter diferentes respostas:

“Pode-se avaliar para: I) racionalizar o processo de tomada de decisões ou simplesmente justificá-lo; II) maximizar a consecução dos objetivos no grupo focal ou conseguir financiamento adicional para a mesma ou outras avaliações;

³³Prof. Gilles Dussault e Prof^a Zulmira Hartz, no curso “Epidemiologia na Avaliação de Programas de Saúde”, durante o III Congresso de Epidemiologia da ABRASCO, realizado em Salvador, Bahia, de 24 a 28 de abril de 1995.

III) otimizar a alocação de recursos ou satisfazer a comunidade científica
(Cohen&Franco: 1994, 281).

Por outro lado, como lembram os professores DUSSAULT e HARTZ, no curso já referido, geralmente os técnicos que conduzem a avaliação pensam que, em virtude desta ter sido tecnicamente bem feita, suas recomendações vão ser “naturalmente” difundidas. Porém, relatórios enormes, redigidos em linguagem técnica, não são absorvidos pelos gerentes, que facilmente os “encostam” em uma gaveta. Para que a avaliação seja realmente norteadora dos processos de decisão e de mudança, é necessário que aqueles que conduziram o processo definam uma “estratégia de difusão” dos achados. Ou seja, é necessário adequar o vocabulário tanto aos gerentes quanto aos usuários dos serviços de saúde (ou beneficiários dos programas), com informações simples e de fácil compreensão, adaptadas à cultura tanto de um como de outro. Neste sentido, recomendar-se-ia que fossem elaboradas várias apresentações do relatório final, com diferentes estilos de redação, para melhor comunicar os achados da avaliação.

A terceira dificuldade que apontamos nos estudos avaliativos consiste em “confundir o fracasso do programa com o do estudo avaliativo”, como coloca SUCHMAN (Suchman: 1967, 73). WEILENMAN também chama atenção que quando a avaliação encontra como resultados “efeitos pequenos ou negativos do programa”, é preciso estar atento para a interpretação destes achados, que poderiam ser explicados de dois modos diferentes, os quais afetam a utilização dos resultado da avaliação: ou atribui-se o problema à questões do estudo avaliativo; ou atribui-se ao fato do programa não ter sido adequadamente implementado. Segundo este autor, foi o segundo tipo de explicação que levou os estudos avaliativos a se deslocarem do enfoque sobre os resultados para o enfoque sobre os processos (Weilenman: 1989, 25-26).

SUCHMAN posiciona-se da seguinte maneira sobre esta questão:

“Um estudo avaliativo que não consegue demonstrar nenhum efeito pode ser bem sucedido enquanto uma avaliação - foi o programa que não foi bem sucedido. É possível, é claro, que as medidas da avaliação não tenham sido suficientemente sensíveis, ou mesmo tenham sido incorretamente escolhidas, para mostrar o efeito, mas, desde que o estudo tenha sido cuidadosamente conduzido, a responsabilidade da prova recai sobre o programa em si mesmo” (Suchman: 1967, 73).

Para finalizar, caberia ainda colocar alguns “fatores limitantes” que caracterizam-se como problemas típicos da avaliação (OPS/OMS: 1985, 14):

- em primeiro lugar, há uma certa “resistência intuitiva à avaliação e seus resultados”, que se evidencia em manifestações de auto-defesa e rechaço (OPS/OMS: 1985, 14). Parte desta resistência pode ser atribuída à tendência dos processo avaliativos criarem um clima de controle, o que leva as pessoas a perceberem o avaliador como uma ameaça³⁴;

- em segundo lugar, a avaliação geralmente tem custos altos, o que dificulta obter recursos financeiros para a mesma, especialmente em períodos de diminuição da atividade econômica, onde a restrição de gastos é uma imposição (OPS/OMS: 1985, 14). Por outro lado, a recessão é uma “faca de dois gumes” para a avaliação, pois é justamente nestes períodos, quando não há outra opção além de cortar custos e examinar mais de perto os “hábitos de despesas” que aumenta o interesse “sobre os métodos de redução dos custos e de melhoramento da eficácia, tal como a avaliação” (Love: 1993);

- além das dificuldades de obter recursos financeiros para as avaliações, nem sempre se dispõe do tempo necessário nem dos recursos humanos, tanto em número quanto em qualificação, para executar o trabalho avaliativo (OPS/OMS: 1985, 14);

COHEN&FRANCO chamam atenção para a necessidade de utilizar técnicas de avaliação adequadas ao “tamanho e recursos” do projeto que está sendo avaliado. Especialmente em pequenos projetos, seria “irracional aplicar técnicas sofisticadas”, adequadas a contextos nos quais se dispõe de maiores recursos. Nestes caso, seria recomendável escolher técnicas que permitissem uma avaliação “o mais rápido possível” (Cohen&Franco: 1994, 281).

Todas as dificuldades apontadas demonstram a complexidade do processo avaliativo, especialmente quando pretende caracterizar-se como uma investigação científica. Estas

³⁴Prof. Gilles Dussault e Prof^a. Zulmira Hartz, no curso durante o III Congresso de Epidemiologia da ABRASCO, já referido anteriormente.

dificuldades são essencialmente derivadas da complexidade da abordagem dos fenômenos sociais e da estreita vinculação da avaliação com as políticas públicas. A habilidade específica que se espera do avaliador, que o distingue do pesquisador acadêmico, é justamente a capacidade de conduzir seu estudo dentro destes limites, de uma forma que possa contribuir, concretamente, para a melhoria destas políticas.

1.3.3. As categorias analítico-operacionais da avaliação

A maturidade de um campo ou científico é evidenciada pela clareza dos seus conceitos. Os conceitos são como a “linguagem” de um campo específico de conhecimento; por esta razão, entendemos que seria importante destacar e analisar aqueles que parecem ser os principais conceitos do campo da avaliação. Antes de iniciar esta seção, gostaríamos de esclarecer o que estamos entendendo por “categorias analítico-operacionais” do campo da avaliação.

As vigas mestras de qualquer construção teórica são os “conceitos”, entendidos como “unidades de significação que definem a forma e o conteúdo de uma teoria” (Minayo: 1992, 92), ou ainda, “unidades de pensamento” (Mario Bunge apud Ferrari: 1982, 92), “construções lógicas que se estabelecem de acordo com um sistema de referência e formam parte dele” (Ferrari: 1982, 92). Para FERRARI, os conceitos “são estabelecidos para representar uma determinada espécie de fenômeno ou um setor do fenômeno”, podendo ser considerados “os instrumento de trabalho do cientista, ou ainda como os termos técnicos do vocabulário da ciência” (Ferrari: 1982, 92). Este autor ressalta que um “conjunto de conceitos” não é o suficiente para formar uma teoria (p. 120), pelo contrário, as teorias é que servem “como fonte de formação de conceitos e classificações” (p. 115).

MINAYO define os conceitos

“...como operações mentais que refletem certo ponto de vista a respeito da realidade, pois focalizam determinados aspectos dos fenômenos, hierarquizando-os. Desta forma eles se tornam um caminho de ordenação da realidade, de olhar os fatos e as relações, e ao mesmo tempo um caminho de criação” (Minayo: 1992, 92).

Os conceitos mais importantes de uma teoria são denominados de “categorias”, que se dividem em dois tipos: analíticas e empíricas ou operacionais. As “categorias analíticas” são “abstrações” realizadas com base no “exame da realidade”, que orientam de forma geral o conhecimento e permitem a interpretação (Testa: 1992, 95). Estas categorias funcionam como “balizas” teóricas que permitem conhecer o objeto de pesquisa nos seus aspectos gerais, comportando “vários graus de abstração, generalização e de aproximação” (Minayo: 1992, 92-94). As “categorias empíricas” são construídas com “finalidade operacional”, visando a coleta de dados no trabalho de campo (Minayo: 1992, 92-94). São “conceitos [...] elaborados com fins práticos: manipulação de problemas ou, na investigação, substrato sobre o qual se assenta o funcionamento das categorias analíticas. Trata-se não de uma derivação a partir da realidade, mas de uma imposição sobre ela” (Testa: 1992, 95).

COHEN&FRANCO demonstram que a avaliação segue o mesmos “processo lógico” da metodologia de pesquisa em Ciências Sociais, ao aproveitar os “procedimentos e técnicas” que permitem “transformar os conceitos abstratos em variáveis mensuráveis”, para transformar os objetivos de um programa ou projeto social em indicadores. A partir do trabalho clássico de Paul Lazarsfeld, que apresenta as quatro fases através das quais os conceitos teóricos são transformados em variáveis operacionais, os autores demonstram que existe uma analogia entre “conceitos e objetivos, por um lado, e entre variáveis e metas, por outro” (Cohen&Franco: 1994, 153; 155).

As quatro fases apresentadas por LAZARFELD³⁵ seriam as seguintes:

1ª) a fase da “representação literária do conceito”, ou seja, “uma abstração derivada da observação de um conjunto de fatos particulares” (Cohen&Franco: 1994, 153). Esta seria “a

³⁵LAZARFELD, Paul. Evidence and Inference in Sociological Research. *Daedalus*, 87(4): 99-130, 1958.

imagem inicial do conceito”, que seria uma “entidade vagamente concebida, que dá algum sentido às relações observadas (Ferrari: 1982, 92);

2ª) a fase da “especificação do conceito, que implica em sua divisão nas dimensões que o integram do conceito” (Cohen&Franco: 1994, 153);

3ª) a fase da “seleção dos indicadores observáveis” (Ferrari: 1982, 92), ou seja, “a escolha de indicadores que permitam medir cada uma das dimensões do conceito” (Cohen&Franco: 1994, 153);

4ª) a fase da “combinação de indicadores em índices” (Ferrari: 1982, 92), ou “a formação de índices para sintetizar os dados resultantes das etapas anteriores” (Cohen&Franco: 1994, 153).

Os autores chamam atenção que esta analogia não é casual, porque “a avaliação é realizada em projetos que foram formulados com base na teoria social, que compartilha com esta, a metodologia utilizada”. Entretanto, existem diferenças importantes entre os dois processos: os conceitos dentro de uma teoria são operacionalizados através de um processo formal “dedutivo analítico”, enquanto que na avaliação o processo através da qual um objetivo geral é decomposto em objetivos específicos, e estes em metas que são avaliadas através de indicadores, ocorre a partir de um processo de definição de “prioridades políticas e da racionalidade técnica” (Cohen&Franco: 1994, 155).

A partir desta analogia, entendemos que os conceitos do campo da avaliação poderiam ser apresentados dentro de uma certa hierarquia, partindo de um grau de generalidade mais alto para um nível mais operacional. Entretanto, devido ao caráter essencialmente aplicado dos estudos no campo da avaliação, entendemos que os conceitos que veremos a seguir funcionam ao mesmo tempo como categorias analíticas e empíricas, que por esta razão optamos denominar de “categorias analítico-operacionais”.

Considerando que a maturidade científica de uma teoria reflete-se na clareza de seus conceitos fundamentais, como vimos inicialmente, poderíamos afirmar que ainda há um bom caminho a ser percorrido no campo da avaliação.

Verificamos que a maioria dos autores consultados referem uma grande diversidade terminológica no que concerne aos conceitos deste campo, associada a uma falta de clareza na

diferenciação destes termos. Para SILVER, "existem tantas definições destes conceitos quanto autores que escrevem sobre o tema" (Silver: 1992, 197).

A questão está bem apresentada por SILVA&FORMIGLI, em artigo de revisão no qual se propõe a sistematizar o assunto:

*"O primeiro problema enfrentado por aqueles que procuram tratar a avaliação de forma sistemática é a enorme **diversidade terminológica** encontrada nos enfoques teóricos sobre o tema, tanto no que diz respeito às possíveis abordagens quanto no que concerne aos seus atributos ou componentes.[grifo nosso]"* (Silva & Formigli: 1994, 81).

A simples proposta das autoras de fazer esta sistematização indicaria, por si só, o estado incipiente da avaliação enquanto campo disciplinar. Além disto, verificamos que a própria forma como os autores referem-se aos conceitos campo da avaliação é variada: "conceitos", "atributos", "noções", "abordagens"; "componentes", "enfoques" da avaliação (Silver: 1992, 197; Silva&Formigli: 1994, 81; Pitta: 1992, 45). Entendemos que estas diversas maneiras de perceber o que estamos entendendo como conceitos básicos do campo da avaliação são indícios da imaturidade teórica deste campo disciplinar.

Ao discutir o significado do que denominam de "atributos" da avaliação - eficácia, eficiência, efetividade, impacto, qualidade, acessibilidade, cobertura, equidade e satisfação do usuário - SILVA&FORMIGLI chamam atenção para as "limitações", "inconsistências" e "inadequações" e "baixa especificidade" destes termos. Para elas, isto ocorre devidos às fragilidades da "concepção sistêmica para análise do real", referencial teórico para a maioria dos autores por elas revisados (Silva&Formigli: 1994).

Mais adiante, as autoras colocam que os problemas do campo referem-se principalmente "à diversidade terminológica discutida, mas também às inadequações e insuficiências metodológicas indicadoras não somente de um **incipiente desenvolvimento teórico do campo**, como mesmo da **ausência de 'acordos de grupo'** visando uma uniformização conceitual [grifos nossos]" (Silva&Formigli: 1994, 81-88). Mais ou menos nesta mesma direção, BATTISTA et al. atribuem esta "confusão terminológica" ao labirinto de teorias, métodos e estratégias de análise no campo da pesquisa em saúde, decorrente da disparidade de disciplinas que concorrem no campo (Battista et al.: 1989, 1155). Ou seja, de

certa forma, ambos estão referindo questões típicas de um campo que se forma através das relações interdisciplinares, que é a nosso ver a causa principal da diversidade terminológica.

Seja qual for a explicação para este problema, ficamos com a impressão de que a questão essencial não é a diversidade de **termos** utilizados - no sentido de vocábulos, palavras - que de uma forma geral são os mesmos, mas sim uma diversidade na atribuição de seus **significados** - no sentido de conceito ou idéia que querem representar -, os quais estes sim apresentam variações, embora na maioria das vezes muito sutis, como veremos a seguir. Ou seja, o problema parece ser relativo à maturação de conceitos, e não à diversidade de palavras.

SILVA&FORMIGLI agrupam, de forma sucinta e didática, o que consideram os principais atributos a serem avaliados nas práticas de saúde e na sua organização social:

“a) relacionados com a disponibilidade e distribuição social dos recursos (cobertura, acessibilidade e equidade); b) relacionados com o efeito das ações e práticas de saúde implementadas (eficácia, efetividade e impacto); c) relacionados com os custos das ações (eficiência); d) relacionados com a adequação das ações ao conhecimento técnico e científico vigente (qualidade técnico-científica); e) relacionados à percepção dos usuários sobre as práticas (satisfação dos usuários, aceitabilidade)” (Silva&Formigli: 1994, 81).

Não iremos tratar de todos estes conceitos, apenas daqueles que consideramos mais gerais. A partir da menção que estas autoras fazem à “clássica ‘tríade’ estrutura-processo-resultado” (Silva&Formigli: 1994, 83), percebemos que os principais conceitos do campo poderiam ser apresentados como “tríades”. Buscamos então estabelecer uma espécie de hierarquia entre estes conceitos, partindo de um plano mais conceitual para um plano mais operacional, seguindo o processo caracterizado por MINAYO e LAZARSELD, que vimos anteriormente.

Selecionamos aqueles que consideramos ser os conceitos básicos do campo da avaliação: estrutura-processo-resultado; eficácia-efetividade-eficiência; indicador-critério-padrão. A primeira destas tríades define as abordagens ou enfoques metodológicos mais gerais do campo da avaliação. A segunda, constitui a conceituação mais geral de onde são definidos

os indicadores. A terceira, tem caráter essencialmente operacional. Além disto, fazemos algumas considerações sobre o conceito de qualidade, que é a essência do campo da avaliação.

1.3.3.1. A tríade estrutura-processo-resultado

Esta clássica tríade conceitual foi introduzida no campo da saúde por Avedis Donabedian, para classificar as abordagens em avaliação, a partir do referencial sistêmico (Silva&Formigli: 1994, 83). Estes conceitos foram desenvolvidos pela economia, associados à “função de produção” do setor industrial (Cohen&Fraco: 1994, 93). Como vimos anteriormente, na seção sobre avaliação de qualidade, estes conceitos correspondem, em nosso entendimento, às noções de “input-process-output”.

Segundo PITTA, este é o quadro conceitual que vem sendo utilizado há mais tempo, tendo “alcançado certa consensualidade” no campo da avaliação (Pitta: 1992, 45). Para SILVER, este é o “marco conceitual mais usado para classificar as diversas metodologias de avaliação de serviços” (Silver: 1992, 202), ou, conforme SILVA&FORMIGLI, estes conceitos referem-se às principais “abordagens (*approach*)” do campo geral da avaliação (Silva&Formigli: 194, 83). Porém, a primeira chama atenção que “embora aparentemente simples de distinguir os componentes estruturais daqueles relativos ao processo e aos resultados, essa distinção nem sempre é tácita” entre os autores (Pitta: 1992, 45).

A confusão começa na forma como os autores utilizam estes termos: ora são referidos como “componentes” do paradigma conceitual de Donabedian, dando a entender que trata-se de conceitos teóricos do campo da avaliação; ora são referidos como diferentes tipos de abordagem metodológica ou “*approach*”, dando a entender que se referem a um enfoque de pesquisa; ora são entendidos como “etapas” do processo avaliativo (Silva&Formigli: 1994, 83-

84; Silver: 1992, 202). As diferentes formas de referir estes conceitos certamente refletem as dificuldades para caracterizá-los, conforme veremos a seguir.

De uma forma geral, entende-se por **estrutura** as “características relativamente estáveis dos prestadores de serviços, das ferramentas e recursos disponíveis e do contexto físico e organizacional no qual trabalham” (Donabedian Apud Silver: 1992, 202). Ou seja, a estrutura corresponde aos insumos ou recursos (“input”) que estabelecem as condições para se desenvolverem os processos, mas também ao contexto organizacional onde os mesmos ocorrem. Neste sentido, COHEN&FRANCO fazem uma distinção, no conceito de estrutura, entre “os recursos e a organização” (Cohen&Franco: 1994, 94).

Segundo DUSSAULT, os recursos que compõe a estrutura são de diversos tipos: “cognitivo (saber, tecnologia), físico (edifícios, materiais, suprimentos, equipamentos, etc.), financeiro e humano” (Dussault: 1994, 479). Observamos que o autor separa o aspecto cognitivo das pessoas que o detém, dando a entender que o conhecimento científico (saber, tecnologia) é um recurso em si.

COHEN&FRANCO distinguem os “recursos” dos “insumos”, embora sejam usados geralmente como sinônimos. Para eles, o termo “recursos se usa no âmbito de uma organização relativamente estável (como a dos efetores de saúde), enquanto que insumos costuma estar ligado a processos que carecem de permanência”. Por esta razão, o primeiro termo estaria associado a “projetos permanentes”, e o segundo aos projetos de “existência mais precária” (Cohen&Franco: 1994, 94).

Além dos recursos **humanos** (tipo e número de trabalhadores), **materiais** (instrumentos, equipamentos, planta física) e **financeiros**, o conceito de estrutura engloba também os aspectos **organizacionais**. PITTA menciona os “recursos administrativos (Pitta: 1992, 45), enquanto DUSSAULT fala em “fatores organizacionais”, como “a divisão e a organização do trabalho (definidas por leis, contratos, convênios, tradições), o quadro jurídico-administrativo da produção e da distribuição dos serviços, do sistema de incentivos (a maneira, o nível e a estrutura de remuneração e de pagamento)” (Dussault: 1994, 479). Para SILVER, a concepção de estrutura inclui a “gama de serviços” oferecidos, bem como as características dos mesmos “que determinam continuidade da assistência, acessibilidade, financiamento e definição da população com direito ao uso” (Silver: 1992, 202). SILVA&FORMIGLI, por sua

vez, consideram que “a própria função de monitoramento é parte da estrutura”, embora este geralmente seja entendido como uma questão relativa ao processo (Silva&Formigli: 1994, 83).

Em resumo, a estrutura corresponderia de uma maneira geral ao conjunto dos “recursos utilizados”: físicos, materiais, humanos, financeiros e administrativos (Pitta: 1992, 45).

A principal crítica ao conceito de estrutura refere-se à colocação das pessoas como meros “insumos” ou “recursos” a serem utilizados no processo, e não como “sujeitos da ação” que constitui o próprio processo. Para SILVA&FORMIGLI este conceito opera uma “homogeneização de fenômenos de natureza distinta, como, por exemplo, recursos humanos e materiais englobados sob o mesmo rótulo” (Silva&Formigli: 1994, 84). DUSSAULT coloca que referir-se à “gestão dos recursos humanos” é uma “expressão ambígua que coloca a contribuição das pessoas no mesmo nível que equipamentos ou dinheiro mas que enfatiza, ao mesmo tempo, a idéia de que as pessoas são recursos com um potencial tradicionalmente esquecido e não utilizado” (Dussault: 1994, 484).

Quanto ao conceito de **processo**, pode ser definido, de uma forma geral, como o “conjunto de **atividades** que se realizam para tentar atingir o objetivo procurado (Cohen&Franco: 1994, 93). Segundo PITTA, a área de processo “implica as atividades desenvolvidas em termos de utilização dos recursos nos seus aspectos **quantitativos** e **qualitativos** com a produção tecnológica processada por seus atores [grifos nossos]” (Pitta: 1992, 45). Referindo-se especificamente ao campo da saúde, SILVA&FORMIGLI definem como o “conjunto de atividades desenvolvidas na relação entre profissionais e pacientes” (Silva&Formigli: 1994, 83).

Enquanto estas definições enfatizam o processo como um conjunto de “atividades”, SILVER define-o como “um **comportamento normativo** que tem duas faces: aquela composta pelos prestadores de serviços e a que se refere aos destinatários dos serviços [grifo nosso]” (Silver: 1992, 202). No que concerne aos prestadores, o processo teria quatro componentes: “reconhecimento do problema ou risco, formulação do diagnóstico ou necessidade, gestão e seguimento”. Ou seja, nesta concepção, o conceito de “processo” corresponde quase que integralmente às etapas do planejamento (só fica faltando a formulação do plano). No que concerne aos destinatários dos serviços (comunidade), o processo “diz respeito à utilização, aceitação (ou não), compreensão (ou não) dos serviços e participação”.

Ou seja, neste caso o conceito corresponde à forma como os usuários/clientes consomem os serviços e se relacionam com estes (p. 202).

Vemos que tanto no entendimento do processo como “atividade” ou como “comportamento normativo”, está referido às relações entre prestadores de serviço e usuários, implicando em aspectos “quantitativos e qualitativos”.

A relação entre a estrutura e o processo é muito bem colocada por SILVER : o conjunto de aspectos que compõem a estrutura determinam “o potencial do sistema”, enquanto que o processo “determina como o sistema realmente funciona” (Silver: 1992, 202).

O conceito de resultado, tradicionalmente, está vinculado a idéia de alcance de um objetivo, ou mais precisamente, do grau de alcance. Entretanto, em uma concepção mais ampliada, os resultados referem-se aos efeitos de uma ação, que podem ter sido intencionais ou não. No campo da saúde, as ações ou intervenções visam mudanças no estado de saúde das pessoas ou populações, como fica bem claro na definição de PITTA: “Os resultados correspondem às conseqüências da atividade da **instituição** ou **profissional** em análise para a saúde de **indivíduos** ou **populações** [grifos nossos]”. (Pitta: 1992, 45)

Segundo SILVER, Donabedian define resultado como “a mudança no estado de saúde atual e futuro do paciente atribuível à assistência médica anterior, incluindo tanto mudanças sociais e psicológicas como as tradicionais físicas e fisiológicas, assim como as de atitude, satisfação e comportamento” (Silver: 1992, 203). Como veremos na seção referente ao conceito de qualidade, Donabedian refere-se em seus trabalhos à assistência médica individual, embora entenda que seus conceitos possam ser extrapolados para intervenções como maior nível de abrangência e complexidade.

Embora a noção de “resultado” seja aparentemente óbvia, observamos que os autores utilizam diferentes termos para se referir aos “efeitos” de uma intervenção sobre a realidade, cujo significado é diferente: “produto”, “eficácia” e “efetividade”, “impacto”.

COHEN&FRANCO caracterizam as diferenças entre os termos “produto”, “efeito” e “impacto”, todos incluídos em uma noção mais ampla de “resultado”. Os produtos seriam “os resultados concretos das atividades desenvolvidas a partir dos insumos disponíveis”, e poderiam ser de caráter material (bens) ou serviços prestados. Os efeitos, por sua vez, seriam

os “resultados da utilização dos produtos do projeto”. O impacto seria a “consequência dos efeitos de um projeto” (Cohen&Franco: 1994, 92-93).

Embora o autor não explicita, percebemos claramente que a noção de produto corresponde à noção sistêmica de “out-put”, ou seja, o produto é o resultado do processamento dos insumos (“input”). A noção de “efeito” corresponde ao sentido clássico do termo “resultado”, qual seja, o alcance de metas previamente determinadas. Entendemos que estas noções correspondam à diferença que existe, na língua inglesa, entre os termos “*out-put*“, que significa “resultado” e “*outcome*“, que significa o “alcance de metas do programa” (Silver: 1992, 200). Ou melhor, “*input*” corresponde à entrada e “*out-put*” à saída, enquanto “*outcome*” é o evento ou resultado final.

As noções de produto e efeito estão voltadas para o funcionamento do programa ou projeto em si mesmo, enquanto a noção de impacto incorpora uma perspectiva populacional, tanto no tempo - os efeitos sobre uma população podem manifestar-se até muito tempo depois do fim do projeto (Aguilar&Ander-Egg: 1994, 40; Weilenmann: 1980: 13) - quanto no espaço, na medida em que o impacto incorpora a noção de “cobertura”³⁶ do programa.

Ou seja, um programa ou projeto pode gerar muitos “produtos” (por exemplo, número de consultas médicas efetivamente realizadas), e isto resultar em um “efeito” pequeno sobre os indicadores de mortalidade (resultado esperado do programa), devido à má utilização deste produto (por exemplo, por uma clientela que não corresponde à população alvo). Por outro lado, mesmo alcançando o resultado esperado (por exemplo, que as consultas sejam utilizadas pela população alvo), o programa pode ter pouco impacto sobre a população, tanto porque tem baixa cobertura como porque há uma inadequação do mesmo em relação ao seu contexto de inserção.

Haveria ainda que tratar dos tipos de efeito das intervenções. SUCHMAN categoriza os tipos de efeito utilizando os mesmos quatro critérios para a classificação dos objetivos: (a) unidade ou multiplicidade de efeitos; (b) intencionalidade ou não dos efeitos; (c) duração dos efeitos; (d) tipo de efeito (cognitivo, atitudinal ou comportamental) (Suchman: 1967, 67). Em relação ao primeiro critério, o autor chama atenção que raramente um programa busca atingir

³⁶Cf. aula da Prof^a Zulmira Araújo Hartz, sobre “Métodos Epidemiológicos aplicados à Avaliação de Programas”, no Curso de Pós-Graduação em Saúde

apenas um objetivo, da mesma forma, uma intervenção dificilmente produzirá um efeito único: geralmente, a ação tem “múltiplos efeitos” (pp. 37, 40). No que concerne à intencionalidade, destaca que as intervenções podem produzir efeitos desejados (ou positivos), mas também efeitos não desejados (ou negativos), além de efeitos “colaterais”, “não esperados” ou “não intencionais”, que não são necessariamente negativos (pp. 37, 40). Quanto à duração, os efeitos podem ser classificados em imediatos, intermediários e de longo prazo (p. 68-70). Os tipos de efeito, por sua vez, relacionam-se com a “área de interesse” das mudanças desejadas, ou seja, mudanças nos conhecimentos, nas crenças, nas atitudes, nas motivações ou nos comportamentos das pessoas que são alvo da intervenção (p.39). Em princípio, todos estes tipos de efeito deveriam ser considerados na avaliação dos resultados de uma intervenção para emitir um juízo sobre a mesma. Seria muito temerário avaliar uma intervenção considerando apenas um efeito, ou apenas os efeitos desejados, ou apenas os efeitos de curto prazo, ou ainda aqueles relacionados apenas a uma área de interesse.

Uma das dificuldades na avaliação de resultados é isolar os efeitos que podem ser específica e unicamente atribuídos à intervenção em estudo. No campo da saúde, SILVER chama atenção que, além do sistema médico assistencial, existe uma complexidade dos fatores sociais e econômicos que influem no estado de saúde das populações. Isto dificulta, na prática de avaliação, identificar quais as mudanças atribuíveis à assistência (Silver: 1992, 203).

Além disto, é necessário lembrar que enquanto alguns efeitos aparecem durante a execução da intervenção, outros podem manifestar-se muito tempo depois do seu término, ou seja, “o impacto do projeto pode continuar mesmo quando este já tenha sido descontinuado há um tempo” (Weilenmann: 1980, 13). Por outro lado, nos projetos sociais existem os efeitos diretos, isto é, aqueles que atingem a população alvo, e os efeitos “indiretos” ou “externalidades”, que são produzidos em consequência do projeto, porém fora do âmbito em que este é realizado, ou seja, são efeitos que atingem outras populações.(Contandriopoulos et al.: 1992; Cohen&Franco: 1994, 167; 183; 185).

Como vimos no início, a caracterização destes conceitos não é fácil. Segundo SILVA& FORMIGLI, o próprio Donabedian teria apontado as dificuldades “para a distinção entre as três categorias do paradigma”, que são mais frequentes “quando se separa o mesmo do âmbito

clínico, para o qual ele foi desenhado, e se tenta utilizá-lo em outros tipos de avaliação”. Para elas, os termos desta primeira tríade tem “baixa especificidade”, sendo necessário explicitar “qual estrutura, qual processo e quais resultados pretende-se estudar”, visto que um mesmo aspecto da prestação de serviços pode ser considerado como estrutura, como processo ou como resultado, citando como exemplo o modelo assistencial, o sistema de referência e contra-referência, o monitoramento. (Silva&Formigli: 1994, 84)

WEILENMANN partilha deste mesmo ponto de vista, especialmente no que concerne às diferenças entre “processo” e “estrutura”. Baseado nas idéias de Miller³⁷, ele coloca que “a distinção entre estrutura e processo freqüentemente não é explicitada, e os dois termos tendem a criar muita confusão”. Para ele, qualquer coisa no mundo pode ser vista deste dois modos, diferentes mas complementares: o ponto de vista do processo implica em considerar “mudança e movimento ao longo do tempo”, enquanto o ponto de vista da estrutura implica em considerar o “arranjo das coisas no espaço e num momento do tempo”. Estes conceitos funcionam como “os dois lados da moeda”, pois os “processos são executados, ou levados ao cabo (*carried out*) pelas estruturas, e as estruturas são sujeitas à mudança e movimento; desta forma, manifestam os processo ao longo do tempo” (Weilenmann: 1980, 38-39).

Além disto, para este autor, o termo “estrutura” tende a criar dificuldades. Ainda referindo as idéias de Miller, ele coloca que é necessário distinguir entre “a estrutura de um sistema” e as “estruturas em um sistema”. No primeiro caso - estrutura **de** um sistema -, o termo estrutura refere-se “ao arranjo espacial e às relações entre as partes do sistema em um determinado momento do tempo”. Estes arranjos e relações “são relativamente persistentes ao longo do tempo”. No segundo caso - estruturas **em** um sistema -, o sentido do termo estrutura refere-se às “partes, entidades, unidades definidas ou vistas estruturalmente”. Ele ressalta ainda que “enumerar as estruturas [neste segundo sentido] não é suficiente para caracterizar totalmente a estrutura do sistema; é necessário considerar as relações relativamente persistentes entre elas” (Weilenmann: 1980, 57).

As sutilezas de significado destes termos ficam ainda mais complexas quando o autor distingue os dois sentidos que podem ter as “partes” de um sistema. Existem as “partes” do sistema que podem ser caracterizadas como “processos”, as quais podem ser interrompidas em

algum momento, as quais são denominadas de “subsistemas”. Mas existem também as “partes” do sistema que são caracterizadas “estritamente em termos estruturais”, e que são denominadas de “componentes” do sistema (Weilenmann: 1980, 39-40).

Transpondo este raciocínio para as grandes linhas de abordagem que vimos anteriormente - Avaliação Normativa e Pesquisa Avaliativa -, verificamos que a primeira enfoca os componentes individuais do sistema ou programa, ou seja, trabalha com as partes “em um sistema” (componentes); enquanto que a segunda enfoca as relações entre estes componentes, ou seja, trabalha com os arranjos e as relações das partes “de um sistema” em um determinado momento do tempo.

1.3.3.2. A tríade eficácia-efetividade-eficiência

A segunda tríade do campo da avaliação são os conceitos de eficácia-efetividade-eficiência, que referem-se aos “efeitos” das ações ou intervenções, conforme vemos na classificação dos “atributos” da avaliação de SILVA&FORMIGLI, mencionada no início desta seção. SUCHAMN caracteriza estes conceitos como “critérios de sucesso ou fracasso” dos programas (Suchman: 1967, 60). Por esta razão, estes conceitos funcionam como categorias analíticas de onde são formulados os indicadores de resultado.

As primeiras ressaltam que freqüentemente há uma superposição entre as definições dos mesmos, da mesma forma que com a noção de “impacto”, mas entendem que tem havido “uma evolução na atribuição de significado a alguns destes conceitos” (Silva&Formigli: 1994, 82).

De uma forma geral, os conceitos de eficácia e eficiência se definem “em relação ao cumprimento de objetivos propostos”, enquanto que o de eficiência corresponde “às relações entre custos e resultado, ou entre resultado e insumos”, ou simplesmente, relação custo/benefício (Silva&Formigli: 1994, 82-83). Segundo BATTISTA et al., as preocupações

³⁷MILLER, J.G. *Living Systems*. New York, McGraw-Hill, 1978.

primárias da Pesquisa Avaliativa são a avaliação da eficácia e da efetividade das intervenções. Somente após essas estarem bem demonstradas é que se poderia incorporar a “análise econômica” à avaliação, ou seja, a análise de eficiência (Battista et al.: 1989, 1158-59).

Como vimos anteriormente, ao tratarmos das linhas metodológicas da avaliação em saúde, as análises de eficácia, efetividade e de eficiência são análises das relações entre os componentes do programa, e correspondem ao tipo de análise feito pela Pesquisa Avaliativa. Assim, as análises de eficácia e efetividade enfocam a relação entre os “serviços” ou “atividades” produzidas e os seus “efeitos”, enquanto a análise de eficiência enfoca a relação entre os recursos utilizados e os efeitos produzidos (custo/benefício). Alguns autores consultados chamam atenção que há uma hierarquia ou uma seqüência lógica nestas análises (Pitta: 1992, 46; Battista et al.: 1989, 1158-59; Clemenhagen&Champagne: 1986, 385, Fig. 1). Esta hierarquia fica bem evidente no Anexo V.

A seqüência lógica entre estas análises (ou entre estes conceitos) seria a seguinte:

1º - A **eficácia** responde à pergunta: “Pode funcionar?” (Silver: 1992, 201). Ou seja, a análise de eficácia deve esclarecer se a intervenção é “capaz de produzir o efeito desejado” (Pitta: 1992, 46). Segundo alguns autores, a forma ideal para determinar a eficácia de uma intervenção é através de um experimento randomizado e controlado (Battista et al.: 1989, 1158; Last: 1988). Com isto, seria garantida “ótima implementação” das medidas interventivas (ou “condições ideais” de implementação), e ao final se examinaria as diferenças de efeito no grupo experimental e no grupo controle, de tal forma que as mudanças observadas no estado de saúde do primeiro (ou qualquer outro estado de coisas que se pretenda mudar) possam ser atribuídas à intervenção (Battista et al.: 1989, 1158-59). Em suma, a eficácia se refere ao “efeito potencial ou o efeito em determinadas condições experimentais” (Silva&Formigli: 1994, 82), ou o efeito “potencial máximo” da intervenção (Pitta: 1992, 46). LAST deixa bem claro o conceito de eficácia em sua definição: “É o grau no qual uma intervenção, procedimento, regime ou serviço produz um efeito benéfico sob condições ideais” (Last: 1988).

Lembramos que alguns autores definem quatro tipos de eficácia, conforme vimos na seção referente ao discurso integrado da avaliação em saúde: eficácia teórica (contexto laboratorial); eficácia dos ensaios (contexto dos ensaios clínicos randomizados); eficácia de

utilização (contexto das práticas normais) e eficácia populacional (contexto comunitário-populacional), sendo que as duas últimas correspondem ao conceito de efetividade, que veremos à seguir (Contandriopoulos et al: 1992).

2° - A **efetividade** responde à pergunta: “Funciona?” (Silver: 1992, 201). Ou seja, a efetividade seria a tradução da eficácia “para uma situação do mundo real” (p. 201) ou o “impacto real” da intervenção (Pitta: 1992, 46), ou ainda o “efeito real num sistema operacional” (Silva&Formigli: 1994, 82). LAST define efetividade como “o grau no qual uma intervenção, procedimento, regime ou serviço específico executa o que foi previsto, sob condições reais” (Last: 1988).

SILVA&FORMIGLI sugerem a semelhança entre os conceitos de efetividade e **impacto**, na medida em que consideram que as diferenças apontadas entre os mesmos não tem uma “fundamentação convincente”. Estas diferenças seriam definidas tanto em relação aos diferentes níveis de objetos da avaliação quanto em relação ao tempo. Segundo DE GEYNDT³⁸, o efeito da intervenção sobre um “grupo populacional” corresponderia à efetividade, enquanto que o “efeito de um ‘sistema’ sobre uma população expressaria o impacto” (De Geyndt apud Silva&Formigli: 1994, 82). RUNDALL³⁹, por outro lado, utiliza a categoria tempo para distinguir os dois termos: “o curto prazo diria respeito à efetividade e o longo prazo, ao impacto” (Rundall apud Silva&Formigli: 1994, 82).

3° - A **eficiência** deveria responder à pergunta: “Vale a pena fazer?” (Silver: 1992, 201). Segundo SILVER, a eficiência se refere aos “efeitos ou resultados em relação ao esforço realizado, expressos em termos de dinheiro, recursos e tempo” (p. 201). Voltando à LAST, este autor define eficiência como “o grau de minimização dos recursos utilizados para prover uma intervenção, procedimento, regime ou serviço de conhecida eficácia e efetividade. É uma medida da economia - o custo dos recursos com os quais um procedimento de conhecida eficácia e efetividade é realizado” (Last: 1988). Baseando-se em Donabedian, SILVA&FORMIGLI definem eficiência como o “máximo cuidado efetivo obtido ao menor

³⁸DE GEYNDT, W. Five approaches for assessing the quality of care. *Hospital Administration*, 15: 21-41, 1970.

³⁹RUNDALL, T.G. Health planning and evaluation. In: *Public Health & Preventive Medicine*. LAST, J.M. & WALLACE, R.B. (orgs). Connecticut, Appleton&Lange (Maxcy-Rosenau-Last), 13ª ed, 1992.

custo”, enquanto PITTA a define como a “relação entre o impacto real e os custos da produção” (Pitta: 1992, 47).

Lembramos que o conceito de eficiência comporta três sub-tipos de análise: custo-benefício, custo/eficácia e custo/efetividade, sendo que no primeiro tanto o numerador como o denominador são expressos em termos monetários, e nos dois últimos o denominador é expresso por índices de resultados ou de qualidade, que não se expressam monetariamente (como anos de vida ganhos, índice de satisfação, indicadores de qualidade de vida) (Contandriopoulos et al.: 1992).

Fica evidente que a análise de eficiência somente poderia ser realizada uma vez que a eficácia e a efetividade de uma intervenção estivessem demonstradas. Não há sentido em avaliar a eficiência econômica de uma intervenção que não é eficaz e efetiva, ou seja, que não funciona.

Além disto, outro aspecto fundamental da análise de eficiência é que este conceito traz, implicitamente, a perspectiva da comparabilidade entre os custos de diferentes maneiras alternativas de intervir na realidade.

Neste sentido, segundo SUCHMAN a pergunta fundamental ao qual busca responder a análise de eficiência é: “Existe uma maneira melhor de obter os mesmos resultados? “ A eficiência diz respeito, portanto, à avaliação “de caminhos ou métodos alternativos em termos de custo - em dinheiro, em tempo, em pessoal e em conveniência pública” (Suchman: 1967, 64). Observamos que ao destacar a “conveniência pública” de uma certa forma de intervenção, o autor vai além da questão econômica, pois está indiretamente contemplando o aspecto político da eficiência.

Da mesma forma, BATTISTA e cols. colocam que a análise de “custo-efetividade” deveria ser feita comparando as diferenças de custo entre duas ou mais intervenções com efeitos (eficácia e efetividade) presumivelmente iguais. Somente assim um programa poderia ser julgado em termos de eficiência (Battista et al.: 1989, 1159). Para estes autores, o desenho experimental da Pesquisa Avaliativa seria o modo mais adequado de realizar a análise de eficiência de um programa.

De uma forma geral, poderíamos afirmar que a tríade estrutura-processo-resultados refere-se às abordagens da Avaliação Normativa, porque enfocam os elementos que compõe

um programa; enquanto que a tríade eficácia-efetividade-eficiência está mais afeta às abordagens da Pesquisa Avaliativa, pois enfoca as relações entre estes componentes.

1.3.3.3. A tríade indicador-critério-padrão

A tríade “indicador-critério-padrão” pertence ao nível mais operacional dos conceitos do campo da avaliação, e está relacionada à questão da medição dos fenômenos que vão fundamentar o juízo avaliativo. É neste nível que os estudos em avaliação buscam superar, através de medidas “objetivas” dos fenômenos, a subjetividade envolvida no “juízo” final de qualquer avaliação.

Se os conceitos que vimos anteriormente apresentam dificuldades quanto à clareza de significado, no caso destes, a confusão parece ser maior. Vemos freqüentemente o uso intercambiável dos termos “indicador” e “critério”, também às vezes simplesmente mencionados como “variáveis”. Porém em relação ao conceito de “norma” ou “padrão” parece haver maior consenso.

Segundo FERRARI, “a operacionalidade [...] inicia-se com a fundamentação adequada e clara dos indicadores”. Para este autor, “a definição operacional [dos conceitos] encoraja o emprego de métodos e técnicas quantitativas de pesquisa” (Ferrari: 1982, 101). Retornado às quatro fases de Lazarsfeld para a transformação de conceitos em variáveis, vemos que a escolha dos indicadores corresponde à terceira e penúltima etapa.

VASCONCELOS⁴⁰ define indicador como “a quantificação da realidade a fim de planejar um modo de interferir nesta própria realidade” (Vasconcelos apud Moraes: 1994, 21). COHEN&FRANCO como “a unidade que permite medir o alcance de um objetivo específico” (Cohen&Franco: 1994, 152).

O aspecto fundamental da operacionalidade dos conceitos é, portanto, a quantificação. Isto fica evidente no raciocínio de KISIL&CHAVES, que reproduzimos a seguir:

“Avaliar é comparar e, por isso, exige a definição de parâmetros, padrões e normas, como referência para a medida da natureza e magnitude dos processos de mudança. Para tanto, são necessários os indicadores, ou seja, variáveis que permitem quantificar determinadas características de uma situação, ou programa e, por comparação, medir as mudanças ocorridas em um dado período, ou as diferenças existentes entre vários programas e situações num determinado momento” (Kisil&Chaves: 1994, 100).

Ou seja, para julgar, é preciso comparar; e para comparar, é preciso “medir”. O pressuposto é que, “pelo fato de serem quantitativos”, os **indicadores** permitem “reduzir a subjetividade do julgamento” (Kisil&Chaves: 1994, 102). Destacamos ainda, neste raciocínio, a utilização do termo “variável” como sinônimo de indicador, fato que observamos em outros autores (OPS/OMS: 1985, 16), e que certamente reflete a analogia entre a operacionalização de conceitos nas Ciências Sociais e a definição de indicadores em avaliação, que apontamos no início desta seção.

A sobreposição do conceito de “indicador” ao conceito de “variável” é muito evidente, como vemos na citação acima. Ao tratar das escalas de valor aplicáveis aos indicadores, COHEN&FRANCO mencionam as mesmas escalas que vemos na literatura de metodologia da pesquisa para a operacionalização de variáveis, que irão definir as “ferramentas estatísticas” a serem utilizadas, quais sejam, escalas de razão, de intervalo, ordinal ou nominal (Cohen&Franco: 1994, 141).

Entretanto, vale lembrar novamente MINAYO, ao colocar que quando se leva em conta “os níveis mais profundos das relações sociais” - como é o caso dos valores subjacentes que norteiam os processos avaliativos - nem tudo pode ser operacionalizado “em números e variáveis” (Minayo: 1992, 10). Observa-se freqüentemente nos programas do setor saúde situações ou atividades que “não se prestam a medição dos resultados em função de objetivos numéricos e previamente determinados” (OPS/OMS: 1985, 14,16). VERGARA aponta que é muito difícil medir “certos elementos intangíveis, tais como: valores, atitudes, sentimentos,

⁴⁰VASCONCELOS, M.M. *O uso de indicadores em epidemiologia*. Rio de Janeiro, UERJ/Instituto de Medicina Social, mimeo, 1977.

etc”, nestes casos, os indicadores apenas “detectam a presença” do aspecto intangível (Vergara: 1980, 63).

De qualquer forma, VASCONCELOS lembra que os indicadores “refletem o **Sistema de Valores** do profissional que os constrói, o que implica em que formule certas hipóteses e perguntas e não outras” (Vasconcelos apud Moraes: 1994, 21).

O segundo aspecto fundamental do conceito de indicador está evidente na própria etimologia do termo: o indicador não é a medida direta do fenômeno que está sendo estudado, mas uma medida indireta, que apenas “indica” a maneira como estes fenômenos ocorrem na realidade social (OPS/OMS: 1985, 14,16). Isto ocorre porque na esfera social, os conceitos correspondem “quase sempre, a um conjunto complexo de fenômenos e não a um fenômeno simples de diretamente observável” (Lazersfeld apud Cohen&Franco: 1994, 153). Em função da complexidade e da impossibilidade, muitas vezes, de medir diretamente as mudanças nos fenômenos sociais, é necessário abordá-los por partes, através de indicadores.

A essência desta questão fica bem demonstrada nestas colocações de FERRARI:

“Os indicadores designam ou selecionam certas propriedades características que tem a ver com o conceito ou com o objeto de estudo ou pesquisa, porque nenhuma ciência consegue lidar de um modo completo na sua total concretização com o seu objeto de estudo. Cada um dos conceitos amplos devem ser decifrados em termos de critérios e indicadores. Por isso, ao se estabelecer os indicadores, apenas se está considerando um parte, ao mesmo tempo que se tenta esboçar as relações entre essas partes. A tarefa do cientista, portanto, reside na sua capacidade de estabelecer indicadores para cada uma das partes, setores, dimensões etc., do fenômeno a ser indagado ou já apreendido” (Ferrari: 1982, 102).

Este autor esclarece que o indicador “não tem relação absoluta com a matéria objeto da pesquisa, mas apenas tem um alto grau de probabilidade de caracterizá-lo, mensurá-lo ou dimensioná-lo”. Tomado isoladamente, cada indicador “tem apenas relação de probabilidade com o aspecto do fenômeno social que se está indagando” (Ferrari: 1982, 102;104). Ou seja, os indicadores seriam “variáveis que ajudam a medir as mudanças que não podem ser diretamente determinadas” (OPS/OMS: 1985, 16). Por isto, o indicador deve ser considerado

“apenas índice, reflexo ou expressão indireta do fenômeno, ou da característica que se analisa, nem sempre suscetível de ser quantificada” (Kisil&Chaves: 1994, 102).

Considerando que cada indicador é apenas um reflexo parcial do fenômeno estudado, entende-se que um indicador, por si só, não seria suficiente para apreendê-lo integralmente, sendo necessário a utilização de “diferentes combinações” de indicadores, que funcionariam de modo complementar, especialmente no caso da avaliação da eficácia e dos efeitos de programas e projetos (Kisil&Chaves: 1994, 102). COHEN&FRANCO recomendam que seja utilizada “uma quantidade elevada de indicadores para garantir que o conceito que se pretende medir esteja sendo medido e, eventualmente, diminuir o efeito negativo derivado da escolha de um mau indicador” (Cohen&Franco: 1994, 153).

Estes últimos classificam os indicadores em “diretos e indiretos”. Os indicadores diretos seriam uma decorrência lógica do objetivo específico: por exemplo, se reduzir a taxa de repetência escolar é um objetivo específico, a variação percentual desta taxa já é, por definição, um indicador deste objetivo. Os indicadores “indiretos” não seriam uma simples tradução do objetivo, mas teriam de ser “construídos (Cohen&Franco: 1994, 152).

De forma geral, os indicadores são referidos aos conceitos que vimos anteriormente: indicadores de estrutura, de processos ou de resultados, ou indicadores de eficácia, de efetividade ou eficiência. WEISS refere-se que os indicadores de “input” são as variáveis independentes do estudo, os indicadores de “outcome” são as variáveis dependentes e os indicadores que intermedeiam os anteriores são as variáveis intervenientes do estudo (Weiss: 1972, 34).

No campo da saúde, são definidas cinco (5) grandes categorias de indicadores: de políticas de saúde, sociais e econômicos, da prestação de cuidados de saúde, da cobertura da atenção primária de saúde e dos estados de saúde (OPS/OMS: 1985, 19). Como vemos, esta classificação está voltada para os objetos da avaliação: os estados de saúde e as intervenções em saúde (políticas, serviços, prestação de cuidados), enquanto que a classificação anterior é referida essencialmente às intervenções, objeto primordial da avaliação em qualquer área de conhecimento.

A necessidade de utilizar vários indicadores faz com que o problema passe a ser então a definição de “critérios de seleção dos indicadores” que possibilitem “decidir quais os limites

ou número deles para poder captar as diversas partes, dimensões da matéria-objeto da pesquisa, no sentido de que a sua reconstrução seja a mais completa” (Ferrari: 1982, 103). Este é o terceiro aspecto na questão dos indicadores.

Para orientar esta seleção, os autores definem as características que deve ter um bom indicador. FERRARI entende que os indicadores devem preencher dois requisitos: “1) têm que ser facilmente determináveis; 2) devem corresponder razoavelmente a um universo de características mais amplas do que quando se utiliza o conceito” (Ferrari: 1982, 102).

Na literatura específica da área da avaliação em saúde, de uma forma geral são discriminadas quatro características que os bons indicadores devem ter: **válidos**, isto é, “devem medir eficazmente o que se supõe que vão medir”; **confiáveis** ou **objetivos**, isto é, “que o resultado deverá ser o mesmo se a medição se efetua em condições semelhantes por pessoas distintas em diferentes momentos”; **sensíveis**, isto é, “que reajam às modificações da situação” e **específicos**, isto é, “que reflitam as modificações unicamente na situação da qual se trata” (OPS/OMS: 1985, 17).

Convém lembrar que estas são características “ideais” dos indicadores (Kisil&Chaves: 1994, 102; OPS/OMS: 1985, 17), porém “na vida real, há relativamente poucos indicadores que respondam a todos esses requisitos, pelo que é preciso considerar seu valor científico com certa humildade” (Kisil&Chaves: 1994, 102).

Vemos que estas características são as mesmas que se atribuem à qualidade dos instrumentos de medida. Outros autores vão além das qualidades do indicador enquanto “instrumento de medida”, incluindo aspectos que referem-se mais especificamente ao seu uso para estudos avaliativos.

LAURENTI et al. apontam cinco (5) características dos indicadores que medem os estados de saúde: a) “que existam dados disponíveis para toda a população da área geográfica considerada”, isto é, a **cobertura** dos indicadores; b) “as definições e procedimentos empregados para construir esses indicadores devem ser os mesmos em todos os países”, isto é, a **padronização** na construção dos indicadores; c) “devem ser de construção fácil e, principalmente, simples de serem interpretados”, isto é, a **simplicidade** dos indicadores; d) “devem, na medida do possível, refletir o maior número dos fatores que, reconhecidamente, influam no estado de saúde”, isto é, a **sinteticidade** ou **abrangência** dos indicadores; e)

“devem ter um bom poder discriminatório, para permitir comparações entre áreas, ou períodos [de tempo] de uma mesma área”, isto é, a **discriminatoriedade** dos indicadores.

MORAES menciona sete critérios de qualidade dos indicadores, baseando-se em trabalho de Jordan Filho et al.⁴¹, que incluem cinco das nove anteriormente mencionados: validade, cobertura, simplicidade, sinteticidade ou discriminatoriedade. Além desta, a autora menciona ainda a “robustez” do indicador, isto é, que ele seja “pouco sensível às deficiências dos dados necessários à sua construção” e a “disponibilidade”, isto é, que deve-se “usar dados habitualmente disponíveis ou de fácil obtenção” (Moraes: 1994, 21). FERREIRA também descreve estas mesmas características, porém denominando algumas de forma diferente: sua descrição “especificidade” corresponde ao que os demais autores denominam de “validade”, da mesma forma que ele denomina de “confiabilidade” o que MORAES refere como “robustez”. Além disto, o autor descreve o conceito de “sensibilidade” de forma muito semelhante ao de “discriminatoriedade”(Ferreira: 1990; 44-45). Nota-se que o conceito de “confiabilidade” ou “robustez” do indicador, conforme estes autores, significa de certa forma o contrário do que foi definido no documento da OPS, e que é o consenso sobre o termo, pois o indicador é confiável quando não varia muito conforme a boa ou má qualidade dos dados. O Anexo XVIII contém a descrição destas características, permitindo comparar as descrições destes dois autores.

O quarto aspecto que destacamos em relação aos indicadores é que talvez o maior viés dos estudos de avaliação seja a seleção do indicador mais fácil. WEILENMANN coloca que nas investigações em avaliação existe uma tendência para selecionar aqueles indicadores que podem “ser medidos com mais precisão e manipulados mais convenientemente nas análises estatísticas - em detrimento de outros fatores que podem estar sendo mais significativos na situação problema” (Weilenmann: 1980, 26). EATON menciona o perigo de limitar os esforços de pesquisa aos aspectos cujos dados estão mais disponíveis, ou que podem ser coletados mais rapidamente, mesmo que não forneçam as informações mais importantes para a avaliação (Eaton: 1993). Para ilustrar este ponto, alguns autores mencionam a parábola do

⁴¹JORDAN FILHO, L. et al. Revisão analítica dos indicadores de saúde. *Revista Brasileira de Estatística*, 35: 22-6, jul./set. 1974.

bêbado⁴², que tendo perdido as chaves do carro à noite, procura-as perto de um poste de luz, não porque tenha mais probabilidade de encontrá-las aí, mas apenas porque está mais iluminado... (Eaton: 1993; Novaes&Paganini: 1994).

Quanto ao segundo conceito desta tríade, o de **critério**, observamos que é utilizado ora como um conceito que está acima (ou antecede) o conceito de indicador, ora como uma opção alternativa ao uso deste termo, ora como um desdobramento do conceito de indicador em um nível mais operacional.

No primeiro caso, vemos a referência à utilização de “critérios” para selecionar os indicadores que irão operacionalizar um conceito, ou seja, a noção de critério antecede a de indicador (OPS/OMS: 1985, 16; Ferrari: 1982, 103).

No segundo caso, vemos que os critérios são considerados uma opção alternativa ao uso dos indicadores:

“Quando não se dispõe de indicadores adequados para determinada situação, apela-se para critérios, que são normas com as quais se medem distintas ações. Podem ser técnicos e sociais, sendo os primeiros, em geral, muito específicos de cada programa” (Kisil&Chaves: 1994, 102).

Observamos que o autor define “critério” utilizando o conceito de “norma”, fato que também observamos em publicação anterior: “os critérios são as normas às quais se recorre para medir ações e que tendem a suscitar uma reflexão para emitir um juízo” (OPS/OMS: 1985, 16). Neste mesmo sentido, KISIL&CHAVES entendem que o principal objetivo dos critérios é “induzir à interpretação dos **valores** por meio da **reflexão**, como base para os pareceres e informes necessários”, pois seria “improvável” a elaboração de pareceres e informes apenas através da “simples soma de valores numéricos resultantes de critérios de avaliação quantitativos [grifos nossos]” (Kisil&Chaves: 1994, 102).

No terceiro caso, isto é, quando o critério é tomado como um desdobramento do indicador, vemos que o critério é o valor que serve de comparação para o valor encontrado, permitindo dar sentido ao mesmo. WEISS coloca que “diferentes pessoas olhando o mesmo

⁴²Cf. NOVAES&PAGANINI, a autoria é de VLADECK, B.C. Quality Assurance through External Controls. In: *Quality Assurance in Hospitals*. GRAHAM, Nancy O. Ed. Aspen Public., Rockville, Maryland, 1990.

dado chegam a diferentes conclusões”. Ela menciona a tradição de utilizar o binômio interpretativo “completamente-somente”: quem está tentando promover o programa apregoa que “completamente 25%” alcançou tal meta, enquanto quem está tentando diminuir o programa apregoa que “somente 25%” alcançou a meta. A pergunta que se coloca é a seguinte: “quanto do progresso em relação à meta demarca o sucesso?”. Para evitar disputas de interpretação, é necessário estabelecer os “padrões de julgamento”, ou seja, os critérios (“yardsticks”) (Weiss: 1972, 32).

Ou seja, os critérios são a base de comparação que permitem, racionalmente, emitir o juízo de valor (OPS/OMS: 1985). Ou seja, para avaliar, é necessário identificar e especificar critérios com os quais os efeitos de uma intervenção poderão ser “emparelhados e comparados” (Weilenmann: 1980, 12).

Há uma distinção entre critérios “técnicos” e critérios “sociais” (OPS/OMS: 1985, 16). Embora sendo recomendado que os critérios devam ser quantificados, há o reconhecimento de que na prática isto nem sempre é possível, sobretudo no caso dos critérios sociais, que seriam “difíceis de cifrar” (OPS/OMS: 1985, 16; Kisil&Chaves: 1994, 102). Nestes casos, o avaliador deveria recorrer à “determinações qualitativas” (Kisil&Chaves: 1994, 102).

É fundamental também lembrar que, com respeito à certas atividades no campo da saúde, “não existem indicadores, nem critérios” para a avaliação, devendo-se então fazer algumas “perguntas pertinentes” para conduzir a mesma (OPS/OMS: 1985, 23; Kisil&Chaves: 1994, 102).

Por fim, gostaríamos de destacar o caráter dialético e histórico dos critérios, quando considera-se que o objetivo das intervenções é a mudança social. Para WEILENMAN, a “natureza e as propriedades” dos critérios para avaliar estas mudanças “não estão necessariamente dadas, e talvez nunca sejam completamente especificadas. Ao invés disto, eles estão sujeitos à mudanças contínuas, de tal forma que a imagem do processo dialético torna-se relevante: tensão e distensão, conflito e pressão permeiam a situação” (Weilenmann: 1980, 12). Especialmente no caso das intervenções sobre problemas mal-estruturados, os critérios de “significância e desejabilidade não estão dados conforme um sentido objetivo”, mas seriam o resultado da “interação entre diferentes perspectivas, [...] sendo passíveis de

mudanças ao longo do tempo, dinamicamente”. Por isto, seria necessário conhecer, em cada momento, “como os critérios foram construídos” (Weilenmann: 1980, 37).

Em relação ao conceito de **normas** ou **padrões** existe maior consenso, sendo entendido de forma geral como “as exigências técnicas que devem ser cumpridas para atingir o objetivo” (Cohen&Franco: 1994, 90). Apesar da objetividade que o uso do termo “técnicas” pretende indicar, o processo de definição da norma ou padrão, que corresponderia à fase mais operacional, é, contraditoriamente, o mais subjetivo.

Alguns autores entendem que os critérios e as normas nos quais se apóiam as avaliações constituem “resumos do verdadeiro e resumos do bem”, de forma que supõe-se que o respeito aos mesmos levaria o programa ou intervenção a ter um efeito positivo (Contandriopoulos et al.: 1992). Da mesma forma, PITTA coloca que a utilização de “critérios de determinação normativa” corresponderia a comparar a intervenção com “um ideal desejado” (Pitta: 1992, 45).

Segundo CLEMENHAGEN & CHAMPAGNE, os padrões (“standars”) podem ser definidos através de “observação empírica ou de julgamento informado (consenso)”. (Clemenhagen & Champagne: 1986, 384).

Para DUSSAULT&HARTZ⁴³, a norma é o “valor do critério que é considerado aceitável, o qual pode variar muito, dependendo do contexto”. Para que a variabilidade contextual da norma permita a comparação, utiliza-se um padrão definido como “padrão-ouro”.

Estes autores formulam um exemplo que facilita a compreensão destes três conceitos:

INDICADOR DE QUALIDADE DE SERVIÇO:	Acessibilidade
CRITÉRIO:	Nº médicos/habitante
NORMA OU PADRÃO:	EUA = 1/300
	Canadá/Uruguai= 1/400
	OMS = 1/1.000
	Certos países = 1/10.000

⁴³No curso já referido.

Observa-se que eles estabelecem uma hierarquia entre os três conceitos, partindo do nível mais geral para níveis mais operacionais. Embora apontem que os conceitos de indicador e critério sejam muito semelhantes, estes professores estabelecem as seguintes definições: o indicador seria a variável que está sendo medida para operacionalizar uma dimensão conceitual.; os critérios seriam as características observáveis do fenômeno que é avaliado; a norma ou padrão seria o valor do critério que é considerado aceitável.

No exemplo acima, observa-se claramente que o conceito de “critério” corresponde ao conceito de “índice”, que seria a quarta e última fase descrita por Lazarsfeld no processo de operacionalização de conceitos. Isto confirma-se na afirmação de GALTUNG⁴⁴, de que “a construção de um índice é semelhante à construção de parâmetros: reduzir a variedade tanto quanto seja possível sem perder a informação essencial (Johan Galtung apud Ferrari: 1982, 104).

VERGARA aponta três passos na construção de indicadores, que evidenciam a semelhança entre o processo de operacionalização de conceitos na pesquisa e de construção de indicadores para a avaliação: o primeiro é a própria definição dos indicadores que servirão para medir o cumprimento das metas; o segundo é a determinação da forma como se fará o cálculo para medi-los, com determinação da “escala de valores ou categorias” que permitem atribuir pontos ao desempenho dos mesmos; e o terceiro comparar estes pontos com o melhor desempenho esperado das metas (Vergara: 1980, 118- 126).

O estabelecimento de normas ou padrões de comparação (“standards”) está vinculado, segundo BERWICK, à busca de qualidade “por inspeção”, ou de uma “abordagem disciplinar”, que implica na definição de “limiares mínimos” de estrutura, processos e resultados. Segundo ele, tem sido demonstrado que nesta abordagem estes limiares mínimos rapidamente se transformam em “teto”, cujo alcance passa ser o esforço máximo das organizações (Berwick: 1989, 53-56).

Em resumo, vimos que a definição de indicadores, critérios e normas ou padrões corresponde ao nível mais operacional da avaliação, havendo uma certa confusão no significado destes termos, especialmente entre os conceitos de indicador e critério, e entre o conceito de indicador e o de variável. Os indicadores devem ser quantificados sempre que

possível, porém sempre estarão apenas “indicando” um aspecto parcial do fenômeno estudado, não podendo ser tomados como uma medida direta do mesmo. Por esta razão, é necessário lançar mão de diversos indicadores para tentar abarcar diferentes aspectos do fenômeno que está sendo avaliado, sendo necessário fazer uma seleção dos indicadores possíveis ou disponíveis, considerando as seguintes características: validade, confiabilidade ou objetividade, sensibilidade, especificidade, simplicidade, cobertura, abrangência, discriminabilidade, robustez, padronização e disponibilidade. É justamente em função de algumas destas características dos indicadores, como a disponibilidade e a simplicidade do uso, que ocorre um dos vieses mais comuns na avaliação, que é a escolha do indicador mais disponível ou mais fácil de usar, em detrimento daquele que expressaria melhor o fenômeno que está sendo avaliado.

1.3.3.4. O conceito de qualidade

Se os conceitos anteriores apresentam dificuldades para serem compreendidos nas nuances de seus significados, o conceito de qualidade é ainda mais difícil de ser caracterizado, apesar de ser considerado por alguns autores como o “aspecto central” do campo da avaliação em saúde (Silva&Formigli: 1994, 81).

Muitos aspectos do conceito de qualidade foram vistos nas linhas metodológicas da avaliação, portanto, aqui trataremos dos aspectos mais gerais do mesmo.

No campo da Filosofia, a “Qualidade” é uma categoria do pensamento, assim como a “Quantidade”, o “tempo”, o “espaço” (Nogueira: 1994, 27-28).

Para o pensamento aristotélico, a qualidade é definida como um “atributo” ou uma “propriedade” das coisas, que poderia ser “de qualquer natureza, boa ou má”. Entretanto, através de uma “transição sutil da linguagem”, o significado do conceito foi adquirindo uma

⁴⁴GALTUNG, Johan. *Teoria y Métodos de la Investigación Social*. Buenos Aires, EUDEBA, 1966.

conotação exclusivamente positiva, como o “conjunto de propriedades adequados a alguém”. Ou seja, a qualidade deixa de ser então uma categoria geral referida aos atributos (bons ou maus) das coisas, e passa a ser considerado como “um juízo valorativo e positivo”. A noção de qualidade passa a ser então algo **positivo e desejável**, tal como idéias valorativas similares de bom, de belo, de útil (Nogueira: 1994, 28-29).

Segundo NOGUEIRA, a Qualidade pode ser entendida como uma “síntese final” de vários “juízos parciais de valor” sobre **cada um** dos atributos das coisas ou produtos. Ao emitir juízo sobre a qualidade de alguma coisa, estamos discriminando “naturalmente” as qualidades (ou atributos) que reputamos “boas” das “más”. Geralmente, ocultamos esta “cadeia prévia de juízos parciais” dos diferentes atributos, expressando apenas a apreciação final sobre o produto: é bom ou ruim (Nogueira: 1994, 28). Ouçamos o autor:

“Na raiz de qualquer avaliação de Qualidade está uma multiplicidade de juízos sobre cada uma das propriedades ou Qualidades de um produto consideradas individualmente ou na relação entre si. Mas o que transparece no final do processo é muito mais a sentença, como uma espécie de veredicto. Qualidade é, portanto, a síntese dos juízos sobre as propriedades de um produto” (Nogueira: 1994, 29).

Ao definir-se a qualidade como um “**juízo sintético**”, coloca-se a mesma no domínio da faculdade humana da “opinião”, ou seja, no domínio da *doxa*, que filosoficamente se opõe ao domínio da *epistéme*, ou conhecimento (Nogueira: 1994, 33). O autor explicita estes dois conceitos foram entendidos, originalmente, no idealismo platônico:

*“Na opinião, a ilação é resultado de um julgamento, sua verdade é presumível, enquanto no conhecimento a ilação corresponde a uma verdade logicamente evidente e necessária. Desde logo, essa dualidade remete a análise a outros pares de categorias estabelecidas pela filosofia: indução **versus** dedução; experiência **versus** razão; crença **versus** ciência; etc”* (Nogueira: 1994, 33).

Ou seja, a faculdade de opinar está permeada “pela dúvida e pela crença”, enquanto que a faculdade de conhecer “necessariamente admite a certeza como seu fundamento”. Conforme este entendimento, a opinião acaba sendo tida “como um conhecimento falso ou de pouca validade”, que se opõe à ciência, entendida como o “único modelo legítimo de conhecimento” (Nogueira: 1994, 32-33).

Embora o senso-comum reproduza exatamente o sentido dado pela Filosofia a estes conceitos, o autor discorda deste entendimento por duas razões.

A primeira, porque levando-se em conta o critério através do qual a filosofia fez a distinção inicial entre *doxa* (opinião) e *epistéme* (conhecimento), as “modernas ciências da natureza” estariam mais propriamente colocadas no domínio da opinião, e não no do conhecimento, pois “fundam-se em formas sistemáticas e socialmente sancionadas de estabelecer juízos, por outras palavras, são **essencialmente indutivas** [grifo nosso]”, enquanto para o idealismo platônico, o verdadeiro conhecimento seria somente aquele “que resulta do **movimento dedutivo da razão** na esfera das quantidades [grifo nosso]”, cujo modelo seriam as ciências essencialmente abstratas, como a matemática e a geometria. As demais disciplinas dependeriam “de uma forma de julgamento sobre atributos da realidade física, que é sempre mutante”, portanto, em sentido estrito, seriam consideradas “da mesma natureza da opinião” (Nogueira: 1994, 33).

A segunda razão, porque, atualmente, a opinião das pessoas, inclusive a denominada “opinião pública”, cada vez mais busca elementos das ciências para fundamentar-se, não podendo mais ser vista como fruto de uma “crença sem arrazoados ou sem evidências”. Portanto, ao tratar da questão da qualidade seria necessário relativizar a contraposição entre “conhecimento científico e opinião” (Nogueira: 1994, 33).

Em suma, o autor destaca no conceito de qualidade a noção de juízo, o que o coloca na esfera da opinião, embora cada vez mais fundamentada em elementos da ciência. O outro aspecto fundamental do conceito de qualidade destacado por NOGUEIRA é sua vinculação com a noção de “valor de uso”, oriunda do campo da economia. Segundo o autor:

*“O conjunto das Qualidades das coisas, consideradas em sua capacidade de satisfazer as necessidades humanas, é o que a economia clássica denomina de **valor de uso**”* (Nogueira: 1994, 36).

Na teoria econômica, o valor de uso refere-se à “utilidade das coisas”, ou seja, às “propriedades [das mercadorias] que são úteis para o consumidor” (Nogueira: 1994, 36-37). HARNECKER coloca: “chama-se valor de uso a todo objeto que responde a uma necessidade humana determinada (fisiológica ou social)” (Harnecker: 1973). Embora originalmente referido à mercadoria, entendida como um “bem material”, o conceito é aplicado também aos serviços.

Entretanto, a utilização do conceito de “valor de uso” para a prestação de serviços não é simples. Na indústria, a qualidade dos produtos pode ser garantida mediante a definição de padrões de Qualidade contidos na “especificação técnica” do produto, que admite uma faixa de variabilidade tolerável, permitindo o controle da produção em massa (Nogueira: 1994, 123). Esta padronização tem dupla via: ao mesmo tempo que padroniza-se o produto, padroniza-se “as necessidades que se pretende fazer com seu uso”. A mudança nas características do produto requer investimentos adicionais, a definição de um novo processo de produção, configurando um “produto novo” (p. 127).

A prestação de serviços, entretanto, implica em uma idéia de particularidade, oposta à de padronização, que deriva do próprio conceito de serviços: “algo que serve a uma pessoa determinada, derivado etimologicamente de *servus*, escravo” (Nogueira: 1994, 123). No setor saúde, esta “particularização” é ainda mais evidente, pois a produção de serviços “deve adaptar-se a necessidades que aparecem inesperadamente e que acabam por representar desvios tanto em termos quantitativos quanto qualitativos” (p. 127). Ou seja, a concepção do produto ocorre “no momento mesmo da produção/consumo”. Levando-se este raciocínio ao extremo, a qualidade da produção de serviços de saúde ocorre no “plano do singular: existem doentes e não doenças e o ato médico reveste-se de Qualidade apenas para **este** paciente, **neste** momento” (p. 128).

Entretanto, a busca de uma “gerenciabilidade” dos produtos e dos gastos em saúde impõe uma certa necessidade administrativa de “uniformização” (Nogueira: 1994; 129).

Neste ponto, o autor diferencia dois padrões de normatividade: a normatividade endógena ou espontânea, que “decorre dos ritmos do trabalho humano, mediado por algum tipo de método”, e a normatividade exógena ou explícita, que é “sempre de natureza técnico

administrativa [e segue] um padrão definido por agentes externos: a administração da empresa, a legislação governamental, etc”, ou seja, é uma “normatividade de tipo sistêmico” (Nogueira: 1994; 123).

O autor aponta que “a prática clínica e a cirúrgica mantém ainda uma grande margem de normatividade de tipo espontâneo”, mesmo em serviços onde já se alcançou “nível elevado de normatização”, que deriva da autonomia destas práticas, como vimos anteriormente, na “Avaliação de Qualidade” (Nogueira: 1994; 129).

Ao destacar as noções de “juízo” e de “valor” como os principais aspectos do conceito de qualidade, NOGUEIRA está apontando os aspectos fundamentais do próprio conceito de avaliação - a avaliação é um juízo de valor -, corroborando a colocação inicial de que a qualidade é o aspecto central da avaliação. Ou seja, embora pareça uma afirmação óbvia, toda avaliação é sempre da “qualidade” de alguma coisa ou ação.

Da mesma forma como fizemos em relação ao conceito de avaliação, NOGUEIRA coloca como ponto central da discussão de qualidade o debate entre a “subjetividade” e a “objetividade” do conceito, que ele conduz ao debater as diferenças entre Joseph Juran e Edward Deming, os dois mentores da “Teoria da Qualidade Total”.

Segundo este autor, a qualidade é definida por Juran como “a adequação ao uso” (Nogueira: 1994, 40). Ou seja, o consumidor de um produto ou serviço irá julgar que o mesmo tem tanto mais “qualidade” quando mais considerá-lo adequado ao seu uso. Neste sentido, NOGUEIRA entende que Juran propõe uma definição tautológica do conceito de qualidade, pois “é como medir algo usando como equivalente aquilo que se quer medir, já que todo bem é, por definição, mais ou menos adequado ao seu uso” (Nogueira: 1994, 41).

Com sua definição, Juran estaria buscando “escapar da lógica do ‘tudo ou nada’” na questão da qualidade, estabelecendo a possibilidade de mensuração da mesma conforme os “graus de adequação ao uso” (Nogueira: 1994, 23), visando em última instância estabelecer “os princípios de mensurabilidade universal da Qualidade”, através de “um juízo que estabelece relações de comparação [e] mede a extensão da utilidade do conjunto dessas propriedades [das coisas]” (p. 40).

Esta visão “quantitativista” da qualidade, aparentemente contraditória, é a mesma de certas correntes das Ciências Sociais. Ao tratar da questão das variáveis qualitativas,

FERRARI adota inicialmente o conceito filosófico, que vimos inicialmente: as variáveis “qualitativas” ou “nominais” seriam “definidas pelos seus atributos”, os quais, por sua vez, corresponderiam “àqueles aspectos não mensuráveis, não numéricos”, que não podem ser contados nem medidos. Entretanto, logo em seguida, o autor coloca que para tratar as variáveis qualitativas “de modo quantitativo”, é necessário que se atribua um “critério de qualidade” às mesmas. Neste ponto, ele cita Ackoff⁴⁵, para quem “a qualidade [...] pode ser definida simplesmente como uma extensão ao longo da escada e em termos da qual o atributo da variável pode ser apreciado” (Ferrari: 1982, 143-44).

De qualquer forma, a questão da “quantificação da qualidade” é de natureza filosófica, conforme vemos em NOGUEIRA:

“Como a Qualidade se relaciona com a quantidade - a Qualidade se expressa por quantidades distintas, sendo a quantidade o atributo primário do ser? Como a Qualidade pode variar num mesmo ser, estabelecendo as quantidades de uma Qualidade” (Nogueira: 1994, 27).

O pensamento de Edward Deming, o outro mentor da “Gestão de Qualidade Total”, vem em sentido contrário. Para ele, conceito de qualidade “não tem a menor necessidade de ser precisado e fixado rigidamente”, pois o essencial na abordagem do tema é a “centralidade do papel do consumidor”, a quem caber emitir o “juízo” da qualidade. Neste sentido, coloca-se claramente a subjetividade do conceito de qualidade na concepção de Deming, para quem “a Qualidade só pode ser decidida em termos de quem avalia”. Provavelmente em função disto, o segundo aspecto essencial na teoria de Deming é a “preocupação em problematizar pedagogicamente o sentido e o uso da própria noção de qualidade” (Nogueira: 1994, 27; 20; 29).

NOGUEIRA adota a abordagem de Deming, ressaltando a subjetividade do conceito de qualidade:

⁴⁵ACKOFF, Russell L. *Planejamento da Pesquisa Social*. São Paulo, Herder, 1967.

“Mas é preciso ser dito e repetido, ecoando as palavras de Deming, que o juiz da Qualidade é o homem. Pode parecer uma declaração muito bombástica ou até destoante para um estudo como este, mas temos de fazê-la: o homem é o supremo árbitro de suas necessidades e daquilo que pode satisfazê-las. A ciência ajuda-o a exercer esta função de arbitrar, mas não pode fornecer fórmula ou instrumento de mensuração objetiva que venha a substituí-lo em tal papel”(Nogueira: 1994, 42).

Em suma, vimos que o conceito de qualidade está centrado nas noções de “juízo”, que implica em um comportamento subjetivo, e de “valor de uso”, que refere a uma necessidade objetiva.

1.4. Os sujeitos da avaliação em saúde: quem avalia.

Cada vez mais, coloca-se no campo das ciências em geral, as questões relativas ao sujeito da investigação, ao seu contexto e suas implicações no trabalho científico. De forma especial, no campo da avaliação, “a perspectiva do avaliador é um componente fundamental que determina os rumos de qualquer trabalho de avaliação” (Silver: 1992, 198). Ou seja, à diferentes sujeitos da avaliação correspondem diferentes perspectivas do objeto avaliado.

O rumo da avaliação é muito influenciado pelo papel dos diferentes atores sociais envolvidos no processo avaliativo. Os problemas da esfera social podem ser vistos por “diferentes perspectivas”, por isto, é necessário explorar inicialmente “de quem são os objetivos que devem ser traduzidos em critérios objetivos para avaliação” (Weilenmann: 1980, 37, 30). Além disto, a execução das políticas públicas envolve diferentes atores sociais que atuam dentro das instituições - os políticos, os técnicos e os burocratas -, que possuem

diferentes racionalidades, muitas vezes conflitantes, e que precisam ser levadas em conta no processo avaliativo (Coehn&Franco: 1994, 64-67).

AGUILAR&ANDER-EGG definem quatro atores principais que devem ser levados em conta numa avaliação, com maior ou menor ênfase, dependendo do seu enfoque, cujos pontos de vista nem sempre é fácil conciliar:

- “- os responsáveis pela pesquisa avaliativa;*
- os que encomendam ou mandam fazer a avaliação;*
- os respondentes da avaliação;*
- os usuários, beneficiários ou destinatários dos programas, projetos, serviços ou atividades que são avaliados”* (Aguilar&Ander-Egg: 1994, 73; 79)

Além disto, é preciso considerar que, sendo a avaliação um “dispositivo de produção de informação”, ela é “fonte de poder” para os atores que a controlam (Contandriopoulos et al.: 1992).

De uma forma geral, observamos na literatura consultada, que a questão do sujeito da avaliação pode ser abordada por dois ângulos, conforme a perspectiva deste sujeito em relação aos objetos de avaliação, que no caso da Saúde Pública, como já vimos, são de duas ordens: os estados de saúde e as intervenções em saúde. Tanto em relação a um como ao outro objeto, os ângulos de abordagem podem ser diferenciados segundo a vinculação ou não do avaliador à instituição promotora do programa ou projeto, bem como sua participação ou não na equipe executora do mesmo (Cohen&Franco: 1994, 111; Aguilar&Ander-Egg: 1994, 44).

De nossa parte, diferenciamos dois ângulos de abordagem do objeto avaliado:

- a perspectiva do sujeito enquanto produtor ou enquanto consumidor de cuidados/serviços/programas de saúde que estão sendo avaliados. Por este critério, temos de um lado os produtores de cuidado de saúde (trabalhadores) e os administradores/gerentes de serviços ou programas, e de outro os usuários dos mesmos, bem como a população em geral;

- a perspectiva de internalidade ou de externalidade em relação à produção dos cuidados/serviços/programas que são objetos da avaliação, que determina o grau de independência do sujeito da avaliação em relação ao objeto que está sendo avaliado. Nesta segunda abordagem, temos de um lado o avaliador que está vinculado à instituição promotora

ou a equipe executora dos cuidados/serviços/programas de saúde, e do outro o avaliar cujo trabalho é desenvolvido autonomamente ao processo de produção dos mesmos.

Veremos a seguir como a questão do sujeito é colocada em relação aos dois objetos da avaliação em saúde. Quando o objeto da avaliação são os estados de saúde, as diferentes perspectivas do produtor e do consumidor de serviços perspectiva ensejam o debate entre as “verdadeiras necessidades de saúde” *versus* “demandas por serviços”. Quando o objeto da avaliação são as intervenções em saúde, a principal questão que destacamos é o avaliador interno *versus* avaliador externo.

1.4.1. O sujeito da avaliação dos estados de saúde.

Em relação à primeira grande ordem de objetos da Saúde Coletiva/Saúde Pública - a avaliação dos “estados de saúde” -, vimos anteriormente que a questão é tratada na literatura referente ao “diagnóstico de saúde” das comunidades ou populações, geralmente oriunda da área da Epidemiologia. A “avaliação das necessidades” de saúde, isto é, a identificação e quantificação dos problemas de saúde, é apenas um dos aspectos do diagnóstico, que inclui também a hierarquização e escolha daqueles que são considerados prioritários para sofrer uma intervenção, ou seja, aqueles que serão objeto de um plano de ação ou de um programa de saúde.

A pergunta que se coloca é quem é o sujeito da avaliação das necessidades de saúde?

Tradicionalmente o “diagnóstico” de saúde é entendido como uma atribuição dos técnicos, que seriam capazes de identificar as “verdadeiras” necessidades de saúde da população utilizando o instrumental científico, o que garantiria a neutralidade do mesmo (Paim: 1979; Sousa Campos: 1991). Segundo ALMEIDA FILHO&ROUQUAYROL, a Epidemiologia “desenvolve tecnologias efetivas para o chamado diagnóstico de saúde da comunidade, oferecendo subsídios para o planejamento e a organização das ações de saúde” (Almeida Filho&Rouquayrol: 1992, 7).

Paralelamente a este entendimento, existe a visão de que “a racionalidade dos políticos se fundamenta em responder às demandas e não às necessidades grupais” (Cohen&Franco: 1994, 68). Entretanto, esta visão vem sendo questionada por alguns autores, gerando um debate entre as “reais necessidades de saúde” *versus* “demandas por serviços de saúde”.

Segundo PAIM, embora se relacionem com necessidades “vitais” ou “naturais”, as chamadas “necessidades de saúde” são definidas por homens que “vivem em determinadas condições históricas”. Portanto, “nem os critérios, nem os processos, nem as técnicas de definição de necessidades de saúde são neutros”, e variam conforme o valor atribuído à vida humana em diferentes conformações sociais. Ou seja, a definição das necessidades “transcende a competência da técnica”. Além do que, mesmo entre os técnicos, não haveria consenso quanto ao “modo de identificar, quantificar, hierarquizar e selecionar as ‘necessidades de saúde’ a serem atendidas” (Paim: 1979, 43; 45).

Por esta razão, seria pertinente indagar “que grupos sociais, num dado momento histórico, estão definindo as necessidades de saúde”, ou seja, quem “têm ou não o poder de enunciá-las e de serem normativos em relação à vida” em cada contexto (Paim: 1979, 43;45)

No mesmo sentido, SOUSA CAMPOS coloca que “as necessidades consideradas essenciais à reprodução da vida são realmente social e historicamente definidas”, e que os problemas sanitários definidos como prioritários “na verdade, são escolhidos e selecionados antes da aplicação de técnicas de programação, a partir de pressões sociais, políticas ou econômicas” (Sousa Campos: 1991, 64, 59).

AGUILAR&ANDER-EGG colocam que “o debate sobre o conceito e as formas de determinar as necessidades sociais não foi resolvido de forma geralmente aceita” (Aguilar&Ander-Egg: 1994, 75). Em artigo que trata das ambigüidades da noção de “necessidade de saúde”, PAIM aponta que as expressões “necessidade de saúde” e “necessidade de serviços de saúde” são utilizadas como sinônimos, porém, a primeira implica em um entendimento mais amplo, de que “as condições necessárias para a obtenção da saúde” vão além das prestação de serviços de saúde (Paim: 1979; 40).

Referindo-se ao trabalho de JEFFER et al.⁴⁶, este autor apresenta uma interessante distinção entre demanda, necessidade e desejo. O conceito de “necessidade social” teria sido introduzido por aqueles autores, e corresponderia à “quantidade de serviços médicos julgada apropriada pelo administrador de saúde ‘bem informado’, capaz de analisar a opinião dos especialistas e a repercussão dos serviços médicos em toda a economia”. A “demanda”, por sua vez, seria “a quantidade de serviços que os membros de uma população desejam comprar para cada combinação de preços, renda, nível de saúde, gostos etc”. Por último, o “desejo” seria “a quantidade de serviços médicos que o indivíduo sente que deveria consumir fundamentada na própria percepção dos seus problemas de saúde”. No Brasil, o autor refere o uso do termo “demanda técnica” com a mesma conotação desta definição de “necessidade social”. Por outro lado, haveria uma preferência do termo “demanda” por parte dos economistas, por entenderem que o conceito de “necessidade” nega a autonomia e a individualidade dos indivíduos, enquanto a noção de demanda implica na possibilidade de escolha entre diferentes opções de serviços, conforme as preferências individuais (Jeffer et al. apud Paim: 1979; 40-42).

A tradicional prerrogativa da esfera técnica para tratar da avaliação das necessidades de saúde está bem caracterizada por BOULDING⁴⁷: “Lo que usted necessita es lo que yo, en mi caracter de consijero profesional suyo, tengo para darte; lo que usted quiere es bastante irrelevante” (Boulding apud Paim: 1979; 41)

Ao discutir a lógica da programação em saúde, SOUSA CAMPOS defende a idéia de que as necessidades de reprodução e manutenção da vida podem ser verificadas “tanto do ponto de vista técnico - pesquisas epidemiológicas - quanto do social - direitos e cidadania”. Entretanto, a chamada “demanda potencial”, isto é, “aquela definida a partir do interesse dos técnicos ou do serviço”, é considerada mais correta e legítima “por ser INDUZIDA a partir do saber epidemiológico, ou administrativo ou clínico” (Sousa Campos: 1991, 85). Ele entende que esta pretensão da esfera técnica corresponde a um “delírio positivista”, no qual se imagina “a possibilidade real de controlar, ordenar e governar a sociedade principalmente segundo os aportes da ‘ciência’” (p.71). Porém, é necessário considerar como legítima também a chamada

⁴⁶JEFFER, J.R; BOGNANO, M.F. & BARFLET, J.C. On the demand versus need for medical services and de concept of shortage. *American Journal of Public Health*, 61: 46-63, 1971.

“demanda espontânea” ou “demanda livre”, isto é, aquela que “depende, em larga medida, do juízo e da consciência de cada indivíduo” (p.85), pois embora inclua “desejos que não teriam ‘legitimidade técnica’”, esta demanda reflete e é consequência da “relação dinâmica entre sofrimento, subjetividade, doença, morte e vontade de viver”, além do que “traduz as aspirações da sociedade por novos padrões de direitos sociais” (pp. 64; 70; 73).

Para o autor, a “‘legítimas’ necessidades de saúde” seriam sempre “socialmente definidas”, portanto estariam além do escopo de um diagnóstico técnico baseado no conhecimento epidemiológico. Além disto, a escolha de prioridades “sempre resultam de políticas que conseguem impor-se em um dado período histórico, e que só o livre jogo democrático conseguiria equacionar de maneira mais equânime”. Neste jogo, a expressão do “conflito entre vários discursos técnicos, entre interesses de distintas classes sociais” estaria garantida pela “democratização da gestão do Estado” (Sousa Campos: 1991, 71; 73).

Nesta mesma linha, o Prof. Gilles DUSSAULT⁴⁸ pergunta: “o que é necessidade: a percepção dos indivíduos (usuários) ou a percepção dos profissionais de saúde?”. Geralmente os profissionais dominam o processo de definição de necessidades, ponto de partida das intervenções em saúde, havendo inclusive diferentes definições conforme a área de atuação dos mesmos. O corporativismo dos profissionais de saúde seria um dos principais problemas do sistema de saúde canadense, identificados pelo diagnóstico da Comissão de Saúde do Canadá, em 1992. Sem controle externo, os profissionais definem os “seus” objetivos de saúde como objetivos de todos, baseados na premissa de que “nós sabemos o que deve ser feito”, e adotando uma ótica interna referida aos “objetivos de serviço”, que nem sempre seriam compatíveis com os “objetivos de saúde”, definidos a partir de uma lógica populacional.

Como vemos, a definição das “necessidades de saúde” não é neutra, transcende a competência da esfera técnica, e está relacionada ao “valor” atribuído à vida humana em diferentes conformações sociais. Além do que, não haveria consenso técnico sobre o modo de realizar o diagnóstico destas necessidades no campo da Saúde Pública.

Obviamente, esta discussão sobre os diferentes sujeitos da avaliação das necessidades em saúde tem como decorrência natural a questão da subjetividade.

⁴⁷BOULDING, K.E. El concepto de necesidad de servicios de salud. *Traducciones* 9: 1-45, C.L.A.M./OPS, Buenos Aires, 1973.

PAIM refere-se explicitamente à “subjetividade dos ‘atores sociais’” no processo de definição de necessidades de saúde (Paim: 1979, 40;42). Ao tratar das diferenças entre os conceitos de necessidade, demanda e desejo, JEFFER et al. estão lidando diretamente com esta questão (Jeffer et al. apud Paim: 1979, 40-42).

BATTISTA et al. também abordam a subjetividade do processo de avaliação dos estados de saúde (*health states*), quando referem a possibilidade de atribuir valores de utilidade ou valores de uso (*utility values*) aos diferentes aspectos nos quais a saúde se manifesta.

Ouçamos os autores:

“Os estados de saúde podem ser descritos em termos de funcionamento físico, mobilidade, principais atividades, estados psicológicos, níveis de angústia e outras dimensões. É possível atribuir valor de uso a cada um deles pelos profissionais de saúde, pacientes, membros da família, ou indivíduos da população em geral. Métodos econômicos [...] tem sido desenvolvidos para atribuir o valor de utilidade para os estados de saúde. [...] As questões mais controversas originária desta abordagem são: quem deveria atribuir os valores de uso (profissionais, pacientes, suas famílias ou parentes, ou o público em geral)? E que tipo de decisões políticas deveriam ser baseadas neste tipo de análise?” (Battista et al.: 1989, 1159).

SOUSA CAMPOS defende que no planejamento e na gestão dos serviços públicos deve-se também “trabalhar com a questão do desejo, das aspirações e necessidades expressas de forma individual através da busca ativa por atenção”, incorporando “aspectos originários dos tradicionais mecanismos de expressão de vontades subjetivas”, valorizando assim a “subjetividade dos cidadãos”. Para ele, não seria ético reprimir esta dimensão subjetiva, ao contrário: “Ético, seria estabelecer um diálogo entre estas aspirações, os discursos técnicos e as políticas de saúde, definindo-se depois os recortes e as prioridades para um período qualquer” (Sousa Campos: 1991, 74; 76). Para o autor

⁴⁸Prof. Gilles Dussault, no curso “Epidemiologia na Avaliação dos Programas de Saúde”, durante o III Congresso de Epidemiologia da ABRASCO, realizado de 24 a 28 de abril em Salvador, Bahia.

“Imaginar um Sistema Nacional de Saúde assentado principalmente no Planejamento e em ‘diagnóstico’ e ‘normas’ oriundas da Epidemiologia, mas que ignorem o sofrimento e a subjetividade, seria construir um modelo burocrático, desumano, incapaz de captar as necessidades de saúde [...]” (Sousa Campos: 1991, 70).

Como vemos, o “diagnóstico de saúde”, ou seja, a identificação e priorização das “necessidades de saúde” remete à questão dos valores dos diferentes grupos que “tem ou não poder para enunciá-las”, como coloca Paim. De qualquer forma, é a partir deste diagnóstico que se definem os objetivos das intervenções em saúde, que são a segunda ordem de objetos da avaliação em Saúde Pública/Saúde Coletiva.

1.4.2. O sujeito da avaliação das intervenções em saúde.

No que diz respeito à segunda grande ordem de objetos da Saúde Coletiva/Saúde Pública - a avaliação das “intervenções em saúde” -, a questão da perspectiva do sujeito da avaliação apresenta-se de forma mais evidente.

Voltando aos dois ângulos de abordagem que vimos no início, vemos que no primeiro caso - a perspectiva do sujeito enquanto produtor ou consumidor do cuidado/serviço/programa de saúde que é objeto da avaliação -, temos, de um lado, os prestadores de cuidados (trabalhadores de saúde) e os administradores/gerentes de serviços e programas; do outro lado, a população usuária ou potencialmente beneficiária dos mesmos.

Referindo-se à avaliação de serviços de saúde, SILVER coloca que a avaliação pode ser feita pela comunidade - que vai estar primordialmente interessada em saber até que ponto os serviços atendem às suas necessidades -, ou pelo prestador de serviços - cuja perspectiva de avaliação vai depender do modelo assistencial adotado (Silver: 1992, 198).

No mesmo sentido, porém referindo-se à avaliação de programas, WEILENMANN aponta a necessidade de considerar que provavelmente mais de um tipo de pessoa estará

interessado nos resultados do programa, ou seja, existem diferentes grupos potencialmente beneficiários dos programas sociais, e cada um deles terá sua estrutura de valores e critérios de avaliação. Conseqüentemente, torna-se necessário incorporar “diferentes perspectivas” no processo avaliativo (Weilenmann: 1980, 37).

PITTA desenvolve uma interessante abordagem sobre os interesses envolvidos na avaliação de qualidade dos serviços de saúde, onde identifica três sujeitos-beneficiários do processo. A autora aborda estes interesses com muita clareza, adotando uma perspectiva integradora, conforme vemos no trecho que transcrevemos a seguir:

*“Para uma gestão pública dos serviços de saúde que busque a progressiva melhoria de qualidade das ações prestadas, tais ‘processos de avaliação de qualidade’ podem atender pelo menos três ordens de interesse: 1) dos **gerentes** de tais serviços que, ao identificar desvios e suas causas, possam introduzir medidas modificadoras tendo a cumplicidade dos seus agentes, eles próprios sujeitos da avaliação: 2) os **trabalhadores** de saúde que podem examinar os seus desempenhos na relação com os consumidores dos seus serviços, redirecionando suas alianças e esforços numa espiral de melhor qualificação da atenção prestada; 3) o desafio de fazer dos **usuários** das instituições sujeitos e parceiros nos processos de avaliação, devolvendo-lhes instrumentos apreendidos com as experiências por eles vivenciadas, transformando-os em cidadãos informados sobre os mais adequados indicadores para pleitear o que mais lhes interesse [grifos nossos]” (Pitta: 1992, 61).*

Na perspectiva mais ampla da avaliação de programas, CONTANDRIOPOULOS et al. desenvolvem um esquema onde identificam sete “atores” implicados em uma intervenção de saúde, os quais podemos tomar como sujeitos potencialmente interessados nos resultados desta intervenção, e conseqüentemente na sua avaliação. Além dos três sujeitos mencionados acima - os “administradores” (gerentes), os “profissionais” (trabalhadores) e os “usuários” - , estes autores apontam como parte interessada os “financiadores”, a “população” (em geral), o “poder público” e, finalmente, os próprios “avaliadores” (Contandriopoulos et. al: 1992).

A utilização da “satisfação do usuário” como indicador de qualidade, tanto para cuidados prestados em nível individual como populacional, é uma tentativa de incorporar a

perspectiva do consumidor (ou usuário) nos processos de avaliação conduzidos pelo pessoal dos serviços de saúde. Entretanto, a forma mais completa para “minimizar a distância que existe entre o avaliador e os beneficiários” é a avaliação participativa. Este tipo de avaliação, na qual a comunidade beneficiária dos serviços ou programas participa do processo avaliativo, é mais adequado para projetos ou intervenções de pequeno porte, onde busca-se “fixar as mudanças propostas criando condições para que seja gerada uma resposta endógena do grupo” (Cohen&Franco: 1994, 114).

Vimos até aqui diferentes perspectivas de avaliação que tem como eixo de referência o fato do avaliador ser ou não agente da intervenção avaliada. Entretanto, mesmo quando se adota uma ótica interna às organizações que conduzem uma intervenção, é possível identificar diferentes perspectivas. Referindo-se a avaliação de projetos sociais, COHEN&FRANCO identificam três “atores sociais” que atuam com diferentes racionalidades, em princípio complementares, mas que “em certas circunstâncias podem ser conflitivas” (Cohen&Franco: 1994, 63):

*“...os **políticos**, que são os responsáveis por tomar as decisões e fixar os grandes objetivos das políticas; os **burocratas**, cuja racionalidade é baseada nos procedimentos, na aplicação de normas e na competência legal; e os **técnicos**, que são orientados pela racionalidade de fins [grifos nossos]”* (Cohen&Franco: 1994, 64).

Para estes autores, cada uma destas racionalidades apresenta âmbitos próprios de atuação: o problema é definir os limites dos mesmos, que quando não são respeitados tendem a produzir frustrações pessoais. De uma forma geral, entende-se que a formulação da “imagem ideal da sociedade”, bem como a definição das metas e dos fins para alcançá-las pertence “àqueles que detém o poder”. O conflito surge, geralmente, quando os técnicos pretendem “invadir espaços próprios do político”, que se observa especialmente nas discussões do “papel político dos planejadores”, que segundo os autores, limitar-se-ia à racionalidade técnica de “conectar adequadamente os meios aos fins”, buscando a máxima eficácia e eficiência na “consecução dos fins politicamente estabelecidos”, não lhes cabendo tomar decisões sobre os mesmos. Por outro lado, em algumas ocasiões “falta o substrato técnico e as decisões são tomadas sem base suficiente para que possam alcançar resultados eficazes”. Segundo os

autores, a maioria dos que decidem “ainda não têm consciência da necessidade de saber o que ocorreu e o que está ocorrendo no presente com as soluções de problemas sociais e, portanto, não sentem necessidade de avaliações”. Citando ARIDA⁴⁹, os autores entendem que “as decisões últimas da sociedade são de caráter político; mas a preparação de qualquer decisão tem que ser técnica”. Os autores endossam a posição da Escola de Frankfurt, de que “é preciso esticar a racionalidade técnica [...] ao seu limite” (Cohen&Franco: 1994, 66-67; 71).

Na segunda abordagem - a perspectiva de internalidade ou externalidade de quem avalia em relação à produção do cuidado/serviço/programa que é objeto da avaliação - coloca-se a questão do avaliador interno *versus* avaliador externo.

A avaliação interna seria aquela “realizada dentro da organização do projeto”, enquanto avaliação externa seria aquela “realizada por pessoas alheias à organização-agente” (Cohen&Franco: 1994, 111-12). Observe-se que a noção de avaliação interna refere-se ao fato do avaliador pertencer à instituição executor do programa, e não necessariamente ser o responsável direto pela execução do mesmo. Neste caso, isto é, quando a avaliação é realizada pelas pessoas que estão diretamente implicadas na execução do programa, trata-se de uma “auto-avaliação” (Ander Egg: 1984, 36).

A discussão sobre as vantagens de uma ou outra perspectiva tem sido causa de debates recentes nos países onde o campo da avaliação está mais desenvolvido⁵⁰ (Contandriopoulos et al.: 1992).

A avaliação é vista frequentemente como uma atividade externa, “realizada sempre por qualquer outra pessoa que não seja aquela que está fazendo o trabalho objeto da avaliação ou a população que dela se beneficia” (Silver: 1992, 198). Provavelmente, esta tendência esteja vinculada ao surgimento da Pesquisa Avaliativa, na década de 60, que conforme veremos nos aspectos históricos, caracterizava-se como um processo *a posteriori* e *post-hoc* em relação aos programas, e que desde então vem sendo considerada a linha metodológica privilegiada no campo da avaliação.

⁴⁹ARIDA, P. “Déficit de cabeças” In: SOLNIK, A. *Os pais do Cruzado contam porque não deu certo*. São Paulo, L&PM, 1987.

⁵⁰Os autores referem-se ao debate que houve entre M.Q. PATTON e C. WEISS, no ano de 1988, em três artigos publicados no periódico “*Evaluation Practice*”, nos quais o primeiro defende as vantagens da avaliação interna, e a segunda as da avaliação externa.

Entretanto, a julgar pelos debates da Conferência de 1993 da Sociedade Canadense de Avaliação, estaria havendo um movimento de inflexão em direção à avaliação interna. Nesta conferência, foi enfatizada “a necessidade de uma associação mais estreita entre os avaliadores e os programas que estes avaliam”, no sentido de “descer à arena”, que vem manifestando-se no “fenômeno da auto-avaliação dos programas” (Eaton: 1993; Love: 1993).

O ponto central da discussão parece ter sido a possibilidade do avaliador interno manter uma postura “objetiva” em relação ao objeto de sua avaliação. EATON coloca muito bem a questão:

“Toda avaliação deve encontrar um equilíbrio entre uma indiferença e uma familiaridade em relação ao programa que é tomado como objeto de estudo. A indiferença contribui com a imparcialidade e a objetividade. A familiaridade conduz a uma melhor compreensão do modo como a linha de operação e a direção de um programa podem conduzir-se na direção do alcance dos seus objetivos. Está claro que toda a tentativa de uma ‘descida à arena’ melhora a compreensão de um programa, mas às custas de uma objetividade diminuída” (Eaton: 1993).

O autor insinua que o envolvimento do avaliador com a execução do programa pode subordinar o andamento da avaliação exclusivamente aos interesses do administradores do programa, excluindo-se os outros sujeitos com interesses no programa, como os trabalhadores e os usuários.

Entretanto, como bem colocam AGUILAR&ANDER-EGG, o papel do avaliador “nunca pode ser neutro”, visto que o processo avaliativo ocorre “num contexto ideológico, organizacional e sóciopolítico definido (Aguilar&Ander-Egg: 1994, 79).

Segundo HARTZ⁵¹, a importância do avaliador interno teria sido descartada durante muito tempo em função da “objetividade” e “neutralidade” do avaliador em relação ao objeto da avaliação. Porém, o fato fundamental a considerar é que, independentemente do avaliador ser interno ou externo, o trabalho que vai ser realizado será uma “encomenda” que deve ser atendida, o que implica em um comprometimento da independência do avaliador em relação às suas conclusões. Além do que, sua escolha como avaliador certamente não terá sido por acaso,

⁵¹No curso “Epidemiologia na Avaliação dos Programas de Saúde”, durante o III Congresso de Epidemiologia da ABRASCO, realizado de 24-28 de abril de 1995, em Salvador, Bahia.

obedecendo a determinados interesses da administração que o contrata. Por esta razão, ela coloca que antes de pensar “quem” deve avaliar - se o “avaliador interno” ou o “avaliador externo”, deve-se considerar “para quem” se destina a avaliação, e que tipo de avaliação é necessária.

De qualquer forma, tanto na perspectiva de avaliador interno como na de avaliador externo, o papel do avaliador contempla uma dualidade de “pesquisador-ator comprometido eticamente com a prática que constitui objeto de sua avaliação” (Hartz&Camacho: 1996). Neste sentido, HARTZ&CAMACHO colocam:

“Aproveitamos para realçar a importância desta ética de participação e questionamento, não como um conhecimento legal e sistemático, mas levando o profissional a se preocupar com a natureza e implicação do que ele faz enquanto ator inserido em um estudo com implicações sobre o desenvolvimento comunitário através do acesso a bens e serviços de qualidade” (Hartz&Camacho: 1996).

Em defesa da avaliação interna, LOVE coloca que a avaliação externa tende a produzir relatórios que “caducam ao mesmo tempo em que são redigidos”, daí porque ele a qualifica de “post-mortem”. A avaliação interna, ao contrário, não se coloca “ao lado da organização, mas de preferência no coração mesmo das operações diárias, à título de processo”, permitindo a comunicação dos problemas “em tempo real”, bem como a produção de recomendações mais relacionadas com o “ambiente dos gestores”, evitando-se assim as “soluções resultantes dos textos”, isto é, desvinculadas da realidade concreta das instituições (Love: 1993).

Esta visão é coerente com a visão da avaliação como uma das “funções de controle” do gerente, ou seja, como uma atividade “inerente à gestão”, caracterizando-se como um processo interno à administração dos serviços e programas de saúde (Clemenhagen & Champagne: 1986, 383-84; Silver: 1992, 198).

COHEN&FRANCO destacam que a avaliação interna “elimina as fricções próprias da avaliação externa”, oriundas da situação de externalidade do avaliador, que tendem a colocar os avaliados em uma “posição defensiva”, por sentirem que está sendo feito “um exame de sua atuação pessoal”. A avaliação interna, ao contrário, é sentida como “uma instância de reflexão sobre o realizado, de mútuo intercâmbio de experiências e, enfim, de aprendizagem”. Além disto, na avaliação interna, o avaliador tende a preocupar-se mais com o “conhecimento

substantivo” daquilo que está sendo avaliado do que com a metodologia da avaliação (Cohen&Franco: 1994, 111-12).

Porém, quando o objetivo da avaliação é verificar a eficácia da intervenção, podendo levar ao questionamento tanto dos objetivos proposto quanto dos meios para alcançá-los, a avaliação interna talvez não seja adequada, pois as organizações tendem a mostrar-se reticentes e resistentes às mudanças (Cohen&Franco: 1994, 113). Estas dificuldades ficam muito claras nas seguintes considerações:

“A avaliação e a organização podem ser termos contraditórios. A estrutura da organização implica estabilidade, enquanto o processo de avaliação sugere mudança. A organização gera compromissos, enquanto que a avaliação inculca o ceticismo. A avaliação se refere à relação entre a ação e os objetivos, enquanto que a organização relaciona suas atividades com seu programa e sua clientela. Ninguém pode dizer com certeza que possam existir organizações auto-avaliadoras, muito menos que cheguem a ser a forma predominante de administração“ (Hamilton Smith apud Cohen&Franco: 1994, 113).

A escolha dos critérios de sucesso para fazer a avaliação do programa pode criar situações de “tensão” entre as necessidades do avaliador e as do administrador responsável pelo programa. Segundo WEILENMANN, o avaliador externo, entendido como “pesquisador”, tende a enfatizar a melhora no *status* da população alvo do programa, enquanto que o administrador tende a enfatizar a regularidade e a eficiência dos serviços prestados. Estas diferentes perspectivas levam à conformação de diferentes imagens-objetivo e, conseqüentemente, a diferentes critérios de avaliação (Weilenmann: 1980, 25-26).

Nesta mesma direção, SILVER entende que os gerentes, de forma geral, tendem a enfatizar mais as questões referentes ao “processo de produção de serviços” e menos as questões relacionadas ao “impacto dos serviços na saúde”. Para ela, “nem sempre as inquietudes dos gerentes correspondem às inquietudes da comunidade, podendo levar ao esquecimento de considerações importantes” (Silver: 1992, 206).

A tensão que se cria entre o avaliador externo e o administrador do programa também está relacionada com a pertinência e a utilização dos resultados da avaliação. A preocupação dos profissionais da Pesquisa Avaliativa com o rigor científico dos critérios de avaliação pode

causar problemas para a utilização dos dados produzidos, pois a utilização dos achados da avaliação é “largamente determinada” por critérios ou assuntos administrativos (Weilenmann: 1980, 25-26).

De qualquer forma, existe a possibilidade das avaliações mistas, isto é, aquela conduzida por uma equipe de trabalho que inclui avaliadores internos e externos da instituição executora, onde combina-se elementos da avaliação interna e externa, buscando ao mesmo tempo superar as dificuldades e preservar as vantagens de cada uma (Ander Egg: 194, 35; Cohen&Franco: 1994, 114).

Para esclarecer alguns aspectos envolvidos na questão da avaliação interna ou externa, recorreremos a uma distinção entre duas áreas de prática de avaliação - a “pesquisa empresarial” e a “pesquisa avaliativa”, que teria ocorrido nos EUA a partir do trabalho de Bernstein e Freeman, em 1975⁵². Esta diferenciação, mencionada por WEILENMANN, está baseada em critérios relacionados à perspectiva do avaliador em relação ao objeto da avaliação, que por serem muito elucidativos resolvemos transcrever, embora reconhecendo que estas modalidades de prática tenham ocorrido naquele determinado momento da história do desenvolvimento da avaliação de programas.

Os critérios envolvem quatro dimensões: (a) as relações entre o patrocinador ou financiador da pesquisa e o pesquisador (ou seja, o avaliador); (b) a filiação institucional do pesquisador; (c) os valores e as abordagens do pesquisador e (d) a duração do estudo (Weilenmann: 1980, 28).

A prática da “pesquisa acadêmica”, enquanto um “modelo ideal”, teria as seguintes características: (a) uma relação frouxa entre o financiador e o investigador (ele aponta como exemplo os prêmios de pesquisa oferecidos por certas agências ou fundações); (b) o pesquisador e sua equipe tendem a estar vinculados ao departamento de uma universidade ou escola; (c) o pesquisador persegue valores de aquisição de conhecimentos, de construção de teorias e de refinamento de métodos e técnicas; (d) o estudo avaliativo tende a ser de longa duração com um orçamento relativamente baixo (p. 28).

A prática da “pesquisa empresarial”, em contraste, teria as seguintes características: (a) haveria uma relação contratual, na qual o patrocinador necessita de respostas à questões

específicas, concernentes às decisões políticas relativas a um programa específico e, por estas razões, pode manter um contato mais próximo com o pesquisador, no curso da investigação; (b) o pesquisador tende a estar vinculado a um instituto ou centro de pesquisa, que pode ou não estar baseado em uma universidade; (c) a abordagem do pesquisador é compatível com as atitudes e o comportamento burocrático, onde o estudo tende a resumir-se na aplicação estrita de métodos e técnicas disponíveis e acessíveis, faltando uma estrutura conceitual para o mesmo, ou seja, o estudo é estruturado principalmente em função do desenho administrativo solicitado; (d) estes estudos tendem a ter menos de dois anos de duração, mas envolvem orçamentos relativamente grandes (p. 28-29).

Para o autor, estas duas modalidades misturavam-se na prática. De qualquer forma, fica evidente que o principal aspecto que distingue estes dois tipos de pesquisa em avaliação é o grau de (in)dependência do pesquisador em relação à gerência ou administração da intervenção que está sendo avaliada.

O principal ponto a destacar é que as condições de execução da “pesquisa acadêmica” exigem a abordagem de “Pesquisa Avaliativa”, enquanto as condições que descrevem a “pesquisa empresarial” - aplicação de métodos e técnicas disponíveis e acessíveis, atitude burocrática - lembram em tudo a linha de abordagem que posteriormente denominou-se de “Avaliação Normativa”.

Estas considerações levam-nos a concluir que a questão do avaliador interno ou externo está intimamente relacionada com a linha metodológica que é utilizada para a avaliação de uma intervenção. A visão da avaliação como um processo interno ao próprio objeto que está sendo avaliado, corresponde em grande parte à linha metodológica da Avaliação Normativa (também denominada de avaliação administrativa). Em contrapartida, a institucionalização da “Pesquisa Avaliativa” como campo científico de investigação reforça a perspectiva de externalidade do processo avaliativo.

Neste sentido, quando o que está sendo proposto é fazer uma avaliação estratégica, isto é, aquela que remete aos próprios objetivos da intervenção, obviamente a atuação do avaliador interno ficará mais comprometida. Da mesma forma, quando o que interessa é fazer uma avaliação do impacto da intervenção, recomenda-se a Pesquisa Avaliativa, que é melhor

⁵²BERNSTEIN, I.N.; FREEMAN, H.E. *Academic and Entrepreneurial Research: The Consequences of Diversity*

conduzida por avaliadores externos, ou, pelo menos, por uma equipe mista. Em contrapartida, no caso da Avaliação Normativa, quando existem critérios e normas de avaliação bem definidos, a atuação do avaliador interno será mais tranqüila.

Dito de outra maneira, quando a finalidade da avaliação é formativa, em geral é feita pelo próprio pessoal do programa; quando a finalidade é somativa, geralmente fica ao cargo de pessoal que não está vinculado diretamente ao programa (OPS: 1985, 10).

Além disto, na escolha de quem deve compor a equipe de avaliação - se pessoal do próprio programa, se uma equipe externa, se equipe mista -, deve-se considerar os seguintes aspectos (Hartz&Dussault: 1995⁵³):

- a necessidade de reforçar uma cultura de avaliação no serviço, que desenvolva a idéia de avaliação como instrumento de trabalho e não como controle externo e punitivo, garantindo a continuidade do processo avaliativo;

- a necessidade de levar em conta que a finalidade de toda avaliação é ser utilizada para uma tomada de decisão, que talvez signifique mudança. Logo, quem não participou do processo avaliativo ou não foi consultado, estará menos comprometido com as decisões que vierem a ser tomadas, e com as mudanças que vierem a ser recomendadas. A pessoa que não foi consultada sente que a avaliação é dos outros, e não dela. Por isto, quando existe um processo de participação ou de consulta ampla na avaliação, embora o processo seja mais demorado, são maiores as chances de que influencie nas decisões e de que as recomendações sejam aproveitadas;

A opção entre a avaliação interna e externa deveria ainda levar em conta o problema dos custos da avaliação. Segundo LOVE, o custo médio de uma avaliação interna é de US\$ 6000, enquanto o uso de uma consultoria externa pode chegar a US\$ 65000, ou seja, dez vezes mais (Love: 1993).

Por fim caberia colocar que, caso não se pretenda que a avaliação seja um trabalho meramente acadêmico, será preciso desenvolvê-la “de forma que possibilite modificar o comportamento dos diversos atores e aprimorar o funcionamento do sistema de saúde” (Silver: 1992, 200). Para isto, como aponta GADOTTI, é fundamental valorizar no processo avaliativo

in Federal Evaluation Studies. New York, Russel Sage Foundation, 1975.

⁵³No curso “Epidemiologia na Avaliação dos Programas de Saúde”, durante o III Congresso de Epidemiologia da ABRASCO, que referimos anteriormente.

“os critérios de representatividade, de legitimidade, de participação da base, de planejamento participativo, de convivência, de identidade ideológica, de consciência política, de solidariedade comunitária, de capacidade crítica e autocrítica, de autogestão e de outros elementos que, em última instância, serviriam para desenvolver a cidadania” (Demo: 1991, Prefácio de Moacir Gadotti).

É nesta perspectiva que coloca-se a chamada **avaliação emancipatória**, na qual “avaliador e avaliado buscam e sofrem uma mudança qualitativa”, ao invés de constituir-se como um “exercício autoritário do poder de julgar” (Demo: 1991, Prefácio de Moacir Gadotti).

Em resumo, vimos que a perspectiva do sujeito da avaliação - produtor ou consumidor do cuidado/serviços/programa avaliado; externo ou interno à execução/prestação destes cuidados/serviços/programas - influi de modo direto na forma como é direcionado o processo avaliativo e no uso que vai ser dado aos seus achados. Ou seja, a forma e a direção do processo avaliativo está relacionada diretamente com os interesses do sujeito da avaliação e à sua subjetividade, que determinam os valores que nortearão o processo avaliativo.

A questão do sujeito da avaliação coloca em relevo os conflitos de interesse que estão por trás dos objetos da avaliação, tanto no caso da avaliação dos estados de saúde quanto das intervenções em saúde. Estes conflitos ocorrem tanto na definição dos objetivos ou prioridades de ação, como na definição dos critérios de avaliação destas ações intervenções. As principais questões que se colocam são: quem define as necessidades? quem define os objetivos? quem avalia os resultados? Por trás destas questões, estão alguns dilemas que tem gerado intensos debates no campo da Saúde Coletiva/Saúde Pública, como a questão das “reais necessidades de saúde da população” *versus* “demandas por serviços”, e a questão da participação dos usuários e da população em geral na avaliação dos serviços de saúde.

2. A TEMÁTICA DA AVALIAÇÃO NO CAMPO DA SAÚDE COLETIVA: ASPECTOS HISTÓRICOS

2.1. A avaliação na área das Ciências Sociais em Saúde: análise das políticas públicas e dos programas sociais

Há uma relação histórica entre o desenvolvimento da avaliação enquanto campo de investigação científica e o desenvolvimento das políticas públicas de saúde, educação e bem-estar social que surgiram nos Estados Unidos, nas décadas de 50 e 60, com o objetivo de reduzir a pobreza, promover o desenvolvimento de comunidades, prevenir a delinquência e o crime, controlar doenças etc (Suchman: 1967, vii, Prefácio de Leonard S. Cottrell Jr; 12, 31; Weilenmann: 1980, 24). Neste contexto surgiu o conceito de “avaliação dos programas públicos”, que de certa forma é “corolário” do papel que o Estado começou a desempenhar na área social (Contrandriopoulos et al.: 1992).

SUCHMAN chega a afirmar que o desenvolvimento da pesquisa em avaliação é paralela ao próprio movimento do serviço público em si mesmo, pois o campo privilegiado de aplicação da avaliação são os serviços públicos de saúde, educação e bem-estar social (Suchman: 1967, 12, 31).

COTTRELL JR. também partilha desta opinião, argumentando que diversas áreas típicas da atuação governamental - saúde pública, educação, comportamento do consumidor, comunicação de massas e opinião pública, treinamento e performance militar, relações étnicas e raciais, saúde mental e desenvolvimento de comunidade - ofereciam “extensivas oportunidades para aplicar a teoria e os métodos de pesquisa das ciências sociais à problemas práticos, muitos deles sendo problemas de determinar como e em que extensão as intervenções planejadas produzem os resultados esperados - e os não esperados” (Suchman: 1967, vii, Prefácio de Leonard S. Cottrell Jr.).

O crescimento do campo da avaliação nas décadas de 50-60, nos EUA, pode ser atribuído a duas ordens de razão: a primeira, refere-se à atuação do Estado norte-americano

diante das cobranças da sociedade em relação às suas políticas públicas, e tem caráter externo ao campo disciplinar das Ciências Sociais, embora influenciando sobre seu desenvolvimento teórico e prático; a segunda, refere-se à incorporação das metodologias quantitativas e do modelo experimental como modelo de cientificidade, sendo de caráter interno às Ciências Sociais, ambas confluindo para o surgimento da chamada “Pesquisa Avaliativa”.

Segundo SUCHMAN, no início da expansão do serviço público as “necessidades óbvias de saúde, educação e bem-estar” não requeriam que os programas provassem sua efetividade. Porém, a medida que foram expandindo-se estes serviços, ocorreram mudanças nas “necessidades e expectativas da população”, que por estar melhor educada e mais sofisticada, passa a demandar “prova científica” da efetividade dos vários programas de governo (Suchman: 1967, 5, 13, 14).

Além disto, nesta época, muitos dos programas sociais tinham caráter experimental, e necessitavam ser avaliados. O governo norte-americano precisava demonstrar ao contribuinte, aos administradores locais e aos legisladores que estes programas funcionavam. Em função disto, acadêmicos e praticantes do campo da educação e do serviço social dedicaram-se à tarefa de demonstrar o sucesso dos programas públicos, geralmente, ao final destes ou quando sua renovação precisava ser justificada.. A avaliação passou então a ser considerada “uma função em si e para si mesma”, separada das funções de planejamento e de implantação de programas (Weilenmann: 1980, 24).

Para conseguir recursos das agências de financiamento, os planos e projetos passaram a incluir “ritualisticamente” uma proposta de avaliação na sua seção final (Suchman: 1967, vii, Prefácio de Leonard S. Cottrell Jr.).

Certamente por esta razão, os economistas, que haviam desenvolvido os principais conceitos da avaliação na área de produção industrial, passaram também a desenvolver “métodos para analisar as vantagens e os custos” dos programas públicos (Cohen&Franco: 1994, 93; Contandriopoulos et al.: 1992). Porém, a insuficiência das abordagens estritamente econômicas para a avaliação dos programas sociais e de educação levou a que o processo de profissionalização do campo da avaliação viesse a ocorrer em uma “perspectiva multidisciplinar” (Contandriopoulos et al.: 1992).

Nesta mesma época, as Ciências Sociais, em sua vertente norte-americana, passaram a adotar uma postura empiricista que enfatizava o rigor do método e da técnica para garantir a validade e a confiabilidade de seus achados, visando garantir que as pesquisas no campo fossem uma atividade livre de valores (“value-free enterprise”). Com isto, não havia mais lugar para a “avaliação” nos moldes tradicionais, que requeria uma atitude “normativa”, isto é, valorativa (Weilenmann: 1980, 24).

A “separação funcional” das atividades de avaliação do planejamento, aliada a esta busca de “cientificidade”, levaram ao surgimento de “uma metodologia de alguma maneira formalizada, institucionalizada” que passa a ser rotulada de “Pesquisa Avaliativa” (Weilenmann: 1980, 14).

Ou seja, enquanto campo formal de estudo, a Pesquisa Avaliativa é filha (“child”) das Ciências Sociais que se desenvolveu nos Estados Unidos, na década de 50 e início de 60, como uma “ferramenta” para elaboração de políticas, o que faz com que seja, ainda hoje, vista como uma forma de “ciência social aplicada”, de forma que seus teóricos e praticantes identificam-se como “cientistas sociais aplicados” (Weilenmann: 1980, 24-27).

Segundo SILVA & FORMIGLI, a adoção da mesma “perspectiva metodológica quantitativa” é a razão pela qual os modelos da Pesquisa Avaliativa “superpõe-se parcialmente a alguns dos tipos de estudos epidemiológicos: ambos têm uma matriz comum, derivada da incorporação da estatística e do método experimental às ciências sociais”. Nesta perspectiva, “o desenho **ideal**” de estudo para demonstrar o sucesso do programa é o desenho experimental, com grupo controle. Se após a intervenção, a diferença entre os dois grupos (experimental e controle) for estatisticamente significativa, considera-se que o programa avaliado foi um sucesso (Silva & Formigli: 1994; 84-86).

Praticamente, todos os autores consultados mencionam que os diferentes desenhos de estudo da Pesquisa Avaliativa foram inspirados na tipologia desenvolvida por D.T. Campbell e J. C. Stanley, em trabalho de 1966 - "Experimental and Quasi-Experimental Designs for Research". Verificamos⁵⁴ que esta publicação é uma re-edição da versão publicada no "Handbook of Research on Teaching", da "American Educational Research Association", em 1963. Ou seja, este trabalho foi patrocinado originalmente pela área de Educação, que segundo

⁵⁴Nas referências bibliográficas do artigo de Silva & Formigli: 1994.

YIN, é uma das raízes do desenho experimental clássico, juntamente com a pesquisa em agricultura (Yin: 1993,77).

Na área da Saúde Pública, vemos em uma referência ao uso do desenho experimental para propósitos avaliativos antes do trabalho de Campbell&Stanley, que referem-se aos modelos quase-experimentais. Segundo SUCHMAN, em 1955, Greenberg e Mattison haviam discutido esta proposta, em trabalho publicado no “Canadian Journal of Public Health”, no qual “enfaticavam o fato de que a pesquisa avaliativa era, ao final das contas, pesquisa, e portanto, para ser significativa, deveria ser conforme aos princípios básicos da metodologia da pesquisa” (Suchman: 1967, 19; 26, Nota de fim de capítulo 21).

Este desenvolvimento da avaliação resultou na profissionalização e institucionalização do campo nos Estados Unidos, com o surgimento de organizações (lucrativas ou não), agências governamentais ou semi-governamentais, institutos universitários e centros dedicados à avaliação. Além disto, surgiram periódicos, livros-texto, antologias, livros de bolso sobre o assunto. (Weilenmann: 1980, 27). Em 1976, foram criadas nos EUA duas áreas básicas de avaliação de programas - “programas de educação” e “outros programas sociais” - unificadas, em meados da década de 80, em uma única Associação Americana de Avaliação de Programas. Especialmente a área da avaliação em Educação está bem desenvolvida nos EUA, com "padrões" de avaliação bem definidos. Também destacam-se os estudos do grupo da psicologia norte-americana, que desenvolveu muitos trabalhos sobre as questões da validade de medida⁵⁵.

Uma rápida revisão na base de dados informatizada "Educational Research Information Center-ERIC", da "American Educational Research Association", revela a grande produtividade científica do campo da educação em avaliação, bem como a ênfase nas questões de medida.. No período 1991-1994, identificamos 6.685 trabalhos indexados sob o descritor "evaluation", classificados em grande grupo abarcado pelo título de "Researches and

⁵⁵Profª. Zulmira Araújo Hartz, na apresentação do tema "Avaliador-Epidemiologista: Desafios da formação 'indisciplinar'", na Comunicação Coordenada 47-"Novas Tendências de Pesquisa em Avaliação de Serviços", no III Congresso Brasileiro de Epidemiologia, realizado de 24 a 28 de abril em Salvador, Bahia, e em aula proferida no Curso de Pós-Graduação em Saúde Coletiva da UNICAMP, em 13.09.94, sobre o tema “Métodos Epidemiológicos aplicados à Avaliação de Programas”.

Measurement". Este grande grupo, por sua vez, divide-se em três subgrupos: 1 - "Research and Theory"; 2 - "Measurement"; 3 - "Tests and Scales".

No primeiro subgrupo - "Research and Theory" -, encontramos o descritor "Educational Assessment", onde estão os trabalhos que visam determinar a ocorrência e interpretar o alcance (ou a consecução) dos objetivos educacionais, tanto em nível nacional, estadual como local, visando o uso das informações para o planejamento, desenvolvimento, formulação de políticas e alocação de recursos. Neste sub-grupo, os trabalhos são classificados conforme os principais aspectos metodológicos do campo da avaliação: "evaluation methods", "educational assessment", "education quality", "formative evaluation", "summative evaluation", "educational research", "surveys", "personnel evaluation", "program evaluation", "program effectiveness", "program validation", "needs assessment", "outcomes of education".

No segundo subgrupo - "Measurement" - encontramos o descritor "Assessment Centers (Personnel)", que além dos artigos mais gerais em avaliação, contém os artigos relacionados à questão de desempenho profissional e/ou desempenho no trabalho. Na lista de descritores deste grupo, encontramos, "job performance", "personnel evaluation", além de "simulation", "situational tests", "evaluation methods".

Na área específica da educação em saúde ("health education"), encontramos 48 relacionados à avaliação. Também chamou atenção o grande número de avaliações de programas de educação de adultos e de desenvolvimento em zonas rurais, que revelam a raiz comum da avaliação na área de pesquisa em educação e em agricultura, apontada anteriormente por Yin.

Vimos até aqui que o grande desenvolvimento teórico-prático da avaliação nos EUA ocorreu no campo das Ciências Sociais, principalmente após a incorporação de um paradigma "quantitativista" e "experimental", que originou a denominada "Pesquisa Avaliativa". Do ponto de vista institucional, o desenvolvimento deste campo teve duas vertentes principais: a pesquisa avaliativa como ciência social aplicada ao planejamento de programas sociais públicos e a pesquisa avaliativa desenvolvida no campo da psicologia e da educação, que acabaram fundindo-se em uma única Associação Americana de Avaliação de Programas.

Porém, este crescimento não ficou restrito aos aspectos de produção de conhecimentos e de ensino, mas estendeu-se à aplicação concreta dos conhecimentos nas práticas de

planejamento das políticas públicas, em todas as suas fases - formulação de políticas, elaboração de programas, definição atividades. A prática da avaliação foi incorporada na própria estrutura institucional do Estado norte-americano, tanto através das avaliações realizadas internamente pela administração federal em seus órgãos, como através das avaliações externas feitas regularmente pelo “United States Accounting Office”, com a finalidade de subsidiar o Congresso dos EUA no controle da administração pública (McQueen: s/d, 32).

A situação no Canadá é semelhante à dos EUA. Segundo MCQUENN, as práticas de avaliação na esfera pública norte-americana influenciaram o desenvolvimento do campo da avaliação no Canadá. Em 1977, a partir das sugestões do “Auditor Geral” do país, foi definida uma “política federal de avaliação”, com a criação, no ano seguinte, do “Office of Comptroller General of Canada”. Além de controlar o manejo financeiro, que é vinculado ao Tesouro, este órgão tem por finalidade fazer a avaliação da “performance” da administração federal, cujos resultados são encaminhados ao Parlamento. Embora a avaliação desta política, treze anos depois, tenha revelado que as avaliações produzidas por este sistema de controle não estavam sendo utilizadas em todo o seu potencial pelos órgãos públicos, a simples implantação de uma “política de avaliação” por parte de um governo federal merece ser destacada (McQueen: s/d, 29-38).

Nos países da América Latina, em geral, o desenvolvimento do campo da avaliação nas Ciências Sociais tem sido bem diferente. AGUILAR&ANDER-EGG entendem que não são muitos “os especialistas latinos que escreveram sobre o tema”, considerando inclusive a produção científica da Espanha (Aguilar&Ander-Egg). Também no Brasil, alguns autores apontam uma relativa pobreza na literatura referente à análise de políticas e programas sociais, quanto comparada com a literatura anglo-saxônica (Draibe: 1989, 13; Bodstein: 1992, 58-59; Coimbra In Abranches: 1989: 66).

Entendemos que esta carência relativa pode ser atribuída basicamente a duas ordens de determinação: a primeira, a natureza autoritária e clientelista do Estado brasileiro, em contrapartida a uma sociedade civil pouco organizada e com uma noção de cidadania ainda muito frágil, o que evidentemente impede a criação de um clima propício à avaliação das políticas públicas; a segunda, a influência do pensamento marxista nas Ciências Sociais latino-

americanas, que em sua vertente mais ortodoxa, adotou durante um certo tempo uma postura cética em relação às políticas de bem-estar social, decorrente de uma visão do Estado como “instrumento da burguesia”, aliada a uma postura de rejeição à certas correntes metodológicas que se desenvolveram na América do Norte, dentre as quais provavelmente acabou sendo incluída a Pesquisa Avaliativa.

Em primeiro lugar, o objeto privilegiado da literatura anglo-saxônica sobre avaliação - a análise de programas sociais - não estava dado para a investigação, na América Latina e no Brasil, como havia ocorrido nos países da América do Norte e Europa. Segundo DRAIBE, “as noções de Estado de Bem-Estar Social ou de Estado Protetor parecem estranhas tanto à realidade quanto às tradições científico-acadêmicas brasileiras no campo da **social policy** ou **public analyses**” (Draibe: 1989, 13).

Está bem demonstrado na literatura que, apesar do discurso oficial, a Seguridade Social no Brasil não se constituiu enquanto política de Bem-Estar Social de fato, caracterizando-se, no máximo, como Seguro Social financiado quase integralmente pelas contribuições dos trabalhadores (Braga&Paula: 1986; Oliveira&Teixeira: 1982). Somente na Constituição Federal de 1988, com a instituição dos direitos sociais de educação, saúde, previdência social, assistência etc., estabeleceu-se a “responsabilidade objetiva”⁵⁶ do Estado neste campo. Obviamente, em não havendo políticas e programas sociais constituídas para funcionarem de fato, isto é, em atendimento ao que se entende por direito de cidadania, não existe o objeto primário sobre o qual constituiu-se o campo da avaliação na América do Norte e na Europa, conforme demonstramos anteriormente nesta seção.

Além disto, a atuação do Estado brasileiro tem sido caracterizada como autoritária, clientelista e corporativista (Braga&Paula: 1986; Oliveira&Teixeira: 1982; Campos: 1990), o que evidentemente não é compatível com a intenção de que suas políticas sejam avaliadas, deixando portanto de existir o requisito básico para que a avaliação seja realizada, isto é, a intenção de ser avaliado. Como bem demonstra SOUSA CAMPOS, a estrutura rígida e centralizada do Estado brasileiro gerou uma tradição de gerência estatal “burocratizada e

⁵⁶Utilizamos este termo no sentido atribuído por Frederich Mosher, conforme vemos citado por CAMPOS: uma responsabilidade objetiva “acarreta a responsabilidade de uma pessoa ou organização perante uma outra pessoa, fora de si mesma, por alguma coisa ou por algum tipo de desempenho [...] Quem falha no cumprimento de diretrizes legítimas é considerado irresponsável e está sujeito a penalidades” (MOSHER, Frederich apud CAMPOS, A.M.: 1990, 33).

autoritária, mas permeável ao clientelismo e às várias formas de utilização privada da coisa pública” (Sousa Campos: 1992, 90).

Segundo CAMPOS, observa-se na América do Norte uma tradição do Estado de prestar contas à sociedade civil, entendida esta prestação de contar em sentido amplo, isto é, não somente contábil, que é denominada de “accountability”. Esta tradição não se verifica no Brasil, a ponto de não dispormos nem ao menos de termo para traduzir esta palavra, o que reflete a própria ausência do conceito em nossa cultura (Campos: 1990, 31).

Para a autora, o Estado brasileiro comporta-se como se não tivesse obrigações perante a sociedade, fato que se deve, em grande parte, à fragilidade do “tecido institucional” da sociedade civil, ou seja, uma “ausência de foros onde as necessidades do cidadão são discutidas, consolidadas, traduzidas em demandas e canalizadas para os órgãos públicos”. Considerando-se que “o exercício da accountability é determinado pela qualidade das relações entre governo e cidadão, entre burocracia e clientelas”, esta ausência de “organizações processadoras de demanda” tem como consequência a “ausência de controles do público sobre o Estado” (Campos: 1990, 35-38).

NOVAES&PAGANINI referem o processo através do qual foi sendo criada, progressivamente, uma “cultura de qualidade” nos EUA, “que foi ocupando vários espaços acadêmicos e intitucionais [...], forçando as comunidades a pressionar os legisladores para aprovação de leis mais complexas, resultando inclusive na criação de entidades particulares, em alguns estados, para avaliar as contas pagas pela seguridades social”. Este tipo de pressão social levou a que, já em 1951, tivesse sido criada uma “Comissão Conjunta de Acreditação de Hospitais”, proposta que recém agora toma corpo no Brasil. Os autores entendem que “talvez por uma tradição da época colonial, nossa cultura foi forjada a admitir que a legislação autoritária, de cima para baixo, pode substituir uma longa etapa de prática social, que deveria anteceder às regulamentações governamentais” (Novaes&Paganini: 1994, xiv).

Enfim, como vimos no início desta seção, o grande motor do desenvolvimento da avaliação na América do Norte foi a necessidade do Estado “provar” ao cidadão contribuinte a eficácia de suas políticas e programas.

A segunda ordem de determinação refere-se à bastante conhecida influência do pensamento marxista nas Ciências Sociais latino-americanas, que manifestou-se no campo da

saúde principalmente através do movimento da Medicina Social, resultando, no Brasil, na proposta da “Saúde Coletiva”, como forma de contraposição à influência da “Saúde Pública” tradicional norte-americana (Donnagelo: 1983; Garcia: 1989; Granda: 1991; Minayo: 1992).

No campo da avaliação em saúde, esta influência manifestou-se, a nosso ver, de duas maneiras: a primeira, por uma certa desqualificação das políticas e programas públicos como objeto de avaliação das Ciências Sociais, a partir de uma visão do Estado como representante da burguesia; a segunda, refere-se às metodologias deste campo.

BODSTEIN chama atenção que “uma análise ortodoxa” da política de saúde e política social no Brasil levou a que se produzisse, nas décadas de 70 e 80, uma literatura que “trabalhava com paradigmas sobre Estado e política social bastante simplificados [baseados] na questão da acumulação e da legitimidade.” O tipo de análise realizada neste paradigma levaria à percepção das políticas sociais “com esse viés de uma política de cooptação, de uma política sempre de caráter reformista.” Para ela, esta visão “de certa maneira, desqualificava o próprio objeto de análise que eram as políticas sociais e as políticas de saúde”. Entretanto, este paradigma teria passado por “todo um esforço de revisão teórica”, de forma que a autora aponta a existência de “uma literatura recente que trabalha com a discussão sobre o estado de bem estar social, o famoso *welfare state*” (Bodstein: 1992, 58-59).

SOUSA CAMPOS também critica a simplificação esquerdista que “tem se constituído em um sério obstáculo para o desenvolvimento de análises marxistas sobre as instituições - império do poder burocrático - e até mesmo para a construção de teorias e de propostas socialistas sobre o bem-governar”. Esta simplificação ocorre entre algumas “vertentes mais ortodoxas”, apegadas às concepções iniciais de Estado do “marxismo tradicional” (Sousa Campos: 1992, 115-16).

A segunda forma como o pensamento marxista influenciou o desenvolvimento do campo da avaliação em saúde está relacionada à aspectos teórico-metodológicos das Ciências Sociais. NUNES coloca que a Medicina Social latino-americana incorpora os **conceitos macro-analíticos** referentes aos fatores estruturais, forças produtivas, relações de produção, que “passam a ser importantes referências conceituais para se entender a saúde, estabelecendo-se uma análise **crítica**, especialmente a partir da segunda metade dos anos 70, aos modelos

teórico-conceituais derivados do funcionalismo e divulgados pela **sociologia médica norte-americana** [grifos nosso]” (Nunes: 1995, 108-109).

A lógica dialética, que embasa o marxismo, mostra-se muito adequada para “o estudo dos fenômenos conflitivos”, oferecendo um excelente instrumental para a análise macro-teórica das políticas públicas; entretanto, não é uma lógica adequada para a análise de “processos lineares”, nos quais procura-se estabelecer de causa-efeito, às quais adapta-se melhor a lógica formal, base do positivismo (Testa: 1992, 51). A lógica dialética tem um “caráter decididamente **não operatório**, no sentido científico” (H. Lefebvre apud Bruyne et al.: 1982, 70), o que dificulta sua aplicação na avaliação dos efeitos das intervenções sobre os estados de saúde, cujo pressuposto é a demonstração de uma relação de causa-efeito.

De uma forma geral, coloca-se que as abordagens marxistas estão voltadas para as “macro” questões, enquanto as funcionalistas ocupam-se das “micro” questões, sendo a primeira mais crítica, e a segunda mais operacional. Ora, segundo HARTZ, a análise de programas necessariamente deve considerar a microteoria e a macroteoria, como enfoques complementares (Hartz: 1996). Neste sentido, ao trabalhar preferencialmente com conceitos macroanalíticos e com uma lógica dialética, o pensamento marxista impõe dificuldades extras para a abordagem da avaliação de programas. Entretanto, a abordagem marxista tem se revelado muito adequada para a análise crítica das políticas públicas, campo onde se manifestam os conflitos, havendo um grande produção teórica nesta linha no campo da avaliação das políticas de saúde.

Por outro lado, do ponto de vista político-ideológico, o pensamento funcionalista coloca a ciência em uma perspectiva de “neutralidade” que, segundo os marxistas, esconde uma proposta política conservadora. O pensamento marxista, ao contrário, assume explicitamente uma proposta política de transformação da sociedade, a partir da visão de que a ciência, como qualquer empreendimento humano, é atravessada por elementos ideológicos.

Esta divergência de caráter ideológico estabeleceu uma “discussão pública” entre as duas correntes de pensamento, que levou a uma confusão entre “o científico revelado nos conteúdos do estudo e o político que se manifesta nos propósitos”. Embora reconhecendo a relação histórica e inevitável entre ciência e política, TESTA coloca desta forma a questão:

“A confusão ocorre quando, a partir do marxismo, acusa-se o funcionalismo de ser um enfoque científico que tende à conservação, quando, na verdade, este é seu propósito político, ou quando a partir do funcionalismo, acusa-se o marxismo de não ser um enfoque científico, sendo apenas uma doutrina subversiva (quer dizer, ter propósitos políticos de mudança)” (Testa: 1992, 51).

Talvez em decorrência desta confusão, acreditamos que em alguns casos a crítica ao funcionalismo induziu uma certa “rejeição” aos métodos quantitativos, bem como sua pretensão de objetividade, que formam a base sobre a qual se estruturou o campo da avaliação nos Estados Unidos, em torno da chamada “Pesquisa Avaliativa”. O pensamento em Saúde Coletiva/Saúde Pública na América Latina, fortemente influenciado pela corrente marxista, manteve uma crítica muito acentuada à produção científica daquele país, entendida geralmente como “representante” do pensamento funcionalista conservador.

Apesar destes fatores limitantes, evidentemente que se produziram trabalhos sobre a temática da avaliação no campo das Ciências Sociais em saúde.

Em uma revisão da produção científica das Ciências Sociais aplicadas à saúde nos últimos trinta e cinco anos (de 1950 a 1985), intitulado "Bibliografía Lationamericana sobre Ciencias Sociales aplicadas a Salud", NUNES refere a existência de 2.749 trabalhos, os quais classifica em quatro grupos temáticos ou categorias: medicina tradicional, serviços de saúde, processo de saúde-enfermidade e formação de recursos humanos (Nunes: 1991, 62).

Os trabalhos relacionados à avaliação em saúde, em nosso entendimento, estão principalmente catalogados na categoria "serviços de saúde", que ocupa o primeiro lugar na produção científica do período estudado, com 36,3% do total. Esta categoria é dividida em doze subtemas, entre os quais destacamos "medicina estatal" e "programas de salud", com respectivamente 13,7% e 12,4% da produção total do grupo temático, e os subtemas "conducta de la población frente a los servicios de salud", "insumos para servicios de salud" e "investigación en salud" com menos de 5% (Nunes: 1991, 63).

Em resumo, vimos que o campo da avaliação desenvolveu-se, primordialmente, a partir da aplicação das metodologias de pesquisa em Ciências Sociais ao estudo dos programas públicos de saúde, educação e bem-estar social. Nos EUA e Canadá, este campo já está institucionalizado e profissionalizado, havendo uma ampla produção sobre o tema na literatura

anglo-francesa em geral. No Brasil, o desenvolvimento da temática da avaliação tem sido menor, tanto em decorrência do padrão autoritário, clientelista e corporativista da atuação do Estado, como em decorrência da inconsistência das políticas sociais, atribuídas em parte à “pobreza política” de uma sociedade civil pouco organizada, ambas as condições evidentemente pouco propícias ao desenvolvimento da avaliação de políticas e programas públicos.

No que concerne especificamente ao campo da Saúde Coletiva, a influência das Ciências Sociais referenciada ao pensamento marxista teve duas conseqüências: primeiro, um certo descaso pela avaliação de programas, para a qual este referencial apresenta algumas dificuldades metodológicas; segundo, a rejeição, pelo menos em parte, da produção da sociologia americana, onde originou-se a Pesquisa Avaliativa.

Entretanto, voltamos a salientar que as críticas marxistas às propostas sistêmicas vem sendo, aparentemente, incorporadas pela literatura anglo-francesa em avaliação, como vimos na parte referente às metodologias, especificamente a incorporação do “contexto” como um elemento da avaliação de programas, bem como da inclusão da avaliação “estratégica” das intervenções.

2.2. A avaliação no Planejamento em Saúde:

etapa do processo de planejamento, instrumento administrativo de controle.

Tradicionalmente, o tema da avaliação no campo da Saúde Pública/Saúde Coletiva tem sido vinculado à área do Planejamento.

Em função disto, tentaremos fazer uma revisão sucinta sobre as grandes correntes teóricas de Planejamento na América Latina, buscando verificar como a temática da avaliação é tratada em cada uma delas, a saber:

- o planejamento normativo, representado pelo chamado Método CENDES/OPS, elaborado em 1965 por uma equipe do Centro de Estudos do Desenvolvimento (CENDES) da Universidade central da Venezuela, com a colaboração da Organização Pan-Americana de Saúde (OPAS);

- o planejamento estratégico, representado, em nosso entendimento, por três propostas: a) a proposta do grupo do Centro Panamericano de Planificación de la Salud, do Chile (CPPS), patrocinada também pela OPAS e apresentada no documento “Formulación de Políticas de Salud”; b) a proposta de Carlos Matus (economista chileno), criador da corrente do Planejamento Estratégico Situacional (PES); c) a proposta de Mário Testa (sanitarista argentino), denominada de “Pensamento Estratégico em Saúde”.

Em geral, Matus e Testa são considerados os responsáveis pela crítica ao planejamento normativo e pela “formulação e pelo desenvolvimento do paradigma estratégico no campo do planejamento geral, e do planejamento de saúde” (Rivera: 1989, 7). Entretanto, incluímos o documento do CPPS-Chile na corrente do planejamento estratégico, por entender que o mesmo faz uma crítica ao aspecto técnico-normativo do Método CENDES/OPS, colocando o desenho de uma estratégia política como o aspecto central de sua proposta.

O trabalho que é considerado o marco inicial no campo do planejamento em saúde na América Latina é o documento "Programación de la Salud. Problemas Conceptuales y Metodológicos", que ficou conhecido como Método CENDES/OPS (OPS/OMS: 1965).

Divulgado em 1965, o documento foi elaborado em pleno contexto das teorias desenvolvimentistas difundidas pela Comissão Econômica para a América Latina (CEPAL) e pela Aliança para o Progresso. Estas teorias se baseavam em dois eixos:

- as idéias de Keynes, encampadas pela CEPAL, que propugnavam a intervenção planejada do Estado para promover o desenvolvimento econômico, onde o Estado atuava como um “macrossujeito” (Rivera: 1989, 13-14) ;

- a preocupação dos EUA com o desenvolvimento social dos países “subdesenvolvidos”, que, através da Aliança para o Progresso, recomendava aos países latino-americanos que aplicassem “formas de planejamento que integrem os setores social e econômico”. Esta recomendação era um dos requisitos para obter empréstimos através do recém-criado Banco Interamericano de Desenvolvimento (BIRD) (Rivera: 1989, 18).

A política da Aliança para o Progresso era, de certa forma, um “transplante” para a América Latina da política de combate à pobreza nos EUA, iniciada no início da década de 60 com a administração Kennedy, que deu origem à inúmeros programas na área da saúde, educação e bem estar-social naquele país. Da mesma forma que no país de origem, vinha acompanhada da exigência de que os Estado nacionais latino-americanos incluíssem a avaliação como etapa final do processo de planejamento.

Para satisfazer a exigência de planejar um desenvolvimento integrado (econômico e social), o método foi elaborado por uma equipe composta por economistas e especialistas em diversas disciplinas da área da saúde, visando “estimular un intercambio activo de ideas entre dichos profesionales”. Com o método, esperava-se conseguir que a política de saúde fosse elaborada “haciendo uso del rico caudal de conocimientos que las ciencias sociales y las matemáticas pueden poner a su disposición”. Além disto, havia a preocupação de que a programação de saúde pudesse integrar-se no plano geral de desenvolvimento do país (OPS/OMS: 1965, i; ii, Prefácio).

O pressuposto fundamental do documento era que a determinação das prioridades em saúde deveria ser feita dentro dos limites da escassez de recursos financeiros disponíveis, a partir de um diagnóstico epidemiológico da situação de saúde e da definição “técnica” das prioridades de ação. A principal questão metodológica era determinar a relação custo/benefício das ações programadas, baseando-se nas técnicas de “insumo-produto” e de “análise de projetos”, ponto que deu origem, posteriormente, às críticas à visão economicista do método (OPS/OMS: 1965, iii).

Inicialmente, o documento propõe ser tomado apenas como um “guia”, baseando-se na idéia de que “programar es una actitud, más que un método”. Entretanto, mais adiante, ao tratar do "proceso programatico", verifica-se que propõe etapas com "límites cronológicos definidos" e funções que poderiam se "precisar con exactitud" dentro do processo de planejamento, evidenciando uma concepção linear deste processo e uma postura normativa. As etapas do planejamento seriam: "1) diagnóstico; 2) la programación propriamente dicha; 3) la discusión y la decisión; 4) la ejecución, y 5) la evaluación y la revisión" (OPS/OMS: 1965, ii, Prefácio; 12-13). A forma como o documento apresenta esta sucessão linear de etapas, cada

qual com uma série de procedimentos matemáticos de cálculo, deu origem às críticas ao caráter “normativo” ou “prescritivo” da metodologia.

Como era de se esperar, a seqüência culminava com uma etapa final de avaliação e revisão. Na parte do documento que trata desta etapa, verificamos que o termo "evaluación" é substituído por "control". O documento limita-se a propor que o programador deva fazer um revisão periódica do plano, tarefa para a qual seria necessária “la información que se deriva del control”. Este controle teria o papel de indicar "en qué medida se están cumpliendo las metas del plan y, si no se están cumpliendo, deve señalar las causas. Se trata pues de un problema de información y de análisis". Através do controle, o planejador teria as informações para revisar o plano em execução e para elaborar novos planos de saúde (OPS/OMS: 1965, 20).

A avaliação fazia parte também da etapa inicial do plano, o diagnóstico de saúde. As quatro tarefas do diagnóstico seriam: a descrição (de preferência em termos quantitativos) da situação atual do objeto programático; o conhecimento dos fatores que determinam esta situação; a análise da situação atual e definição das perspectivas futuras e, finalmente, a “avaliação” da situação, isto é, a determinação se é ou não satisfatória (OPS/OMS: 1965, 12-13). Para executar esta quarta tarefa do diagnóstico, seria necessário "la determinación de un patrón apropiado de comparación" (p. 15).

Em suma, além da coleta de dados sobre a situação de saúde, o diagnóstico de saúde compreenderia três tarefas principais : a explicação, o prognóstico e a avaliação da situação de saúde (p.42-79).

A tarefa do programador não se limitaria a descrever a situação atual de saúde, mas “explicarla, a la luz de la influencia que ejercen sobre ella los distintos factores condicionantes” , tarefa para a qual o programador não se poderia basear exclusivamente nos conhecimentos epidemiológicos, visto que “la epidemiologia no ha llegado a cuantificar la importancia de cada uno de los factores condicionantes ni el grado de influencia que ejercen sobre el estado de salud” (OPS/OMS: 1965, 42).

Em seguida, o programador deveria estabelecer um prognóstico da situação de saúde. Baseado na explicação da situação atual e tendo como pressuposto a “no modificación de la política de salud”, o programador deveria perguntar-se o que ocorreria com a situação de saúde da comunidade “digamos en los próximos 10 años” (OPS/OMS: 1965, 44).

Por fim, o programador deveria avaliar a situação de saúde, ou seja, “expresar un juicio objetivo sobre si la situación es satisfactoria o no”. O documento esclarece que “cualquier juicio de esta naturaleza siempre envuelve comparar el estado de la situación observada con otro estado que se considera aceptable” e alerta que o resultado a ser obtido desta análise dependerá dos “critérios” utilizados (OPS/OMS: 1965, 47). Nesta fase de diagnóstico, o programador deveria avaliar ainda os instrumentos (ou procedimentos), as tarefas (ou atividades), as técnicas e os rendimentos (desempenhos), bem como “la eficiencia con que se emplean para atacar cada daño” (p. 48).

Na definição de prioridades de ação, o documento institui os clássicos critérios de magnitude, transcendência e vulnerabilidade. A magnitude e a vulnerabilidade seriam operacionalizadas por procedimentos oriundas da Epidemiologia, que ficam bem demonstrados no documento. Entretanto, em relação à determinação da transcendência, o documento assume que não proporá uma norma para medir a transcendência social da saúde e resolver assim o problema das prioridades, visto que este critério remete à escala de valores da sociedade, devendo a questão ser resolvida em cada “caso concreto” (OPS/OMS: 1965, 6).

Embora criticado por sua visão economicista e normativa do planejamento em saúde, o método foi amplamente divulgado, de forma que este documento fundamentou a terminologia básica da avaliação no campo da Saúde Pública: cobertura e concentração (de atendimentos/consultas ofertadas nos serviços), prioridades e metas (do planejamento), magnitude, transcendência e vulnerabilidade (dos problemas de saúde), eficácia, eficiência e impacto (das intervenções programadas).

Dez anos depois, em 1975, a Organização Pan-Americana de Saúde publica um novo documento que passa a ser o segundo marco na teoria de planejamento em saúde na América Latina - "Formulación de Políticas de Salud", elaborado pelo grupo do Centro Panamericano de Planificación de la Salud (CPPS), do Chile.

Como o próprio título do documento indica, sua ênfase principal era na questão da formulação de políticas de saúde. O pressuposto fundamental do método era que as chances de que um plano viesse a dar certo dependiam essencialmente de uma boa "avaliação" de suas possibilidades na esfera política. Na parte inicial do documento são colocados alguns conceitos básicos da metodologia: política, processo político e sistema político. Ao invés da

programação formal, linear, burocratizada do documento anterior, a formulação do plano de saúde passa a ter um caráter “estratégico”.

No que tange à avaliação enquanto etapa do diagnóstico em saúde, o Método CENDES/OPS substitui, de forma radical, o “diagnóstico epidemiológico”, que pretende ser objetivo, pela avaliação subjetiva “del político”, cujo estilo intelectual “utiliza más la síntesis que el análisis”. Neste sentido, a própria essência do método de planejamento proposto seria “traducir e interpretar este modelo mental [do político], haciéndole secuencial y explícito” (OPS/CEPAL: 1975, 35). O passo inicial do processo começaria então com a “síntese” realizada pelo “político” que é assim descrita:

“Su conocimiento de una situación presente que considera insatisfactoria o perfectible y la posesión de una escala de valores evoca en él una situación futura deseada (imagen-objetivo). Tanto en la identificación de problemas actuales como en la construcción del futuro deseado, el proceso de ideación está orientado por valores, es decir, ideológicamente” (OPS/CEPAL: 1975, 35).

O processo de planejamento constitui-se essencialmente da análise da **factibilidade**, **coerência** e **viabilidade** das diferentes proposições políticas formuladas, com vistas ao alcance da imagem-objetivo. Ao final desta “análise política” das possibilidades alternativas do plano, seria feito o “diseño de una estrategia” para implementar a alternativa escolhida (OPS/CEPAL: 1975, 51).

Por fim, após a formulação das políticas e do desenho da estratégia, viria a etapa de “formalização” das mesmas através da elaboração de dois documentos escritos (o plano propriamente dito): um, de caráter reservado, que conteria as informações que não pudessem ser reveladas para não prejudicar a estratégia do plano; e outro, de caráter público, que conteria o esquema geral do plano e seus elementos técnico-administrativos (OPS/CEPAL: 1975, 61-63). Tanto um como outro documento deveriam conter os procedimentos de avaliação. O documento “secreto” (o primeiro), incluiria “las medidas de control o corrección a adoptar” do ponto de vista político. No documento público (o segundo), mais uma vez consta, como componente final do “esquema geral” do plano, o item “control y evaluación de la orientación y conducción de las políticas y sus resultados”. Este item seria constituído pelo

"diseño de un sistema de información que provea conocimientos de interés político e informe acerca de la realización de las acciones propuestas, de los resultados obtenidos y de los beneficios logrados. Debe permitir una apreciación oportuna del funcionamiento del sistema y de los productos entregados a la población, con el fin de revisar o confirmar la orientación de la política y la conducción que se ha establecido" (OPS/CEPAL: 1975, 66-67).

Ou seja, da mesma forma que na proposta anterior, os procedimentos de avaliação estão descritos como procedimentos de controle, que forneceriam as informações retroalimentadoras sobre a execução do plano. Percebe-se, entretanto, que a visão de “controle” administrativo-burocrático é substituída por uma visão de estratégica de um “sistema de informações” que aporte conhecimentos de **interesse político**.

A metodologia da CEPAL contribuiu para superar a visão estritamente economicista do processo de planejamento em saúde, bem como se contrapôs frontalmente à pretensão de “objetividade técnica” da proposta anterior, deslocando a questão do estabelecimento de objetivos e prioridades para o campo político, isto é, o campo dos interesses e valores. Conseqüentemente, a ênfase da avaliação deslocou-se dos procedimentos burocratizados de controle da execução do plano e da verificação do cumprimento de metas para a esfera ampliada da disputa política, acrescentando à terminologia do planejamento em saúde os conceitos de "viabilidade", "factibilidade" e "imagem-objetivo".

Os representantes mais expressivos da corrente teórica do planejamento estratégico na América Latina, porém, são Carlos Matus e Mário Testa.

MATUS, economista chileno, foi ministro da economia do governo de Salvador Allende, no Chile. A partir desta experiência concreta de governo, criou a Teoria do Planejamento Estratégico Situacional (PES), com a finalidade de instrumentalizar, de uma forma geral, os governos da América Latina, especialmente na instância nacional⁵⁷. Embora

⁵⁷Na verdade, a proposta de Matus inclui três métodos de planejamento, adequados para diferentes “espaços” de governo (nacional, estadual, local) e diferentes “níveis” de gerenciamento (diretivo, intermediário, base): 1- Planejamento Estratégico Situacional (PES), para os níveis diretivos nacional e estadual; 2- ZOPP (sigla referente ao método “utilizado pela GTZ do Governo da ex-Alemanha Ocidental), recomendado para os níveis intermediários nacional e estadual, bem como o nível diretivo local; 3- Método Altadir de Planificação Popular (MAPP), recomendado para o nível intermediário do espaço local e para o nível de base dos três espaços de governo. MATUS, Carlos. *Programa do Curso “Governo e Planificação”*, s/d, p. 72

sua proposta de planejamento não seja específica para o setor saúde, o autor vem sendo adotado como referência nesta área.

Ele amplia a base conceitual do planejamento, pois parte do pressuposto de que existem quatro esferas de "domínio" da realidade sobre a qual se quer planejar: Político, Econômico, Cognitivo e Organizacional (Matus: s/d, 88-89). A cada um destes âmbitos corresponde um tipo de recurso, que são disputados por diferentes "atores sociais"⁵⁸, visto que são (sempre) recursos escassos (p. 110). A partir destes pressupostos, o autor desenvolve uma metodologia de planejamento que tem por eixo uma "teoria da ação social" cujos aspectos fundamentais são os seguintes⁵⁹:

- a realidade social é complexa e diante disto é preciso considerar que muitos aspectos da mesma não podem ser previstos e/ou controlados, ou seja, o futuro é incerto;

- a vida social é um "jogo de poder", onde diferentes jogadores ou "atores sociais" fazem suas "apostas" (isto é, encaminham seus projetos) buscando "vencer o jogo". Neste sentido, "o papel do plano é melhorar o sistema de apostas" do ator/jogador que planeja sua ação em um clima de incerteza (Matus: s/d, 53).

A metodologia de planejamento proposta por MATUS consiste, essencialmente, em avaliar muito bem a "situação" inicial - isto é, a distribuição do "poder" entre os atores - e a partir daí definir o "cenário" mais provável do futuro, desenhando então uma "estratégia" de ação que permita "acumular poder" no jogo social, e assim garantir a implantação do plano. O planejamento seria então o processo pelo qual um ator social estabelece um curso de ação estratégica (plano), em direção a uma "situação-objetiva" futura que deseja alcançar. O plano em si (ou "programa direcional") seria uma "grande aposta de final aberto", considerando que a maioria dos problemas sociais não são predizíveis nem previsíveis, mas incertos (Matus: s/d, 57). O eixo do raciocínio estratégico é a análise da viabilidade do plano, referida aos quatro âmbitos da realidade mencionados anteriormente (Matus: s/d, 109-110).

No PES, o planejamento é visto como um processo dinâmico, onde o conceito de "etapas" que se sucedem linearmente é substituído pelo conceito de "momentos", que podem e devem ser constantemente retomados, como um processo contínuo ou "en cadena que no

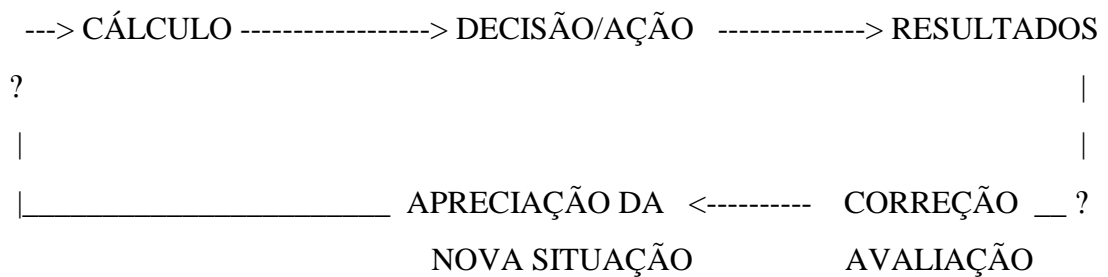
⁵⁸O autor define quatro "tipos" de atores sociais: partidos políticos, forças sociais, personalidades e governos institucionais (MATUS: s/d, 84-85).

⁵⁹Ver também MATUS, Carlos. *Política, Planificación y Gobierno*. Fundacion Altadir, Caracas, Venezuela, s/d.

tiene comienzo ni termino definido” (Matus: 1989, 371). Para efeito didático, os quatro momentos são apresentados de forma linear: o ponto de partida do processo é o "momento explicativo", seguido dos momentos "normativo", "estratégico" e "tático-operacional".

Coerentemente com esta proposta, a metodologia não prevê um momento “final” de avaliação do plano, mas deixa entrever que a avaliação é um processo permanente que retroalimenta a implantação do mesmo. Não havendo um “lugar” específico na obra de Matus para tratar a temática da avaliação, tivemos dificuldade de localizá-la nas obras consultadas.

Em nosso entendimento, o que talvez melhor corresponde à avaliação, na metodologia de Matus, é a noção de “cálculo interactivo”, que gera o “juicio estratégico” (Matus: 1989, 361; 185). Para este autor, a “repetição constante do cálculo de planificação é uma defesa contra a incerteza”. Ele resume esta idéia no seguinte esquema (Matus: s/d, 43):



Ainda explicando o conceito de “cálculo interactivo”, Matus coloca o seguinte:

“Debo calcular a mayor velocidad y con capacidad de anticipación del cambio situacional real. Ese cálculo debe constituirse en un proceso incesante y permanente de prealimentación sobre el futuro y de retroalimentación sobre el pasado reciente. Así mi plan se parece más a la estrategia en un juego que a un diseño normativo” (Matus: 1989, 186).

Apesar da dificuldade, localizamos o termo “avaliação” (evaluación) e alguns conceitos correlatos em alguns dos “momentos” do processo de planejamento de Matus.

No momento explicativo, quando é formulada a "explicação situacional ou diagnóstico", os problemas que deverão ser objeto da ação do plano são identificados e explicados através do desenho da “árvore de problemas”. Os elementos da mesma são o “vetor

de definição do problema (VDP)” e os “nós críticos” (Matus: s/d, p. 33). Cada problema é descrito por um VDP mediante a “enumeração de descritores [...] necessários e suficientes para precisar o conteúdo do problema” (p. 45). Os “nós críticos” seriam os pontos do problema onde uma ação causaria maior impacto sobre o mesmo, de forma que “cada nó crítico demanda uma ação sobre ele”, isto é, uma “operação” (p.34). O projeto de governo é “modificar o VDP (Vetor de Definição do Problema) dos problemas até alcançar o VDR (Vetor de Definição de Resultados) que propõe a situação-objetivo” (p.30). Quando o autor enumera os “requisitos” que devem cumprir os descritores de um problema (p. 15), percebe-se uma evidente similitude entre o conceito “descritor” com o tradicional conceito de “indicador”.

Há duas diferenças fundamentais entre o “diagnóstico” de Matus e o das propostas anteriores: a primeira, é que este diagnóstico seria apenas uma "explicação" da realidade, não tendo a pretensão de corresponder a um retrato exato da mesma; a segunda, é que o “sujeito” do planejamento deve sempre considerar a existência de "outros" sujeitos que têm outras "explicações" (isto é, outras avaliações) da "situação" inicial. (Matus: s/d, XIV).

Também no momento normativo, quando discute as “operações” a serem definidas no programa direcional (ou plano propriamente dito), verificamos que o autor refere-se a procedimentos de avaliação, como a discussão dos “critérios” para avaliar as operações: valor normativo, efetividade, confiabilidade e eficiência, e eficácia como um dos itens que compõem a efetividade (Matus: s/d, 68; 70-71).

Da mesma forma, no momento estratégico, é recomendada a simulação (sala de situações) como procedimento de “avaliação” das diferentes estratégias alternativas (Matus: s/d, 104), onde são avaliadas as “trajetórias de ação” possíveis (p. 108).

No momento tático-operacional, quando trata dos "subsistemas de direção do plano", o autor define um subsistema de "petição e prestação de contas", que não se refere ao aspecto financeiro do plano, mas "aborda a evolução da marcha do processo de mudança situacional" através de "momentos, procedimentos e critérios de avaliação" que deverão ser empregados pelos gerentes do plano, nos diferentes níveis de hierarquia do sistema de direção (Matus: s/d, 121). Ele utiliza, em alguns momentos, o termo "monitoramento".

Passamos então para o outro teórico do planejamento em saúde na América Latina. TESTA participou diretamente da elaboração da proposta CENDES/OPS e foi consultor da

equipe que elaborou a proposta da CEPAL. Entretanto, posteriormente tornou-se um crítico destas metodologias, por considerá-las "normativas" ou "prescritivas", colocando-se como adepto da “linha de planejamento estratégico”, como Carlos Matus (Testa: 1992, 109). Apesar de endossar boa parte das posições deste, Testa considera que ele “substitui a normatividade técnico-econômica por uma normatividade política, o que reduz a eficácia de sua proposta”. Testa diferencia-se de Matus na medida em que adota uma auto-referência de “militante de base”, enquanto este último coloca-se na perspectiva de “dirigente de processo”, isto é, na perspectiva de governante (p. 108).

Para ele, a reflexão crítica (e autocrítica) sobre a crise do planejamento em saúde na América Latina tem duas vertentes explicativas: a visão da crise como uma questão interna do planejamento em saúde - “reiterados fracassos dos planos formulados em quase todos os países” -, ou como um reflexo da crise (geral) nos países capitalistas dependentes - “manifestação mais profunda da desarticulação do capitalismo dependente” (Testa: 1992, 99). Assumindo a segunda posição, Testa coloca que na América Latina “o objetivo central do planejamento deve ser o de contribuir para promover a criação de condições que possibilitem a realização de ações que levem a mudanças sócio-políticas nos países”, cujo mecanismo principal seria o aprofundamento das contradições dentro das instituições de Estado (pp. 110, 112).

A discussão crítica sobre o planejamento, conduzida, segundo ele, principalmente na “segunda metade da década de sessenta e a primeira de setenta” (p. 99), teria assumido um duplo caráter: a análise de “métodos” e “objetivos” e a referência a “problemas mais próximos da epistemologia”. Neste enfoque epistemológico, coloca-se a questão da tríplice racionalidade do planejamento: política, administrativa e científica (Testa: 1992, 127, Nota nº 21).

Em função desta tríplice racionalidade, o planejamento trata de três formas de poder: 1) o poder técnico, que lida com diferentes "ordens" de informação (médica, sanitária, administrativa e marco teórico), as quais circulam em vários “âmbitos” (docência, investigação, serviços, administração superior e população); 2) o poder administrativo, o qual "visto da perspectiva da acumulação de poder no setor [saúde], é sintetizado quase completamente, nos países capitalistas, através das diversas formas de financiamento" ; 3) o

poder político, que para ele é "a capacidade de desencadear uma mobilização", a qual depende basicamente de uma "visão de realidade" ou "ideologia" (Testa: 1992, 118-124).

Para atingir seus objetivos de mudança, o planejamento funcionaria como um "mecanismo facilitador de situações que desencadeiem ou acelerem processos de mudança social", através da busca do "aprofundamento das contradições" nas instituições de Estado, pois estas são o *locus* de formulação das políticas (Testa: 1992, 112).

O elemento central de sua teoria é o "postulado de coerência", raciocínio através do qual estabelece-se "uma relação necessária entre **propostas, métodos** para alcançá-las e organização das **instituições** que se encarreguem disso [grifos nossos]" (Testa: 1992, 110). A eficácia da ação institucional seria alcançada "através da adequação entre propósitos institucionais e formas de organização" (p. 154).

Na formulação deste postulado, Testa busca respeitar a coerência entre três esferas fundamentais: ; a político, que determina os propósitos de governo; a teoria (ciência), que determina os métodos para alcançar estes propósitos; e a história, que determina as organizações ou instituições através das quais se executam estes métodos. A articulação entre estas três esferas dá-se através das três racionalidades do planejamento - a política, administrativa e técnica (Testa: 1992, 116). Tentamos resumir o complexo raciocínio de TESTA no esquema do Anexo XIX.

TESTA não faz uma discussão específica sobre avaliação, da mesma forma que Matus. Porém, como as instituições de Estado são o foco principal da ação de planejamento, a terminologia da avaliação - eficiência, eficácia - é colocada em relação à atuação global das mesmas, e não ao plano, programa ou serviços. A eficácia institucional seria alcançada pela análise das relações entre os efeitos e tempos de demora das ações entre as três esferas de poder - político, administrativo e técnico (Testa: 1992, 123-24).

Em resumo, a prática discursiva do Planejamento em Saúde na América Latina, e conseqüentemente no Brasil, em relação à temática da avaliação, tem-se dado historicamente em torno dos seguintes eixos conceituais:

- a avaliação refere-se aos âmbitos técnico, organizacional, econômico/financeiro e político;

- a avaliação é um conceito que permeia o processo de planejamento sob diferentes “nomes” ou “termos”. Na perspectiva linear do planejamento normativo, a avaliação corresponde a "etapas" do processo: no momento inicial do plano (sob o nome de diagnóstico de saúde), no decorrer da implantação do plano (como função de controle) e no final do plano (sob o próprio nome de avaliação). Na perspectiva dinâmica do planejamento estratégico, a avaliação aparece em "momentos" recorrentes, geralmente sob o nome de “análise” da situação ou do contexto, mas também com o seu próprio nome (“evaluación”), ou embutida dentro dos conceitos de controle e de monitoramento;

- a avaliação é uma função de "controle administrativo" ou de "monitoramento" da execução do plano, e esta função está estreitamente vinculada à questão dos sistemas de informação.

Nas metodologias de planejamento analisadas, o papel da avaliação pode ser resumida no esquema conceitual da Teoria de Sistemas, que está na base da maioria das abordagens deste campo, como vimos nas seções referentes ao conceito e às metodologias de avaliação:

Avaliação Inicial -----> Intervenção -----> Avaliação Final
(Diagnóstico de Saúde) (Plano) (Resultado ou Impacto)

"Insumos" -----> "Processo" -----> "Produto"

Ainda que claramente vinculado ao marxismo, entendemos que mesmo as idéias de Testa podem ser abarcadas por este esquema, geralmente identificado com as idéias funcionalistas. Ele demonstra dar crédito à teoria dos sistemas em dois momentos: quando utiliza conceitos da mesma e quando afirma que há uma “coerência lógica” na utilização do planejamento normativo pelo Estado soviético, pois na (então) URSS, “a história e o plano se continuam” no traçado do futuro (Testa: 1992, 98). Além disto, a própria formulação de “coerência” de seu postulado atribui um papel “sistêmico” ao método de planejamento, ainda que este seja entendido como historicamente condicionado e que deva estar politicamente comprometido com o marxismo.

Para encerrar esta análise, caberia lembrar que, no campo do Planejamento em saúde, as práticas de avaliação são pouco freqüentes no Brasil. Em nosso entendimento, SILVA & FORMIGLI resumem a situação quando colocam que "como componente indispensável do processo de planejamento, a avaliação, entre nós, é muito lembrada, pouco praticada e, quando realizada, não é muito divulgada, sendo dificilmente utilizada para a tomada de decisões" (Silva & Formigli: 1994, 87). SILVER confirma esta colocação: "a avaliação é uma das atividades menos praticadas no planejamento" (Silver: 1992, 197).

Da mesma forma que vimos no histórico da avaliação em Ciências Sociais, as causas estão enraizadas nas características dos Estados latino-americanos em geral, e do Brasil em especial.

Ao analisar a contribuição do planejamento normativo, desenvolvido na vigência das teorias desenvolvimentistas da década de 60, RIVERA aponta o "alcance limitado" de suas propostas, devido aos obstáculos representados pelos grupos de poder tradicionais, "que atuam de acordo com critérios tradicionais de alocação de recursos", e também à presença de "um forte aparato administrativo tradicional, inerte, clientelista", que atua como um "sério entrave à modernização que o planejamento comporta." Em função disto, ele conclui que "o planejamento na América Latina possui um desenvolvimento teórico muito mais rico do que as suas possibilidades de aplicação reais" (Rivera: 1989, 20-21).

Apesar da influência cada vez maior do planejamento estratégico no setor saúde do Brasil, consideramos que em nosso país, em qualquer das três instâncias de comando do Sistema Único de Saúde, os obstáculos apontados por RIVERA continuam determinando que o uso do planejamento, e conseqüentemente da avaliação, continue sendo apenas "virtual" (Rivera: 1989, 20-21).

Ou seja, o pouco desenvolvimento da avaliação dentro do Planejamento em Saúde pode ser atribuído mais às resistências dos tradicionais grupos políticos de poder em tornar transparente o processo de gestão - implicação natural dos processos avaliativos - e menos a questões de ordem epistemológica/metodológica. Esta quase ausência de práticas de avaliação no processo de planejamento reflete-se também na gerência dos serviços. Como bem lembra SOUZA CAMPOS, no processo de formação tradicional no campo da administração, o gerente é concebido apenas como "uma correia de transmissão das determinações gerais de

quem detém o poder (governante, proprietário, conselhos de direção, etc.)”, restando-lhe apenas “o consolo da servidão voluntária” (Souza Campos: 1992, 194).

Em resumo, na vertente normativa do planejamento em saúde na América Latina, a avaliação resume-se à última fase do que AGUILAR&ANDER-EGG referem como “estrutura básica do procedimento” de programação (Aguilar&Ander-Egg: 1994, 22, 28). Além disto, o conceito de avaliação que encontramos nos textos avaliados confunde-se com as noções de controle administrativo ou monitoramento.

Na vertente estratégica do planejamento em saúde, a noção de avaliação confunde-se com o próprio raciocínio estratégico, com a única diferença que a avaliação visa primariamente fornecer informação e produzir conhecimento sobre a ação organizada, visando aprimorá-la, enquanto o raciocínio estratégico visa aumentar o poder do ator que planeja.

2.3. A avaliação na área da Epidemiologia: a qualidade da atenção à saúde, a organização dos serviços, o impacto dos programas

Na área da Epidemiologia, o desenvolvimento da temática da avaliação vem ocorrendo através de duas abordagens complementares - a de avaliação de qualidade e a de a avaliação de serviços e programas de saúde -, que já vimos anteriormente, na seção referente às linhas discursivas gerais da avaliação em saúde.

A maior parte da produção científica sobre a abordagem epidemiológica da avaliação está na literatura anglo-francesa, especialmente do Canadá, EUA e Inglaterra.

Na vertente da avaliação de serviços de saúde, há indícios de que a produção canadense foi a precursora. A reorganização do sistema de saúde no Canadá (Reforma Sanitária), divulgada através do “relatório Lalonde”⁶⁰, de 1974, pela OMS, deu origem a uma redefinição do papel da Epidemiologia, que passa a incluir os fatores relacionados à organização dos serviços médico-sanitários como um dos determinantes do processo saúde-

⁶⁰Segundo informação pessoal da Prof. Zulmira Hartz, Marc Lalonde foi Ministro de Saúde do Canadá.

doença. Este modelo ampliado de determinação do processo saúde-doença está baseado no enfoque de risco, e inclui quatro fatores determinantes: biológicos, ambientais, estilo de vida e organização dos serviços de saúde (Hartz: 1993, 9).

Em 1976, nos EUA, Allan Dever ele escreveu um artigo onde desenvolve a matriz conceitual canadense, que passa a ser conhecida como “modelo epidemiológico para a análise da política de saúde”, que, segundo HARTZ, desencadeou “uma série de críticas pela colocação, em igual nível de importância, dos determinantes sociais e biológicos dos problemas de saúde, bem como pela possibilidade de culpabilizar o estilo de vida das vítimas” destes problemas. Para ela, embora estas críticas tenham construído uma das maiores contribuições da Epidemiologia latino-americana para “a tentativa de (re)orientação das políticas sociais voltadas para a saúde”, tiveram como efeito secundário um menor investimento da área para integrar “no ensino e na pesquisa epidemiológica” os fatores relacionados à organização dos serviços médico-sanitários (Hartz: 1993, 9).

Na Inglaterra, KNOX et al., em publicação de 1979, adotam a mesma perspectiva de Dever. Segundo estes autores, embora a idéia de aplicar a Epidemiologia à Administração de serviços de saúde não fosse nova, seu desenvolvimento sistemático teria sido recente (Knox et al.: 1986, 1). Para eles, enquanto a Epidemiologia tradicional trabalhava no campo da manutenção da saúde e da prevenção, alguns epidemiologistas estariam voltando sua atenção para os efeitos do cuidado médico (“health care”) na incidência e na severidade de certos processos mórbidos, passando a focar seus estudos sobre a eficácia dos procedimentos diagnósticos e terapêuticos. Mais recentemente, este enfoque teria sido ampliado, surgindo a preocupação com o desempenho dos serviços de saúde em seu conjunto, originando o enfoque no planejamento do cuidado à saúde (“health care planning”), como um aspecto em separado do planejamento em saúde em geral. Para os autores, estes enfoques demonstram o reconhecimento da importância do cuidado à saúde como um determinante da própria saúde, embora algumas vezes esta importância tenha sido mal-compreendida ou exagerada (p. 2-3).

Na América Latina, o marco de referência para este novo enfoque da Epidemiologia foi o seminário internacional sobre o tema "Usos y Perspectivas de la Epidemiologia", organizado pela OPAS na cidade de Buenos Aires, em 1983 (Hartz: 1993, 9). Provavelmente, este evento

fazia parte de uma estratégia desta organização para divulgar os aspectos positivos do “modelo epidemiológico para análise da política de saúde”, que vinha sendo duramente criticado.

Neste seminário, foram consagrados os três “âmbitos” ou “usos” da Epidemiologia: “control de las enfermedades”, “evaluación de las condiciones de salud” e “planificación de los servicios” (OPS/OMS: 1984, vii). Na publicação que reúne os documentos do seminário, verifica-se que há um reconhecimento de que o conhecimento epidemiológico já vinha sendo aplicado para orientar o planejamento de serviços, além do tradicional enfoque no controle de doenças e no diagnóstico de saúde das populações/comunidades. Porém, estas aplicações da Epidemiologia foram colocados como “moderno enfoque”, como “aplicaciones novedosas” decorrentes de uma “visión más amplia de la epidemiología y su aplicación a los servicios de salud” (OPS/OMS: 1984, 4).

Este caráter de novidade devia-se, pois, a uma concepção mais dinâmica e ampliada da Epidemiologia, como podemos verificar no discurso de abertura do Dr. Carlyle Guerra de Macedo, então diretor da OPAS, que de certa forma resume as proposições fundamentais do seminário:

*"Nos referimos a la epidemiología como función de inteligencia; **la epidemiología como función de evaluación**, no solo de los problemas de salud y de los servicios de salud, sino de todo el contexto que tiene importancia para la comprensión de los fenómenos de salud; a la epidemiología que va mas allá del problema de salud específico o de la enfermedad, para considerar a la sociedad como fuente de explicación para los problemas y para las soluciones de salud; a la epidemiología que anticipe el conocimiento de las necesidades, que identifique y determine las condiciones de riesgo y que deba orientar la definición de prioridades y la utilización de los recursos disponibles, convirtiéndose así en instrumento para la planificación y para la conducción estratégica de los sistemas de salud.[grifo nosso]" (OPS/OMS: 1984, 8).*

Reforçando essas colocações, destacamos também a fala do Ministro de Saúde Pública e Meio Ambiente da Argentina, Dr. Horácio Rodrigues Castells, em seu discurso de boas vindas aos participantes do seminário:

“Con el enfoque epidemiológico, la evaluación de la salud y de los servicios de atención - así como de su organización y funcionamiento - adquieren una nueva dimensión, que posibilita una visión sistémica de la salud de una población, de su entorno social, de los factores condicionantes y del verdadero valor de nuestro accionar[grifo nosso]” (OPS/OMS: 1984, 3).

Vemos, nestes discursos, que a Epidemiologia é colocada basicamente como “função de avaliação”, através da qual exerceria seu papel principal dentro da Saúde Pública, qual seja, instrumentar o planejamento dos serviços de saúde.

Para GALVIS, diante do desafio apresentado por Guerra de Macedo em seu discurso, as atividades da Epidemiologia deveriam ser reorientadas “buscando su aplicación en el proceso administrativo” (Galvis: 1984, 238).

Segundo o já referido ministro argentino, o ponto de partida para este enfoque teria sido a publicação de Morris “Usos da Epidemiologia”, ao final da década de 50, que dedicava alguns capítulos “a la evaluación de servicios de salud”, os quais haviam assinalado “fructíferos conceptos, que permitieron experiencias precursoras en los años siguientes”. Ele assinalava ainda que quase 15 anos depois desta publicação, na VI Conferência Internacional de Epidemiologia, realizada na Iugoslávia, em 1971, o tema havia sido precisamente “Usos de la Epidemiología en la Planificación de los Servicios de Salud”. Lamentava, entretanto, que a América Latina apresentava “una cierta debilidad” neste aspecto, tomando-se em conta o desenvolvimento de experiencias concretas e o andamento do programas (OPS/OMS: 1984, 4).

Neste mesmo seminário, Guilherme Rodrigues da Silva lembrava que no Plano de Ação para as Américas, de 1982, a OPS havia definido três metas gerais para instrumentalizar a estratégia Saúde para Todos no ano 2000, a saber: a) o estabelecimento de setores e programas prioritários; b) o desenvolvimento da infra-estrutura de saúde e c) a criação de **sistemas de avaliação e auditoria** (OPS/OMS: 1984, 63).

A partir do Seminário de Buenos Aires, passou-se a enfatizar a aplicação da Epidemiologia na “avaliação dos serviços, programas e tecnologias”, entendida como o “quarto” uso desta disciplina (MS/FNS/CENEPI: 1992, 3; Hartz: 1993).

Este novo modelo aplicativo da Epidemiologia, voltado para a avaliação do impacto dos serviços, programas e tecnologias de saúde, deu origem ao surgimento de “novas áreas de pesquisas, como a de Pesquisa em Serviços de Saúde (“Health Services Research”)” (Novaes: 1995, 49) Nesta linha, destacamos a produção científica dos pesquisadores vinculados ao "Groupe de Recherche Interdisciplinaire in Santé (GRIS)", da Univesidade de Montreal, no Quebec: Clemenhagen, Champagne, Battista, Dussault e outros, já citados na seção referente às metodologias de avaliação em saúde.

Da mesma forma que nas Ciências Sociais, o desenvolvimento da temática da avaliação no campo da Epidemiologia refletiu-se no plano institucional, conforme vemos na breve cronologia, feita por NOVAES⁶¹, dos eventos relativos à avaliação em saúde nos EUA:

1968 - cria-se o "Instituto Nacional de Pesquisa em Serviços de Saúde", que passa a ser o eixo de definição no campo da Avaliação em Saúde;

1977 - cria-se, neste Instituto, o "Escritório de Aplicações Médicas e de Pesquisa";

1989 - cria-se a "Agência de Política e Pesquisa em Atenção à Saúde", extinguindo-se o Instituto.

Estes órgãos teriam estimulado o uso das grandes bases de dados (secundários) nas pesquisas de avaliação em saúde, que foi facilitado pela disseminação da informática. A ênfase na utilização dos dados produzidos pelos serviços de saúde teria provocado um deslocamento dos estudos de "eficácia" (randomizados, controlado), com base nos pacientes atendidos na clínica, para os estudos de "efetividade", com base populacional. Entretanto, teria havido um certo recuo destas proposições, quando os Epidemiologistas passaram a advertir sobre os problemas no uso de dados secundários, incluindo nesta crítica até mesmo o aproveitamento dos prontuários médicos, pois estas bases não levavam a "dados verdadeiros". Voltou-se então a recomendar que as pesquisas em serviços de saúde fossem realizadas a partir de dados primários, produzidos no próprio contexto da investigação.

O desenvolvimento das pesquisas aplicadas à avaliação de serviços, na área da Epidemiologia, levou ao surgimento da figura do avaliador como “um profissional que analisa

⁶¹Profª H. Maria D. Novaes, da Fac. de Medicina da USP/SP, na apresentação do tema "Epidemiologia na Avaliação em Serviços de Atenção Médica: novas tendências na pesquisa", na Comunicação Coordenada 47, do III Congresso Brasileiro de Epidemiologia, realizado em Salvador, Bahia, de 24 a 28 de abril de 1995.

a qualidade da assistência como um fator de proteção/risco inserido entre os determinantes de saúde” (Hartz: 1996).

Em nosso país, a situação é bastante diversa. Segundo SILVA & FORMIGLI , existe "uma grande quantidade de informações registradas rotineiramente pelos serviços que não são utilizadas nem para a análise da situação de saúde, nem para a definição de prioridades, nem para a reorientação de práticas". Talvez devido a esta ausência de uma "prática cotidiana” de avaliação nos serviços de saúde no Brasil, as autoras consideram que o campo “enfrenta dificuldades metodológicas e operacionais não respondidas completamente no plano da investigação" (Silva & Formigli: 1994, 87, 81),

Porém, consideram que o quadro estaria se modificando nas últimas décadas, e para exemplificar citam alguns estudos, que classificam dentro do campo da avaliação em saúde, produzidos desde o ano de 1973 (Silva & Formigli: 1994, 87). Verificamos que são ao todo quatorze trabalhos citados, o que não chega a ser muito, considerando que trata-se de um artigo de revisão sobre o tema. Entretanto, observa-se uma diversidade das abordagens temáticas que indica a tendência de crescimento deste campo.

A análise da temática da "Avaliação" em eventos recentes na área de Epidemiologia confirma esta impressão. Nos três Congressos Brasileiros de Epidemiologia promovidos pela ABRASCO - em Campinas, 1990, em Belo Horizonte, 1992 e em Salvador, 1995 - percebemos uma tendência de crescimento da produção científica na área.

Nos anais do I e do II Congresso, verificamos que na temática "Serviços de Saúde - Avaliação de Serviços", foram classificados vinte e três (23) trabalhos no primeiro congresso e trinta e quatro (34) no segundo. Em termos proporcionais, houve uma discreta ampliação no volume de trabalhos: 6,08% no primeiro (de um total de 378 trabalhos apresentados) e 6,43% no segundo congresso (total de 528 trabalhos). No III Congresso Brasileiro de Epidemiologia, quarenta e um (41) trabalhos foram classificados na área de "Avaliação em Epidemiologia" (considerando-se apenas os “posters”). Porém, devido talvez ao grande aumento no número de trabalhos apresentados - somente “posters” foram 1.145 - houve um redução proporcional da área temática (3,6%).

Também observamos no III Congresso (abril/95), um grande volume de apresentações de temas relacionados ao assunto, com uma afluência de público bastante grande, o que

provavelmente denota um interesse pelo assunto. Nas áreas de “Avaliação” e “Qualidade”, foram sete comunicações coordenadas, três painéis e duas palestras. Também na área de “Sistemas de Informação” para o planejamento e a gerência, percebe-se um crescimento em relação aos dois congressos anteriores, tendo sido apresentados uma mesa redonda, três comunicações coordenadas, dois painéis e quatro palestras, além de uma oficina sobre o tema.

Considerando que a maior parte dos participantes destes Congressos está vinculada aos serviços de saúde, é de se esperar que o interesse pela temática da avaliação possa levar a uma mudança da situação desalentadora da avaliação em saúde no país.

A produção de conhecimentos sobre a avaliação na área da Epidemiologia pode ser resumida, de uma forma geral, nos seguintes aspectos:

- a avaliação dos estados de saúde, que mais recentemente vem evoluindo do tradicional “diagnóstico de saúde” para os estudos que enfatizam as “desigualdades” destes estados entre os diferentes grupos populacionais⁶²;

- a avaliação como instrumento para nortear a organização dos serviços de saúde e os modelos assistenciais, sendo que os estudos mais recentes enfatizam a “equidade” na oferta de serviços de saúde⁶³;

- a avaliação para informar o processo de decisão gerencial (com o desenvolvimento correlato da área de “Sistemas de Informação em Saúde”)⁶⁴

- a avaliação de qualidade da assistência prestada, incluindo os tradicionais estudos sobre o “cuidado médico”, na linha de Donabedian, e os estudos sobre padrões de acreditação de hospitais⁶⁵;

- a avaliação de tecnologias em saúde;⁶⁶

- a avaliação de programas de saúde.⁶⁷

⁶²Ver: VICTORA, Cesar G.; BARROS, Fernando C. & VAUGHAN, J. Patrick. *Epidemiologia da Desigualdade*. São Paulo, Ed. Hucitec, 1988.

⁶³Ver: KADT, Emanuel de & TASCA, Renato. *Promovendo a Equidade. Um novo enfoque com base no setor saúde*. São Paulo-Salvador, Ed. Hucite/Cooperação Italiana em Saúde, 1993.

⁶⁴Ver: MORAES, Ilara Hammerli Sozzi. *Informações em Saúde: da prática fragmentada ao exercício da cidadania*. São Paulo-Rio de Janeiro, Ed. Hucitec/Abrasco, 1994.

⁶⁵Ver: NOVAES, Humberto de Moraes & PAGANINI, José Maria. *Padrões e Indicadores de Qualidade para Hospitais (Brasil)*. Washington, D.C., OPAS, 1994, Edição Especial para UNIMED do Brasil, Congresso Latino-Americano de Hospitais.

⁶⁶Ver: MARQUES, Marília Bernardes. *Ciência, Tecnologia, Saúde e Desenvolvimento Sustentado*. Série Política de Saúde n° 11, Rio de Janeiro, FIOCRUZ/Núcleo de Estudos em Ciência e Tecnologia(NECT), 1991.

Para encerrar, verificamos que a temática da avaliação na área da Epidemiologia está basicamente localizada no “início” do processo de atenção à saúde, isto é, no diagnóstico dos estados de saúde, e no “final”, isto é, na avaliação do impacto dos serviços prestados e dos programas implementados, tanto porque estas são as duas ordens de objeto da Saúde Pública/Saúde Coletiva, e a Epidemiologia é a disciplina mais especificamente ligada ao processo saúde-doença neste campo, tanto porque, como vimos anteriormente, a metodologia epidemiológica constitui-se, essencialmente, dos mesmos componentes da Pesquisa Avaliativa.

⁶⁷Ver: HARTZ, Zulmira M.A. *Avaliação do Programa de Saúde Materno-Infantil (PMI) em uma área do Nordeste do Brasil*. Tese de Doutorado, Montreal (Canadá), mimeo, 1993.

DISCUSSÃO E CONCLUSÕES

Uma vez que neste trabalho propusemo-nos a fazer apenas uma reflexão sobre a temática da avaliação no campo da Saúde Coletiva, não caberia apresentar “conclusões” propriamente ditas. O que pretendemos fazer aqui é retomar e aprofundar algumas idéias, que julgamos fundamentais em nosso trabalho, e que podem ser tomadas como “teses” que levantamos, mas que talvez tenham ficado diluídas no desenvolvimento do tema.

O primeiro ponto é relembrar o que coloca Demo: “a avaliação é um dos desafios científicos que mais escancaram os limites da ciência” (Demo: 1992, 70-71). Por ser uma forma de pesquisa social aplicada, a avaliação põe em evidência todos os dilemas e impasses da pesquisa sobre os fenômenos sociais.

As questões fundamentais da pesquisa em avaliação são as mesmas que colocam-se, em um plano epistemológico, para a prática científica em geral, e especialmente para as ciências sociais. Estas questões podem ser resumidas na discussão sobre a cientificidade ou não de estudos que busquem contemplar os aspectos subjetivos da realidade, através de procedimentos qualitativos de coleta de dados, discussão esta que é fruto da tradição positivista da ciência moderna. Neste sentido, a avaliação coloca, de forma crucial, a problemática das ciências sociais em geral.

A pretensão de avaliar “isentamente” leva à busca incansável por critérios quantitativos, supondo que estes são mais “objetivos”, portanto, neutros. Entretanto, a própria tarefa de avaliar consiste essencialmente em um procedimento subjetivo: emitir juízo de valor sobre algo. Ainda que este valor possa ser transformado em números, a noção de “juízo” remete, sempre, à subjetividade do juiz/avaliador.

A ambigüidade do conceito e da tarefa da avaliação está inscrita na própria etimologia do termo. O conceito de avaliação - atribuir juízo de valor a algo - comporta uma subjetividade intrínseca que permeia a prática e a produção de conhecimentos no campo, embora todo o esforço da avaliação consiste na busca da objetividade.

Esta ambigüidade pode ser assim explicada: se avaliar é fazer juízo de valor, é, inevitavelmente, um procedimento subjetivo; porém, para avaliar algo é necessário “medir”

este valor, e a tarefa de medir implica em determinar indicadores e formular instrumentos objetivos de medida.

A avaliação é assim vista como um processo que tem duas dimensões: técnica e social. A primeira, reproduz o esforço de objetivação característico da ciência; a segunda, está carregada padrões de valores construídos, subjetivamente, na vida social.

A busca da objetividade leva à quantificação e à experimentação, típicas das metodologias quantitativas. A subjetividade do campo dos valores, entretanto, só pode ser apreendida qualitativamente, pois refere-se às representações sociais.

Esta aparente obviedade, ao que tudo indica, está sendo reconhecida na tendência atual da pesquisa em avaliação de adotar modelos de estudo que integrem as metodologias quantitativas e qualitativas, tanto através da pesquisa sintética de Contandriopoulos e cols. (1994), onde incluem-se os estudos de caso, conforme a concepção de Yin (1994), quanto através da metodologia denominada de triangulação, na qual são combinados diferentes métodos e procedimentos de coleta e análise de dados em um mesmo estudo.

Disto resulta que os estudos avaliativos implicam na utilização de diferentes critérios de validação do conhecimento produzido, o que torna-os ainda mais complexos. A validade é um aspecto fundamental da avaliação, a ponto de Suchman (1967) considerá-las como conceitos paralelos. A medida é a base do juízo avaliativo, por isto, os estudos em avaliação enfatizam tão fortemente a questão da validade dos dados.

Este é o primeiro grande recorte de nosso trabalho, de caráter metodológico: os valores sociais, que fundamentam as intervenções, são aspectos subjetivos da realidade, logo, são apreendidos mais adequadamente através das metodologias qualitativas; a medição dos efeitos destas intervenções, bem como a tentativa de colocar estes valores em uma escala de medida, definindo critérios objetivos de “bom”, requer o uso das metodologias quantitativas.

Entretanto, este recorte é feito para logo desfazer-se. Uma vez que os campos disciplinares comportam-se, geralmente, como “feudos” de uma ou outra metodologia, ao integrar em seus estudos tanto os procedimento quantitativos como qualitativos, a avaliação coloca-se essencialmente como uma área de pesquisa “indisciplinar”.

A tarefa de avaliar implica na necessidade de recorrer ao instrumental teórico-metodológico de diferentes áreas disciplinares não somente pelo seu duplo caráter subjetivo-objetivo, mas também pela diversidade e complexidade dos objetos da avaliação em saúde.

Coloca-se então o segundo recorte deste trabalho, que refere-se às duas naturezas dos objetos da avaliação em Saúde Coletiva: os estados de saúde e as intervenções em saúde.

O objeto da avaliação no campo da Saúde Coletiva pode estar em qualquer ponto de um esquema conceitual no qual, partindo-se da avaliação de uma situação inicial - o diagnóstico de saúde - encontra-se justificativa para fazer uma intervenção em saúde - que requer uma avaliação continuada durante sua execução - cujos efeitos sobre a situação inicial serão observados, para verificar se os objetivos desejados foram alcançados, isto é, a melhora no estado de saúde - avaliação final. Este esquema lógico-conceitual do “círculo avaliativo” tem origem na Teoria de Sistemas, onde a idéia central é a noção de retroalimentação, como subsídio à decisão sobre as intervenções.

A avaliação, portanto, não enfoca apenas cada um destes objetos isoladamente, mas coloca-se como um procedimento fundamental para vincular os dois: as necessidades de saúde, identificadas no diagnóstico, e as intervenções de saúde, moldadas para atender estas necessidades, visando modificar o estado de saúde inicial. Ou seja, mais uma vez, aquilo que está separado é integrado pela avaliação.

O terceiro recorte do estudo, que entrecruza o dos objetos da avaliação em saúde, refere-se à tradicional divisão dos níveis individual e coletivo, que tem fundamentado a separação entre as práticas clínicas e as de Saúde Pública. Vimos que esta separação é artificial, pois, na realidade concreta das intervenções em saúde, é necessária a visão da integralidade do processo saúde-doença, que expressa-se tanto no nível individual quanto no nível coletivo, os quais devem ser tomados como níveis de análise, e não como divisão intransponível entre práticas.

Enfim, a natureza “indisciplinar” da avaliação é reforçada pela própria natureza do campo da Saúde Coletiva, por duas razões: primeiro, porque a Saúde Coletiva é, em si mesma, um campo interdisciplinar; visto que a integralidade do fenômeno saúde-doença somente pode ser apreendida através dos métodos oriundos de várias disciplinas; segundo, porque os fenômenos tomados como objeto de avaliação neste campo são de grande amplitude e

complexidade, exigindo conhecimentos interdisciplinares para que possam ser abordados e compreendidos em sua totalidade.

A discussão do sujeito da avaliação, por sua vez, traz à tona uma série de dilemas que acompanham o campo da Saúde Coletiva: quem define as necessidades de saúde que justificam a intervenção? quem define os objetivos selecionados para formular a intervenção? quem avalia os resultados, e com que critérios?

O contexto da pesquisa em avaliação é altamente conflitivo. Há conflitos de interesse na definição dos objetivos das intervenções, há conflitos de interesse na definição dos critérios de avaliação destas intervenções.

O que está por trás destes dilemas é a discussão de quem deve conduzir a avaliação: o sujeito da intervenção ou o avaliador? Em outras palavras, o sujeito da avaliação deve ser o sujeito da ação, que é o destinatário primordial da avaliação; os beneficiários da ação, destinatários finais; ou um avaliador externo à tudo? O sujeito da ação deve ser o sujeito da avaliação?

A compreensão do processo avaliativo como um processo de aprendizagem, que propicia a produção de conhecimentos sobre as intervenções organizadas, indica que a avaliação participativa, ou a avaliação emancipatória, ainda que mais demoradas, parecem ser o melhor caminho para que a avaliação cumpra sua finalidade - influir na ação. Caso contrário, a avaliação pode resumir-se a um “controle cibernético” ou a um “exercício autoritário do poder de julgar”, entendida como um processo externo, fora da responsabilidade do sujeito da ação.

Ao tratar dos sujeitos da avaliação, entramos em uma polêmica que vem, de longa data, conformando-se no campo da Saúde Coletiva: a questão das “reais necessidades de saúde” *versus* “demandas por serviços de saúde”.

Para boa parte dos planejadores e epidemiologistas, as crescentes demandas da população por serviços de saúde seriam determinadas principalmente pela propaganda que induz o consumo de serviços, dentro de um processo de medicalização da sociedade. As reais necessidades de saúde da população, por sua vez, somente poderiam ser identificadas através de diagnósticos epidemiológicos, que como vimos, nada mais são do que avaliações. Ou seja, caberia aos técnicos, munidos de instrumental epidemiológico “objetivo”, definir quais as

prioridades em saúde. Isto implica em considerar que a avaliação que a população faz de suas necessidades de atenção em saúde, as quais geram as demandas por serviço, não correspondem às suas necessidades “reais”, porque seriam determinadas por questões “subjetivas”. Novamente, identificamos por trás deste debate, a questão subjetivo-qualitativo *versus* objetivo-quantitativo, onde a objetividade da dimensão técnica coloca em plano secundário a subjetividade dos valores sociais.

Nesta discussão, vemos que, contraditoriamente, atribui-se aos técnicos, e não à população, o papel de conhecer suas “reais necessidades de saúde”. A demanda da população por serviços de assistência curativa, muitas vezes, não é considerada legítima, não apenas pelo reconhecimento de que influem nesta demanda a medicalização da sociedade e fatores de consumo, definidos pelas regras de mercado, mas também, em alguns momentos, como se a assistência curativa não fosse legítima em si mesmo.

Quanto à seleção das prioridades de ação, vemos que, na literatura latino-americana, o planejamento em saúde apresentou, historicamente, duas racionalidades na definição dos critérios de seleção das necessidades identificadas no diagnóstico: a racionalidade técnica (ou normativa) e a racionalidade estratégica (ou política).

Na corrente conhecida como planejamento normativo, cujo documento pioneiro ficou conhecido como método CENDES/OPS (1965), a racionalidade técnica fundamental tinha origem na área econômica, que estabelecia como critério de seleção de prioridades de ação a “eficiência”, ou seja, a alternativa de ação que proporcionasse mais benefícios ao menor custo. Os critérios epidemiológicos entravam, na verdade, apenas como um instrumento para determinar a magnitude dos problemas de saúde e a sua vulnerabilidade à intervenção. O terceiro critério - a transcendência do dano à saúde - por remeter à questão dos valores sociais, era, corretamente, deixado para ser discutido em cada contexto social específico.

Na corrente conhecida como planejamento estratégico, abandona-se a Epidemiologia em função de seus limites, e volta-se para a política. O critério de definição de prioridades passa a ser a força de pressão política que os grupos de interesse exercem sobre os governantes do setor saúde, conforme vemos no documento “Formulación de Políticas de Salud” (1975). A idéia da estratégia política evolui para a noção de raciocínio estratégico na definição de prioridades de ação, cuja finalidade, para Carlos Matus, é acumular poder e construir maior

capacidade de governo, e para Mario Testa, dar maior poder de pressão política às classes economicamente subalternas, tradicionalmente desfavorecidas na partilha dos recursos.

Por fim, é a Epidemiologia que volta-se para o Planejamento, propondo a adoção do critério de risco para a identificação dos grupos populacionais mais expostos, aos quais dever-se-ia proporcionar mais atenção. O pressuposto deste critério é que a desigualdade dos estados de saúde exige a equidade na oferta de serviços, ou seja, mais atenção para quem mais precisa.

Em resumo, no campo da Saúde Coletiva, tradicionalmente a Epidemiologia tem buscado “argumentos científicos”, através da tentativa de formular diagnósticos “objetivos” (entenda-se quantificados) das “reais necessidades da população”. Entretanto, diante desta mesma população e do contexto político no qual são realmente definidas as prioridades de saúde, verifica-se que este argumento somente tem se justificado perante sua própria lógica. As correntes do Planejamento em Saúde latino-americano, de um modo radicalmente diferente, tem proposto abandonar esta pretensão de objetividade, partindo direto para um “diagnóstico situacional”, que nada mais é do que um diagnóstico dos interesses políticos dos diferentes grupos que tem acesso à discussão sobre a definição de prioridades. Ou seja, propõe-se assumir radicalmente a subjetividade dos juízos de valor, tornando o processo de seleção de prioridades totalmente subjetivo, no qual a simples existência da “vontade de algo”, ou interesse, é justificativa suficiente para uma intervenção coletiva. Ainda que esta vontade esteja embasada em um ideal de bom, de justiça, é de questionar-se a forma como se propõe que esta vontade prevaleça sobre o conjunto das pessoas - através do acúmulo de poder, ou seja, do argumento da força.

Por trás destas discussões, o que está de fato em questão é a intencionalidade da ação humana, ou seja, o motivo real que está atrás dos argumentos trazidos a público para justificar uma intervenção em Saúde Coletiva. É justamente neste ponto que observamos a principal omissão da literatura em avaliação - a falta de uma teoria da ação social, que permita discutir e avaliar estes motivos. Consideramos que esta falta é muito importante, pois o objeto privilegiado da avaliação é justamente a ação social organizada com vistas à mudança.

A literatura em avaliação tem omitindo-se em tratar dos valores subjacentes à determinação dos objetivos das intervenções, porque estes pertenceriam à esfera dos formuladores de políticas e tomadores de decisão. Esta postura implica em alienar-se da

própria finalidade e do próprio contexto do processo avaliativo, totalmente imerso no campo subjetivo dos valores. Daí, talvez, a fuga para a “medição”, que, como vimos, não se confunde com avaliação.

Neste sentido, caberia levantar a proposta de Jürgen Habermas, conforme a vemos em RIVERA (Rivera: 1991. 10-35). Segundo este último, Habermas distingue três tipos básicos de racionalidade na ação social: a normativa (ou instrumental, tecnocrática), a estratégica (ou política) e a comunicativa, embasada na noção de direito e de solidariedade.

De uma certa forma, a proposta tradicional de ação programática da Epidemiologia, no campo da Saúde Coletiva, pode ser enquadrada na racionalidade tecnocrática, na medida em que subordina o “outro” - a população - à racionalidade do técnico. A proposta do Planejamento Estratégico em Saúde é enquadrada na racionalidade estratégica, que coloca o outro, isto é, aquele que tem uma proposta alternativa de ação, como “inimigo”. Em ambas, a ação é orientada para o “êxito” ou “sucesso”.

A proposta de Habermas parte do pressuposto que o outro não é incompetente nem é um inimigo a ser derrotado, mas um interlocutor que deve ser respeitado como “sujeito de direitos”, com o qual temos que agir “comunicativamente”, isto é, buscando, através do diálogo, validar nossos argumentos perante o mesmo, e chegar assim à uma proposta de ação. Nesta “terceira via”, a ação é orientada pelo entendimento, através de um processo comunicativo onde os atores tentam “harmonizar” seus planos de ação (Rivera: 1991, 16), discutindo seus argumentos, em um diálogo onde o “opositor” é substituído pelo “outro”, igual em direitos, e através do qual se desnuda o que está por trás do ideal de “bom” de cada um, ou seja, seus valores e interesses.

Entendemos que a proposta da “quarta geração” de avaliadores, de negociar os resultados da avaliação com todos os interessados, bem como as propostas de avaliação emancipatória e avaliação participativa, vem ao encontro desta nova visão da ação do homem, menos tecnocrática e menos belicosa, que propõe o diálogo entre sujeitos de direito.

Ainda que carente de uma teoria da ação social, a finalidade essencial da avaliação é produzir conhecimentos sobre a ação social organizada, ou seja, as intervenções. A noção ampliada de “programa” inclui diferentes níveis de intervenção, que caracterizam objetos de avaliação com diferentes amplitude e complexidade, na área da Saúde Coletiva.

O tipo de avaliação que vai ser feita, e os métodos de pesquisa a ser utilizados, dependem muito da amplitude e complexidade dos objetos de estudo, mas, principalmente da finalidade da avaliação. Os objetivos estratégico, formativo e somativo correspondem, de certo modo, à avaliação antes, durante ou final, respectivamente.

Tradicionalmente, os objetivos da avaliação das intervenções eram dois: o objetivo formativo, onde o foco da avaliação são a estrutura e os processo, e que corresponde, de certa forma, à avaliação durante a execução da intervenção, ou avaliação contínua; e o objetivo somativo, que enfoca os resultados, e que corresponde à avaliação ao final da intervenção.

A inclusão mais recente do objetivo estratégico indica, a nosso ver, um disposição em abandonar a tradicional omissão dos estudos avaliativos em relação ao questionamento dos objetivos das intervenções, pois coloca-os como foco da avaliação, constituindo-se, desta forma, como um avaliação inicial, que indica a pertinência ou necessidade da intervenção.

Esta compreensão levou-nos a estabelecer uma analogia entre a “análise estratégica” (Contandriopoulos et al.: 1994, 24), a análise de pertinência (Kisil&Chaves: 1994, 105) e a “análise de situação” do Planejamento Estratégico Situacional (Matus: 1989). A diferença é que, enquanto nas duas primeiras a finalidade essencial é avaliar as condições que tem a intervenção de modificar o problema, nesta última, o objetivo fundamental é verificar as condições de “poder” do ator social para implementar a ação que está se propondo a fazer.

Na pesquisa avaliativa, a análise estratégica é um procedimento técnico, que pressupõe a transparência dos objetivos. No planejamento estratégico, a análise de situação é um procedimento político, que parte da visão do outro como um “inimigo”, e pressupõe que o sucesso da ação está fundamentado no “segredo” de seus verdadeiros objetivos, como forma de garantir o “ataque” mais certo, impedindo a implantação de uma proposta de ação alternativa.

O entendimento de que os problemas sociais são mal estruturados é a base na qual baseia-se a ação estratégica, cujo pressuposto é que a realidade funciona como um “sistema”, ou uma “rede”, onde os problemas estão inter-relacionados. A concepção sistêmica do “círculo avaliativo”, no qual o processo avaliativo pode iniciar em qualquer ponto, é semelhante à formulação de Matus (1989) de que os “momentos” do planejamento podem ser retomados em qualquer fase, de forma recorrente. A ação normativa, ao contrário, pressupõe, grosso modo, a

linearidade dos fenômenos sociais, que permitiria estabelecer um modelo causal, relativamente bem definido, entre as variáveis que influem no problema.

No campo da Saúde Coletiva, identificamos duas grandes linhas discursivas através da qual são tratadas as questões da avaliação em saúde - a Avaliação de Programas e a Avaliação de Qualidade.

Vimos que a Avaliação de Qualidade desenvolveu-se originalmente no setor industrial, como uma forma de controlar a qualidade dos produtos (bens materiais), através de mecanismos de controle do processo de produção. A Avaliação de Programas, por sua vez, historicamente, está voltada para a avaliação dos programas sociais públicos, oriundos das políticas estatais de bem-estar social, onde encontra-se a Saúde Pública.

Procuramos demonstrar que estes discursos, ao serem incorporados pelo campo da saúde, voltam-se, respectivamente, para o nível individual da prestação de cuidados de saúde e para o nível coletivo das intervenções. No primeiro caso, por influência do pensamento de Avedis Donabedian, para quem a qualidade era essencialmente um atributo do cuidado médico. No segundo caso, por uma visão de que a Saúde Pública deveria estar voltada exclusivamente para a avaliação da “ação programática”, em nível coletivo.

Donabedian refletia, provavelmente, a influência da noção de garantia de qualidade conforme havia sido formulada, originalmente, no setor industrial, onde a qualidade do “produto” (bem material) era garantida através do controle do processo de produção. Embora não pretendesse limitar seus estudos à prática clínica, para ele, a melhora no estado de saúde era um “produto” do cuidado médico, por isto, este deveria ser o objeto privilegiado da avaliação.

A proposta de uma “estrutura integrada de avaliação em saúde”, formulada por autores canadenses, busca integrar estas duas “raízes” da avaliação em saúde. Destacamos dois aspectos no que denominamos de “discurso integrado” sobre a avaliação em saúde.

O primeiro aspecto é que, ao deslocar a questão da qualidade do âmbito da garantia de qualidade, conduzida nas auditorias médicas, para o âmbito da responsabilidade do gerente em saúde, esta proposta passa a tratar a qualidade como um atributo do conjunto dos processos de trabalho que contribuem para a oferta de serviços, e não apenas como o resultado isolado das

práticas clínicas. Em função disto, os autores propõe que as análises devem processar-se sempre em dois níveis: o nível individual e o nível agregado.

A qualidade passa a ser referida à organização dos serviços de saúde como um todo, refletindo provavelmente a influência da teoria da Gestão de Qualidade Total. A avaliação de qualidade deixa de ser, então, a identificação dos desvios no desempenho dos profissionais, através da comparação com padrões previamente estabelecidos, e a avaliação de programas deixa de ser meramente uma comparação dos resultados obtidos com as metas planejadas, e ambas passam a ser o resultado de uma negociação sobre as metas e dos desempenhos possíveis de alcançar diante da condição concreta dos serviços. A busca de qualidade é vista essencialmente um processo de aprendizagem coletiva sobre os processos de trabalho.

Neste contexto, o papel do avaliador aproxima-se da concepção da “quarta geração” de avaliadores, onde os resultados da avaliação são negociados entre todos os interessados, e onde a avaliação emancipatória busca constituir-se “num processo e num projeto em que avaliador e avaliando buscam e sofrem uma mudança qualitativa” (Demo: 1987, Prefácio de Moacir Gadotti).

A principal vantagem desta proposta para o contexto do SUS é que reforça a visão da qualidade como uma questão global do sistema de prestação de serviços, onde a qualidade é uma responsabilidade dos gerentes e o juiz final é o cidadão usuário.

A principal crítica à esta nova forma de tratar a qualidade é o abandono da “imagem ideal” dos resultados que se quer alcançar, substituindo-a pela “imagem do possível” diante das condições dadas, o que, de certa maneira, implica em uma aceitação e acomodação ao *status quo*, contrariando o espírito de crítica e desconformidade que é a essência da avaliação⁶⁸.

O segundo aspecto que destacamos na proposta de integração dos canadenses é que a mesma resulta em uma nova divisão, entre a “Avaliação Normativa” e a “Pesquisa Avaliativa”.

A chamada Avaliação Normativa está dirigida para os componentes de um programa - problema, objetivos, recursos, serviços realizados, efeitos - buscando avaliar como cada um

⁶⁸Esta crítica foi levantada no debate após o painel sobre o tema “A Epidemiologia na Avaliação dos Serviços de Saúde: a discussão da Qualidade”,

deles comporta-se à luz de normas definidas previamente. Está estreitamente vinculada à gerência, daí também ser designada de “avaliação administrativa”, sendo que a maioria das práticas de avaliação podem ser enquadradas nesta linha. É um tipo de avaliação que pode e deve ser feita pelos próprios responsáveis pela intervenção.

A chamada “Pesquisa Avaliativa” está voltada para as relações entre os componentes dos programas, buscando estabelecer como um deriva do outro e, acima de tudo, no caso da saúde, estabelecendo um eixo entre os estudos sobre os estados de saúde e sobre as intervenções em saúde. Esta linha está vinculada à idéia da avaliação como prática de investigação científica, sendo sua proposição fundamental a aplicação do método científico aos procedimentos de avaliação. Pela complexidade de suas análises, a Pesquisa Avaliativa é melhor conduzida por um profissional em avaliação.

Poderíamos afirmar que a Avaliação Normativa corresponde à uma visão “anatômica”, estática, estrutural dos componentes do programa, enquanto que a Pesquisa Avaliativa corresponde à uma visão “funcional”, dinâmica das relações entre estes componentes. A primeira corresponde ao que Aguilar&Ander-Egg (1994: 37, 49) denominam de “avaliação formal”, que enfoca a coerência interna entre os componentes do programa, buscando verificar se o que está sendo realizado corresponde ao programado; e a segunda corresponde à “avaliação substantiva”, que enfoca a coerência externa do programa, isto é, busca verificar se as mudanças observadas na situação problema são resultantes da intervenção, vinculando o programa com as necessidades que se propõe a resolver.

Entendemos que existe uma certa correspondência entre esta tipologia e a divisão que, historicamente, estabeleceu-se no campo da avaliação, a partir dos primórdios da Pesquisa Avaliativa, onde distinguiu-se duas correntes fundamentais:

- a pesquisa avaliativa “clássica”, que enfocava o impacto, após a implantação do programa, e que corresponderia a atual “Pesquisa Avaliativa”, que tem objetivos somativos;

- a pesquisa avaliativa “formativa”, que enfocava os processos, durante a implantação do programa, e que corresponde à atual “Avaliação Normativa”, que obviamente tem objetivos formativos.

durante o I Encontro Paulista de Epidemiologia, realizado na Faculdade de Medicina da USP, São Paulo, de 9 a 11 de março de 1995.

Como o enfoque nos processos corresponde à forma privilegiada da avaliação de qualidade - no caso da saúde, tradicionalmente, o processo de trabalho médico, em nível individual -, e o enfoque nos resultados corresponde à forma tradicional da avaliação de programas - as intervenções em nível coletivo -, a abordagem integrada une processos e resultados, bem como individual e coletivo, ao unir avaliação de qualidade e avaliação de programas.

De qualquer forma, a distinção entre Avaliação Normativa e Pesquisa Avaliativa parece que estabeleceu uma linha divisória fundamental no campo da avaliação. Quando falamos da primeira, estamos no campo do Planejamento; quando falamos da segunda, estamos em um campo de pesquisas multidisciplinar.

Finalmente, chegamos aos aspectos históricos. O desenvolvimento histórico do campo da avaliação revela que, ao incorporar a metodologia científica das Ciências Sociais, a avaliação desprende-se do Planejamento e passa a constituir-se como um campo específico de estudo. Inicialmente, reproduzindo o próprio processo de planejamento e programação, a avaliação restringe-se a medir os efeitos da ação para verificar o grau de alcance das metas (estudos baseados em objetivos), constituindo-se essencialmente como “ciência social aplicada” ao planejamento, com a finalidade essencial de subsidiar as decisões sobre a continuidade, melhoramento ou interrupção dos programas sociais. Posteriormente, a avaliação evolui como uma forma de estudar a ação visando a mudança social, concebida de forma ampla, passando a ser vista como um instrumento para refinar teorias sobre a ação social organizada, cujo objetivo fundamental é contribuir para o conhecimento e formulação de teorias.

Coerentemente com o que aponta a literatura latino-americana, a temática da avaliação em Saúde Coletiva é ainda incipiente, no Brasil, tanto do ponto de vista das práticas em serviço quanto da produção de conhecimento, evidentemente, quando se compara com o que relata a literatura anglo-francesa. O campo da avaliação está bem desenvolvido e institucionalizado em outros países, especialmente nos EUA, Canadá, Inglaterra e França, onde existem associações profissionais de avaliadores, bem como periódicos e publicações específicas sobre o tema.

Como toda produção de conhecimento é norteada por um interesse concreto, poderíamos atribuir este desenvolvimento relativamente incipiente à própria maneira como o Estado brasileiro tem conduzindo, historicamente, suas intervenções sociais. Boa parte da literatura estrangeira em avaliação foi produzida como subsídio ao desenvolvimento da políticas sociais de combate à pobreza, no caso dos EUA, e no contexto do *Welfare State*, no caso do Canadá e da Europa. De uma forma geral, poderíamos afirmar que a literatura sobre avaliação em saúde é desenvolvida especialmente nos países que desenvolveram sistemas nacionais de saúde, de caráter público.

Ora, sabemos que esta condição de Estado de Bem-Estar Social nunca chegou a existir, de fato, no Brasil, e, de direito, somente após a Constituição Federal de 1988. A atuação do Estado brasileiro tem caracterizando-se pelo clientelismo e autoritarismo - características que evidentemente tornam os governos pouco inclinados a querer fazer ou a aceitar que suas intervenções sejam submetidas à avaliação, pois isto tornaria transparente estes mecanismos na forma de conduzir a “coisa pública”. Especificamente no campo da saúde, o corporativismo de boa parte das categorias profissionais, lideradas e inspiradas pela postura da categoria médica, tampouco demonstra disposição para a avaliação, menos ainda à aceitar uma avaliação externa, o que reforça a forma como as instituições públicas gerenciam a saúde.

A forma clientelista e autoritária como tem atuado o Estado brasileiro certamente é o fator determinante pela qual o campo da avaliação ainda está relativamente pouco desenvolvido em nosso país.

Entretanto, acreditamos que parte da responsabilidade pode ser atribuída a fatores internos do próprio desenvolvimento das Ciências Sociais na América Latina, que refletem-se no campo específico da Saúde Coletiva no Brasil, e dizem respeito à influência do pensamento marxista, o qual, ao recusar o modelo funcionalista da sociologia norte-americana, acabou sendo pouco permeável aos conhecimentos produzidos no campo da avaliação.

Por outro lado, a literatura canadense mais recente sobre avaliação demonstra que estão sendo incorporadas críticas fundamentais que o pensamento marxista faz à visão funcionalista do processo avaliativo. Isto é observado principalmente no conceito ampliado de intervenção, que inclui o próprio contexto, e no surgimento da proposta da análise de implantação, que

contempla o contexto como elemento fundamental a ser avaliado no funcionamento das intervenções.

Em nosso entendimento, o contexto atual de implantação do SUS reproduz algumas condições que estimularam o desenvolvimento da Pesquisa Avaliativa nos EUA, na década de 60. Ao instituir a saúde como um direito de cidadania, a Constituição Federal instituiu também um novo patamar de relacionamento da população com as instituições públicas, onde esta passa a cobrar um desempenho dos serviços de saúde que possa garantir, efetivamente, a plena realização deste direito. Ou seja, a constituição pretende instalar, tardiamente, no Brasil, um Estado de Bem-Estar Social, no qual o cidadão entende como um direito seu a cobrança da efetividade das instituições públicas.

Além disto, como a proposta do SUS é de mudança social, é preciso avaliar os efeitos de sua implantação. O novo modelo assistencial, delineado nas diretrizes do SUS, compete com o antigo modelo de assistência à saúde, de forma semelhante à competição que estabeleceu-se entre os “novos” programas sociais e os “velhos” programas, que Suchman menciona ao historiar o desenvolvimento da Pesquisa Avaliativa nos EUA. Para o autor, os “programas que estavam já estabelecidos levavam a vantagem de ter tradicionalmente um suporte, tanto dos grupos administrativos quanto do público, que já estava acostumado com tais serviços”, de forma que a avaliação cumpria um papel fundamental na validação dos novos programas perante a população, que demandava “provas” de sua efetividade (Suchman: 1967, 6).

Ora, aparentemente, esta é a situação na qual se encontra o SUS atualmente: provar sua efetividade na melhora da assistência à saúde, e, conseqüentemente, nas condições de saúde da população.

É justamente para “provar” a efetividade de novos modelos assistenciais que entendemos ser fundamental o papel da análise de implantação. Este tipo de análise, que pode ser considerada uma nova “versão” da avaliação de processos, parte de uma visão ampliada de programa, que tem duas direções: uma vertical, que ao aprofundar-se inclui diferentes níveis de intervenção na realidade, a partir do pressuposto que qualquer ação organizada para alcançar um fim pode ser avaliada como um programa; e uma direção horizontal, que ao ampliar-se inclui como um “componente” de análise o próprio contexto de implantação do

programa. Embora incorporando a visão “interna” da avaliação de processos, a análise de implantação inclui a externalidade do contexto, permitindo avaliar o funcionamento dos processos de uma determinada modelagem de intervenção em contextos específicos.

Por outro lado, ao focar o funcionamento interno dos programas, a análise de implantação permite contemplar de forma privilegiada a questão da qualidade.

Acreditamos que esta é a outra contribuição fundamental que o campo da avaliação pode dar ao SUS, neste momento, direcionando os estudos para a questão da qualidade na administração dos serviços, programas, planos. É através da avaliação dos processos de gerência em andamento que pode haver um salto de qualidade na prestação de serviços, possibilitando que a avaliação possa ser, de fato, incorporada como um processo pedagógico, assumindo seu caráter “formativo”.

Enfim, supondo que as diretrizes da integralidade, descentralização e participação da população favoreçam a transparência da gerência do SUS, entendemos que o contexto atual seja mais favorável para o desenvolvimento do campo da avaliação.

A avaliação no contexto do SUS deve, em nosso entendimento, considerar três aspectos fundamentais:

- primeiro, que a noção de integralidade da assistência à saúde supera a tradicional separação assistência curativa/assistência preventiva, individual/coletivo, que justificava a oposição entre as práticas clínicas e as de Saúde Pública. Ou seja, a avaliação no SUS deve integrar ambas as práticas;

- segundo, que por caracterizar-se como um sistema em construção, que comporta diferentes modelos de atenção à saúde conforme as peculiaridades loco-regionais, a ênfase nos estudos avaliativos deve ser na análise de implantação destes modelos, pois este tipo de avaliação permite identificar tanto os elementos internos ao funcionamento dos serviços e programas, como os elementos externos, contextuais, que influem na efetividade destes modelos de atenção. Com isto, a avaliação passa a ser o instrumento privilegiado para produzir conhecimento sobre diferentes modelos de atenção à saúde, permitindo determinar sua efetividade tanto em contextos específicos como sua eficácia potencial ;

- terceiro, que o juiz da qualidade do cuidado à saúde não pode mais ser exclusivamente a corporação médica, mas deve ser principalmente o usuário dos serviços, na condição de cidadão com direito à saúde, diretamente ou através das instâncias colegiadas.

Diante da relativa carência de estudos em avaliação voltados para as intervenções de saúde no contexto do SUS, e considerando a “novidade” relativa do campo da avaliação no Brasil, esperamos que este estudo possa contribuir para a sistematização do tema. Mas acima de tudo, esperamos que estar contribuindo para o desenvolvimento de uma “cultura de avaliação” nas instituições públicas, de forma que a atuação do Estado passe a nortear-se pela necessidade de prestar contas à população.

SUMMARY

This study is a review and a reflection about evaluation in Public Health.

There has been an increasing interest on this subject-matter, in Brazil, in recent years, seeming that a new area of knowledge is being constructed in the field of Public Health practices and investigation.

This construction is necessarily built up through the relationship between different subject-matters: firstly, because the knowledge in the area of Public Health is multidisciplinary; secondly, because the objects of study in evaluation in health are so wide and complex that it demands knowledges from different areas to be understood.

This review shows that evaluation research field is recently developed in our country, when compared with scientific production in North America and Europe. There has been, however, an increasing national scientific production, with studies in literature review and even field research studies, with many approaches, that takes objects of investigation in different levels of depth and complexity, ranging from studies in health care through health policy evaluation, including: technological assessment research in health procedures and assistance; health services organization and programs; and health systems planning and management.

This study about evaluation was developed in two sections: theoretical and methodological aspects; and historical background.

The structure of the first section is the same of a research project: purposes, subject-matter, methodology, and study conductors. In the methodology, we discuss great questions on evaluation research: those related to Social Sciences problems - subjectivity *versus* objectivity; qualitative *versus* quantitative methods; study validity; and those related to specific evaluation questions - selection of evaluation model and other typical problems of evaluation research, as scientific rigor *versus* political demands; "availability" of programs and ethical aspects, most of these problems derived from the political context where evaluation process takes place. Two major lines of evaluation in Public Health are identified: quality and program assessment. Besides, the concept and task of evaluation are discussed, and the primary analytical concepts of the evaluative studies: structure-process-effects; efficacy-effectiveness-efficiency; indicator-criteria-standard and quality concept.

In the second section, a review is made on evaluation subject-matter in the three basic areas of Public Health: Social Science, Health Planning and Epidemiology. Evaluation research field is very well developed and institutionalized in some countries, specially Canada, USA and UK, mainly due to the interest of the state administration in evaluating public policies. In Latin America and Brazil, the authoritarian and patronizing pattern of the state governmental policies has restrained the development of evaluation.

The Brazilian National Health System, which is being developed since 1988, tends to favour the development of evaluation research in Public Health.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUILAR, Maria José & ANDER-EGG, Ezequiel. *Avaliação de Serviços e Programas Sociais*. Petrópolis, RJ, Vozes, 1994.
- ALMEIDA FILHO, Naomar de. & ROUQUAYROL, Maria Zélia. *Introdução à Epidemiologia Moderna*. Belo Horizonte/Salvador/Rio de Janeiro, COOPMED Editora-ABRASCO, 1992, 2ª ed.
- ALVES, Paulo César. A perspectiva da análise social no campo da saúde. In: CANESQUI, Ana Maria (org). *Dilemas e Desafios das Ciências Sociais na Saúde Coletiva*. São Paulo-Rio de Janeiro, Ed. HUCITEC-ABRASCO, 1995.
- ANDER EGG, Ezequiel. *Evaluacion de Programas de Trabajo Social*. Buenos Aires, Editorial Hvmánitas, 1984.
- ASSOCIAÇÃO BRASILEIRA DE PÓS-GRADUAÇÃO EM SAÚDE COLETIVA (ABRASCO)/Comissão de Epidemiologia. *Plano diretor para o desenvolvimento da epidemiologia no Brasil*, maio de 1989.
- , *Anais do 1º Congresso Brasileiro de Epidemiologia. Epidemiologia e Desigualdade Social: os desafios do final do século*. Rio de Janeiro, ABRASCO, 1990.
- , *Qualidade de Vida: compromisso histórico da Epidemiologia (Anais do II Congresso Brasileiro de Epidemiologia)*. Belo Horizonte, Editora COOPMED/ABRASCO, 1994.
- , *Resumos do III Congresso Brasileiro de Epidemiologia, II Congresso Ibero-Americano, I Congresso Latino Americano e 1ª Mostra de Tecnologia em Epidemiologia - EPITEC*. Salvador, Bahia, Brasil, 1995a.
- , *II Plano Diretor para o Desenvolvimento da Epidemiologia no Brasil, 1995-1999*. Rio de Janeiro, 1995b.
- BERWICK, Donald M. Continuous Improvement as an ideal in Health Care. *The New England Journal of Medicine*, 320 (1): 53-56, jan/1989.
- BATTISTA, Renaldo N.; CONTANDRIOPOULOS, André-Pierre; CHAMPAGNE, François; WILLIAMS, J. Ivan; PINEAULT, Raynald & BOYLE, Pierre. An integrative framework for health-related research. *Journal of Clinical Epidemiology*, Grã-Bretanha, 42 (12): 1155-1160, 1989.

- BODSTEIN, Regina Cele. Políticas de Saúde e Cidadania: uma contribuição para a discussão sobre Serviços de Saúde. In: SPINOLA, A.W.P. et alli (org). *Pesquisa Social em Saúde*. São Paulo, Cortez Editora, 1992
- BRAGA, José Carlos de Souza & PAULA, Sérgio Goés de. *Saúde e Previdência: estudos de política social*. São Paulo, Ed. HUCITEC, 1982, 2ª ed.
- BRASIL. *Constituição. República Federativa do Brasil*. Brasília, Centro Gráfico do Senado Federal, 1988.
- BRASIL. Ministério da Saúde. Fundação Nacional de Saúde. Centro Nacional de Epidemiologia. *Proposta de Capacitação de Recursos Humanos em Epidemiologia*. Brasília, mimeo, junho de 1992.
- BROOK, Robert H. & LOHR, Kathleen N. Efficacy, Effectiveness, Variations, and Quality: Boundary-crossing Research. *Medical Care*, 23 (5): 710-722, 1985.
- BRUYNE, Paul de; HERMAN, Jacques; SCHOUTHEETE, Marc de. *Dinâmica da Pesquisa em Ciências Sociais*. Rio de Janeiro, Ed. Francisco Alves, 1982, 2ª ed.
- CAMPOS, Anna Maria. Accountability: quando poderemos traduzi-la para o Português? Rio de Janeiro, Revista de Administração Pública, 24(2): 30-50, fev./abr. 1990.
- CHIZZOTTI, Antonio. *Pesquisa em Ciências Humanas e Sociais*. São Paulo, Cortez Editora, 1991.
- CLEMENHAGEN, Carol & CHAMPAGNE, François. Quality Assurance as part of Program Evaluation: Guidelines for Managers and Clinical Department Heads. *Quality Revieww Bulletin - Journal of Quality Assurance* (Joint Commission on Accreditation of Hospitals), 12 (11): 383-87, nov, 1986.
- COHEN, Ernesto & FRANCO, Rolando. *Avaliação de Projetos Sociais*. Petrópolis, RJ, Vozes, 1994.
- COIMBRA, Marcos Antônio. Abordagens Teóricas ao estudo das Políticas Sociais. In: ABRANCHES, S.H.; SANTOS, W.G.; COIMBRA, M.A. *Política social e combate à pobreza*. Rio de Janeiro, Jorge Zahar Editor, 1989, 2ª ed.
- CONASEMS (Conselho Nacional de Secretários Municipais de Saúde). *Sistema Único de Saúde*. Publicações Técnicas - N° 2. Dezembro, 1990.
- CONTANDRIOPOULOS, André-Pierre; CHAMPAGNE, François; DENIS, Jean-Louis; RAYNALD, Pineault. A avaliação na área da saúde: conceitos e métodos. Mimeo, 1996 (no prelo). Versão revisada e traduzida de artigo publicado In: LEBRUN, T.; SAILLY,

- J.C.: AMOURETTI, M.; CREGE, L. (ed.). *L'Evaluation dans le domaine de la santé: concepts et méthodes*. (Atas de colóquio). Montreal, Canadá, 1992, p. 14-32.
- CONTANDRIOPOULOS, André-Pierre; CHAMPAGNE, François; POTVIN, Louise; DENIS, Jean-Louis; BOYLE, Pierre. *Saber preparar uma Pesquisa. Definição, Estrutura, Financiamento*. São Paulo-Rio de Janeiro, Editora HUCITEC/ABRASCO, 1994.
- COSTA, Dina Czeresnia & COSTA, Nilson do Rosário. Teoria do Conhecimento e Epidemiologia: um convite à leitura de Jonh Snow. In: COSTA, D.C. (org.). *Epidemiologia: teoria e objeto*. São Paulo, HUCITEC-ABRASCO, 1990.
- DEMO, Pedro. *Metodologia científica em Ciências Sociais*. São Paulo, Ed. Atlas S.A, 1981.
- . *Avaliação Qualitativa*. São Paulo, Cortez Editora/Editora Autores Associados, 1987.
- . *Pesquisa: Princípio Científico e Educativo*. São Paulo, Cortez Editora/Editora Autores Associados, 1992, 3ª ed.
- DONABEDIAN, Avedis; WHEELER, John R.C.; WYSZEWLANSKI, Leon. Quality, Cost, and Health: An Integrative Model. *Medical Care*, Vol. XX (10): 975-992, 1982.
- DONABEDIAN, Avedis. The seven pillars of quality. *Archives of Pathology Laboratory Medicine*. Ann Arbor, Michigan, Health Administration Press, 1990.
- DONNAGELO, Maria Cecília F. A pesquisa na área da Saúde Coletiva no Brasil - a década de 70. In: ABRASCO. *Ensino em Saúde Pública, Medicina Preventiva e Social no Brasil*. Vol. 2. Rio de Janeiro, abril de 1983.
- DRAIBE, Sônia Miriam. O Welfare State no Brasil: características e perspectivas. *Ciências Sociais Hoje*, ANPOCS, 1989.
- DUNCAN, Bruce & SCHIDT, Maria Inês. *Epidemiologia Clínica: A aplicação do método epidemiológico na pesquisa e na prática médica*. Porto Alegre, mimeo, 1986.
- DUSSAULT, Gilles. Gestão dos Recursos Humanos e Qualidade dos Serviços de Saúde. *Educación Médica y Salud*, 28 (4): 478-489, 1994.
- EATON, Frank. Questions ressortant de la Conférence 93 de la SCE. Les nouveaux visages du monde le l'evalation. *La Société Canadienne d'Evaluation. Bulletin de Liaison*, 13 (2), septembre 1993.
- ECO, Humberto. *Como se faz uma tese*. Ed. Perspectiva S.A., São Paulo, 1983

FARR, Robert M. Representações Sociais: a Teoria e sua História. In: JOVCHELOVITCH, Sandra & GURARESCHI, Pedrinho (orgs.). *Textos em Representações Sociais*. Petrópolis, Rio de Janeiro, Ed. Vozes Ltda., 1994.

FERREIRA, Jair. Vigilância Epidemiológica em Hanseníase. In. *Hanseníase: epidemiologia e controle*. LOMBARDI, Clóvis (coord.). São Paulo, Imprensa Oficial do Estado: Arquivo do Estado, 1990, Capítulo III.

FLETCHER, Robert H.; FLETCHER, Suzanne W. & WAGNER, Edward H. *Epidemiologia Clínica: bases científicas da conduta médica*. Porto Alegre, Artes Médicas, 1989

FOUCAULT, Michel. *A arqueologia do saber*. Petrópolis, Editora Vozes, 1972.

GALVIS, Yolanda Torres de. Nuevas orientaciones de la epidemiologia. *Educación Médica y Salud*, 18 (3): 238-259, 1984.

GARCIA, Juan César. Medicina e Sociedade: as correntes de pensamento no campo da saúde. In: NUNES, E.D. (org.). *Juan César Garcia. Pensamento social em saúde na América Latina*. São Paulo, Cortez Editora/ABRASCO, 1989.

GONÇALVES, Ricardo Bruno Mendes. *Tecnologia e Organização Social das Práticas de Saúde: características tecnológicas de Processo de Trabalho na Rede Estadual de Centros de Saúde de São Paulo*. São Paulo, Ed. HUCITEC/ABRASCO, 1994.

GRANDA, Edmundo. Breves anotaciones sobre la investigación de la práctica y el saber. In: FRANCO, S. & NUNES, E.D. *Debates en Medicina Social*. Quito/Washington, ALAMES-OPS/OMS, 1991.

HARNECKER, Marta. *Os conceitos elementais do materialismo histórico*. 1973.

HARTZ, Zulmira Maria de Araújo. A Epidemiologia e a organização dos serviços. *Proposta - Jornal da Reforma Sanitária*, setembro de 1993.

----- . “Avaliador-Epidemiologista: desafios na formação ‘insdisciplinar’”. In: *III Congresso Brasileiro de Epidemiologia, II Congresso Ibero-Americano, I Congresso Latino Americano e 1ª Mostra de Tecnologia em Epidemiologia - EPITEC*. Salvador, Bahia, Brasil, ABRASCO, 1995.

----- . Análise de Implantação do Programa de Saúde Materno-Infantil no Nordeste do Brasil. Rio de Janeiro, 1996 (No prelo).

HARTZ, Zulmira Maria de Araújo & CAMACHO, Luiz Antônio Bastos. Formação de Recursos Humanos em Epidemiologia e Avaliação dos Programas de Saúde. Rio de Janeiro, *Cadernos de Saúde Pública*, 1996 (No prelo)

- IANNI, Octavio. A Sociologia e o mundo moderno. *Tempo Social: Revista de Sociologia*, USP, 1 (1): 7-27, São Paulo, I sem. 1989.
- JICK, Todd D. Mixing Qualitative and Quantitative Methods: Triangulation in Action. In: MAANEN, John Van (ed.). *Qualitative Methodology*. Beverly Hills/London/New Delhi, Sage Publications, 1985, Fourth Print.
- JOVCHELOVITCH, Sandra & GUARESCHI, Pedrinho (orgs.). *Textos em Representações Sociais*. Petrópolis, Rio de Janeiro, Ed. Vozes Ltda., 1994.
- KISIL, Marcos & CHAVES, Mario (editores). *Programa UNI: Uma Nova Iniciativa na Educação dos Profissionais de Saúde*. Battle Creek, Estados Unidos da América, Fundação W. K. Kellogg, 1994.
- KNOX, E.G. (ed.); ACHESON, R.M.; ANDERSON, D.O.; BICE, T.W. & WHITE, K.L. *Epidemiology in health care planning. A guide to the uses of a scientific method*. Oxford (Inglaterra), International Epidemiological Association/WHO, 1986.
- LAST, John. *A Dictionary of Epidemiology*. New York, Oxford University Press, 1988, 2nd Edition..
- LOVE, Arnold. Les initiés: les nouvelles et les opinions des évaluateurs internes. *La Société Canadienne d'Évaluation. Bulletin de Liaison*, 13 (2), septembre 1993.
- MATUS, Carlos. *Programa do Curso "Governo e Planificação"*. Mimeo, s/d.
- . *Política, Planificación y Gobierno*. Caracas, Venezuela, OPS/OMS, Fundacion Altadir, data provável 1989.
- MCQUEEN, Cameron. Program Evaluation in the Canadian Federal Government. In: Hudson, J.; MAYNE, J. & THOMLISON, R. *Action Oriented Evaluation in Organizations. Canadian Practice*. Toronto, Ontario, Wall&Emerson, Inc., Cap. 2, pp. 28-47, 1992.
- MERHY, Emerson Elias. *A Saúde Pública como política: um estudo de formuladores de políticas*. São Paulo, Ed. HUCITEC, 1992.
- MINAYO, Maria Cecília S. Interdisciplinaridade: uma questão que atravessa o saber, o poder e o mundo vivido. *Medicina*, Ribeirão Preto, 24 (2): 70-77, abr/jun, 1991.
- . *O desafio do conhecimento*. Pesquisa Qualitativa em Saúde. São Paulo-Rio de Janeiro, Editora HUCITEC-ABRASCO, 1992.
- MORAES, Ilara Hammerli Sozzi de. *Informações em Saúde: da prática fragmentada ao exercício da cidadania*. São Paulo-Rio de Janeiro, HUCITEC/ABRASCO, 1994.

- NEMES, Maria Inês Baptistella. Ação Programática em Saúde: recuperação histórica de uma política de Programação. In: SCHRAIBER, Lilia Blima (org). *Programação em saúde hoje*. São Paulo, Ed. HUCITEC, 1990.
- NOGUEIRA, Roberto Passos. *Perspectivas da Qualidade em Saúde*. Rio de Janeiro, Qualitymark Ed., 1994.
- NOVAES, H. Maria D. Epidemiologia e Avaliação em Serviços de Atenção Médica: novas tendências na pesquisa. In: *III Congresso Brasileiro de Epidemiologia, II Congresso Ibero-Americano, I Congresso Latino Americano e 1ª Mostra de Tecnologia em Epidemiologia - EPITEC*. Salvador, Bahia, Brasil, ABRASCO, 1995, p. 49 (Resumo, 186).
- NOVAES, Humberto de Moraes & PAGANINI, José Maria. *Padrões e Indicadores de Qualidade para Hospitais (Brasil)*. OPAS, Edição Especial para UNIMED do Brasil/Congresso Latino-Americano de Hospitais, 1994.
- NUNES, Everardo Duarte. Trayectoria de la Medicina Social en America Latina: elementos para su configuración. In: FRANCO, S. & NUNES, E.D. *Debates em Medicina Social*. Quito, ALAMES, Washington. OPS/OMS, 1991.
- NUNES, Everardo Duarte. A questão da Interdisciplinaridade no estudo da Saúde Coletiva e o papel das Ciências Sociais. In: CANESQUI, Ana Maria (org). *Dilemas e Desafios das Ciências Sociais na Saúde Coletiva*. São Paulo-Rio de Janeiro, Ed. HUCITEC-ABRASCO, 1995.
- OLIVEIRA, Jaime A. de Araújo & TEIXEIRA, Sonia M. Fleury. *(Im)P: revidência Social: 60 anos de história da previdência no Brasil*. Petrópolis, Vozes/ABRASCO, 1989, 2ª e.
- OPS/OMS. *Programación de Salud. Problemas conceptuales e metodologicos* (Publicaciones Científicas N° 111). Washington, D.C., 1965
- . *Usos y Perspectivas de la Epidemiologia. Documentos del Seminario sobre Usos y Perspectivas de la Epidemiologia*. Buenos Aires (Argentina). 7-10/noviembre/1983. (Publicación N° PNSP 84-47). Washington, D.C, 1984.
- . *Evaluación de los Programas de Salud - Programas de Formación de Personal y Programas de Prestación de Servicios. Módulo auto-didáctico*. (Serie Desarrollo de Recursos Humanos N° 59). Washington, D.C., 1985.
- OPS/CEPAL. *Formulación de Políticas de Salud*. Santiago, Chile, 1975.
- PAIM, Jairnilson Silva. As ambigüidades da noção de necessidades de saúde. *Planejamento*, Salvador, Bahia, Fundação de Pesquisa, CPE, 1 (1): 39-46, set./out. 1979.

PITTA, Ana Maria Fernandes. Avaliação como processo de melhoria da qualidade de serviços públicos de saúde. *Revista de Administração Pública*, Rio de Janeiro, 26 (2): 44-61, abr./jun. 1992.

QUEIROZ, Maria Isaura Pereira de. Relatos orais: do “indizível” ao “dizível”. *Ciência e Cultura*, 39 (3): 272-286, março de 1987.

RIVERA, Francisco Javier Uribe (org). *Planejamento e Programação em Saúde: um enfoque estratégico*. São Paulo, Cortez Editora, 1989.

RIVERA, Francisco Javier Uribe. *O agir comunicativo e a planificação estratégica no setor social (e sanitário): um contraponto teórico*. Rio de Janeiro, Tese de Doutorado apresentada à Escola Nacional de Saúde Pública, mimeo, 1991.

SCHRAIBER, Lilia Blima. Pesquisa qualitativa em saúde: reflexões metodológicas do relato oral e produção de narrativas em estudo sobre a profissão médica. *Revista de Saúde Pública*, São Paulo, 29 (1): 63-74, 1995.

SILVA, Maria Lígia V. da & FORMIGLI, Vera Lúcia A. Avaliação em Saúde: Limites e Perspectivas. *Cadernos de Saúde Pública*, Rio de Janeiro, 10 (1): 80-91, jan/mar, 1994.

SILVER, Lynn. Aspecto Metodológicos em Avaliação dos Serviços de Saúde. In: *Planejamento Criativo. Novos Desafios em Políticas de Saúde*. (E. Gallo, F.J.U. Rivera & M.H. Machado, orgs), Rio de Janeiro, Editora Relume-Dumará, 1992.

SOUSA CAMPOS, Gastão Wagner de. *A Saúde Pública e a defesa da vida*. São Paulo, Ed. HUCITEC, 1991.

----- . *Reforma da Reforma. Repensando a Saúde*. São Paulo, Ed. HUCITEC, 1992.

SUCHMAN, Edward A. *Evaluative Research. Principles and practice in Public Service & Social Action Programs*. New York, Russel Sage Foudation, 1967

----- . In CARO, Francis G. *Readings in Evaluation Research*. Russel Sasse Foundation, New York, 1977, 2ª ed., p. 48

TESTA, Mario. *Pensar em Saúde*. Porto Alegre, Editora Artes Médicas/ABRASCO, 1992.

THIOLLENT, Michel. *Crítica Metodológica, Investigação Social e Enquete Operária* (Coleção Teoria e História, Vol. 6). Editora Polis, São Paulo, 1987, 5ª ed..

VERGARA, Mario Espinoza. *Evaluacion de Proyetos Sociales*. Buenos Aires, Editorial Hvmanitas, 1980.

WEILENMANN, Alexander. *Evaluation research and social change*. Ghent (Bélgica), UNESCO, 1980.

WEISS, Carol. *Evaluation Research. Methods of Assessing Program Effectiveness*. New Jersey, Prentice-Hall, Inc., 1972.

YACH, Derek. The use and value of Qualitative Methods in health research in developing countries. *Social Science and Medicine*, 35 (4): 603-612, Grã-Bretanha, 1992.

YIN, Robert K. *Applicatons of Case Study Research*. Newbury Park, CA, Sage Publications, 1993.

YIN, Robert K. Discovering the Future of the Case Study Method in Evaluation Research. *Evaluation Practice*, 15 (3): 283-190, 1994.

ANEXO I

RELAÇÃO ENTRE A TEORIA DE SISTEMAS, O PLANEJAMENTO E A AVALIAÇÃO EM SAÚDE (Nível Individual ou Nível Agregado/Coletivo)

A Teoria de Sistemas

Insumos -----> Processo -----> Produto
(Input) (Process) (Output)

Os componentes da intervenção

Estrutura -----> Processo-----> Resultado

A lógica da intervenção

Problema ou Necessidade -----> Ação -----> Efeito

A avaliação da intervenção

Avaliação Inicial -----> Avaliação Durante-----> Avaliação Final
(Diagnóstico) (Controle) (Efeitos)

↑
Retro-alimentação
(Feedback)

ANEXO II

Fonte: Contandriopoulos et al. *Saber Preparar uma Pesquisa. Definição, Estrutura, Financiamento.* São Paulo-Rio de Janeiro, Ed. HUCITEC/ABRASCO, 1994, p. 20.



ANEXO IV

**A INTERDISCIPLINARIDADE DA AVALIAÇÃO
- NOVAS ÁREAS DE PESQUISA -**

**Pesquisa do Cuidado à Saúde
(*Health Care Research*)**

Os estados de saúde

**As intervenções
em saúde**

**Fonte: BATTISTA et. al. An integrative framework for Health-Related
Research. *Journal of Clinic Epidemiology*, 42 (12): 1155-1160,
1989.**

ANEXO V

FÓRMULA BASE DO MODELO EXPERIMENTAL

	ANTES	DEPOIS
EXPERIMENTAL	a	b
CONTROLE	c	d

Se a diferente entre a e b é maior que a diferença entre c e d , o programa é um sucesso.

Fonte: WEISS, Carol. *Evaluation Research. Methods for Assessing Program Effectiveness*. Englewood Cliffs, New Jersey, Prentice-Hall, 1972, p. 61.

MODELO EXPERIMENTAL NA ANÁLISE DAS INTERVENÇÕES

GRUPO	TEMPO	
	ANTES	DEPOIS
COM PROJETO	X	X'
SEM PROJETO	Y	Y'

Fonte: COHEN & FRANCO. *Avaliação de Projetos Sociais*. Petrópolis, RJ, Vozes, 1994,

p. 127.

ANEXO V - Arq. A5.Doc

FÓRMULA BASE DO MODELO EXPERIMENTAL> VER TAMBIÉN
EM:

VERGARA, Mario Espinoza. *Evaluación de Proyectos Sociales*. Buenos Aires, Editorial Hymnitas, p. 96.

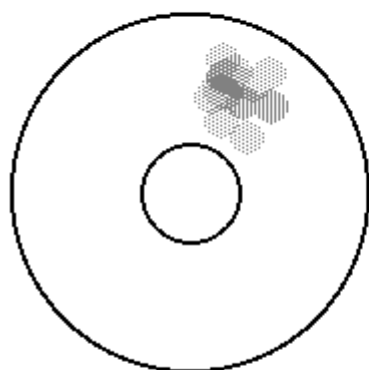
ANEXO VI

QUADRO COMPARATIVO DOS MODELOS DE ESTUDO EM RELAÇÃO AOS DESENHO EXPERIMENTAL

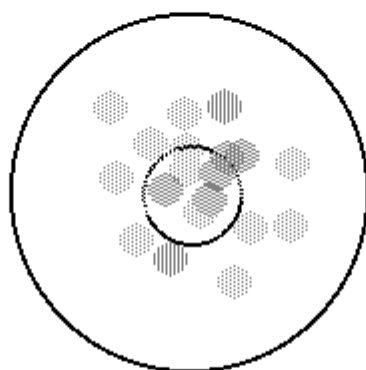
SITUAÇÃO	TIPO DE DESENHO		
	EXPERIMENTAL	QUASE-EXPERIMENTAL	E
Número suficiente de “sujeitos” produzindo múltiplos pontos de dados	<input type="checkbox"/>	<input type="checkbox"/>	
Número limitado de variáveis de interesse (menor que o número de pontos de dados)	<input type="checkbox"/>	<input type="checkbox"/>	
Investigador pode controlar e manipular as variáveis de interesse	<input type="checkbox"/>		
Investigador pode distribuir os sujeitos aleatoriamente no grupo experimental e no grupo controle	<input type="checkbox"/>		

Adaptado de YIN, Robert K. *Application of Case Study Research*. Newbury Park Publications, 1993.

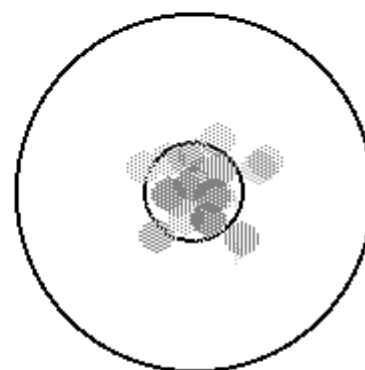
O Esquema do Tiro ao Alvo



Confabilidade



Validade



Ambas

BEST, John W. Research In Education.

Prentice-Hall, INC., Englewood Cliffs, New Jersey, 1970.

ANEXO VIII

A QUESTÃO DA VALIDADE NAS METODOLOGIAS QUANTITATIVAS

VALIDADE EXTERNA : capacidade de generalizar resultados para outras populações (amostragem)

VALIDADE INTERNA : capacidade de generalizar resultados para a população alvo. Evitar erros
- aleatórios (acaso) > estimativa estatística
- sistemático (viés) > controle desenho do estudo

1)VIÉS DE SELEÇÃO: grupo experimental/controle

2)VIÉS DE CONFUSÃO: variáveis de confusão não consideradas no modelo teórico do estudo

**3)VIÉS DE AFERIÇÃO: qualidade das medidas
construção de indicadores**

**A) CONFIABILIDADE DAS MEDIDAS (*Reliability*)
precisão e fidelidade do instrumento**

B) VALIDADE DAS MEDIDAS (*Validity*)

- **VALIDADE DE CONSTRUÇÃO (Plano
Conceitual)**

- **VALIDADE DE CONTEÚDO (Elenco
de variáveis)**

- **VALIDADE DE CRITÉRIO (Ponto de corte
das Normas ou Padrões)**

. **Sensibilidade**
. **Especificidade**

ANEXO IX

TIPOLOGIA DA AVALIAÇÃO*

CRITÉRIOS	TIPOS
MOMENTO DA AVALIAÇÃO	<ol style="list-style-type: none">1. Avaliação antes ou <i>ex ante</i>2. Avaliação durante, contínua, concorrente, concomitante, de gestão, monitoramento3. Avaliação depois ou <i>ex post</i>
FUNÇÃO DA AVALIAÇÃO	<ol style="list-style-type: none">1. Avaliação Formativa2. Avaliação Somativa
PROCEDÊNCIA DOS AVALIADORES	<ol style="list-style-type: none">1. Avaliação Externa2. Avaliação Interna3. Avaliação Mista4. Auto-Avaliação
ASPECTOS OU COMPONENTES AVALIADOS (do programa)	<ol style="list-style-type: none">1. Avaliação Formal (áreas de coerência interna)2. Avaliação Substantiva (área de coerência externa)

*Cf. AGUILAR, Maria José & ANDER-EGG, Ezequiel. *Avaliação de Serviços e*

Programas Sociais. Petrópolis, Vozes, 1994.

ANEXO X

TIPOS DE FENÔMENOS

NÚMERO DE VARIÁVEIS	RELAÇÕES ENTRE VARIÁVEIS	
	DEFINIDA	INDEFINIDA
SIMPLES	1.Simples Definido	2.Simples Indefinido
COMPLEXO	3.Complexo Definido	3.Complexo Indefinido

TIPOS DE PROBLEMAS:

PROBLEMA BEM-ESTRUTURADO: 1

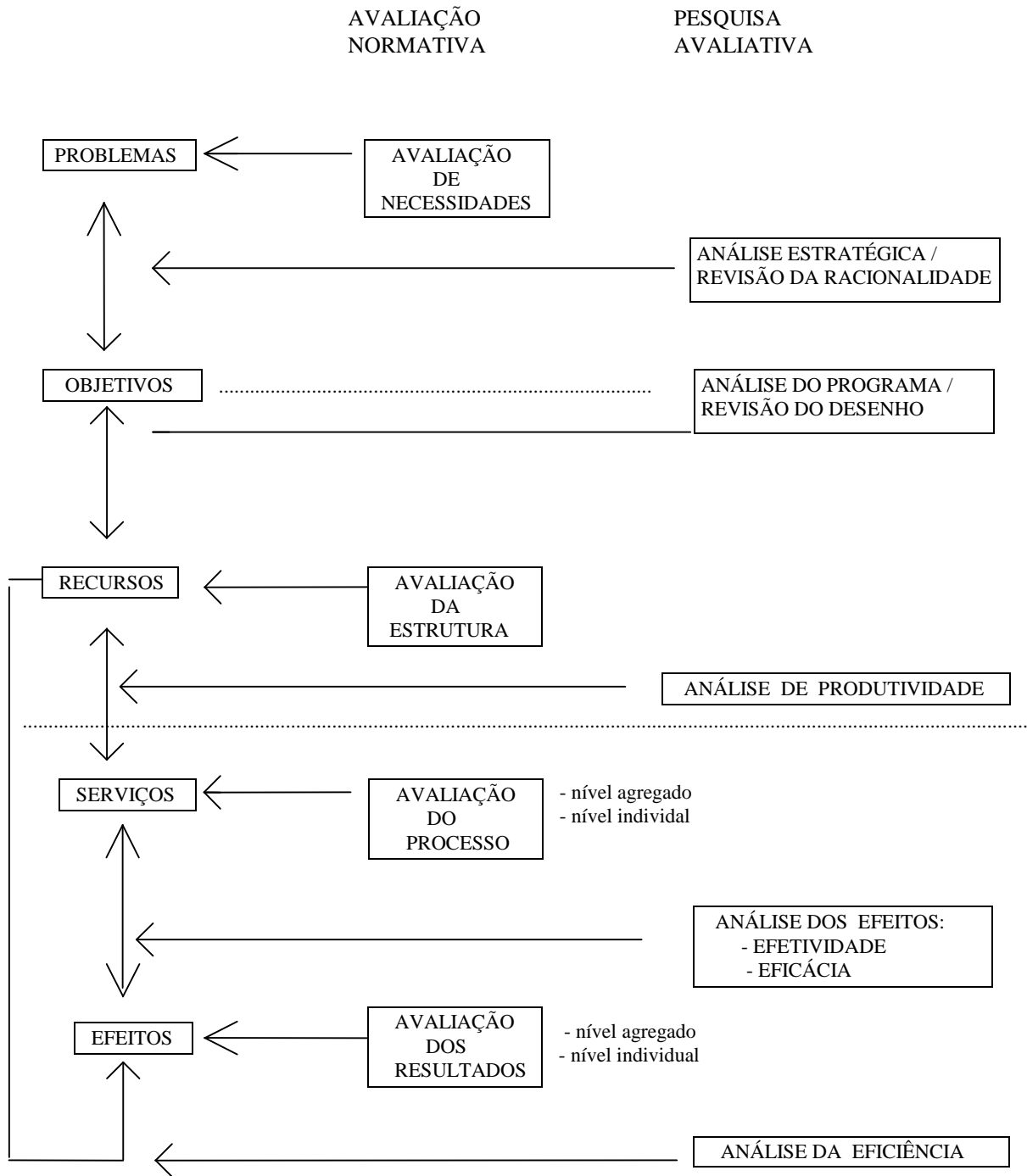
PROBLEMA QUASE-ESTRUTURADO: 2 e 3

PROBLEMA MAL-ESTRUTURADO: 4

Cf. TESTA(1992) e MATUS (1989)

ANEXO XI

ESTRUTURA CONCEITUAL PARA A AVALIAÇÃO DE PROGRAMAS DE CUIDADOS DE SAÚDE*



Adaptado de CLEMENHAGEN&CHAMPAGNE. Quality Assurance as part of Program Evaluation.: guidelines for managers and clinical department heads. *Quality Review Bulletin*, Journal of Quality Assurance, 12 (11): 383, nov/86.

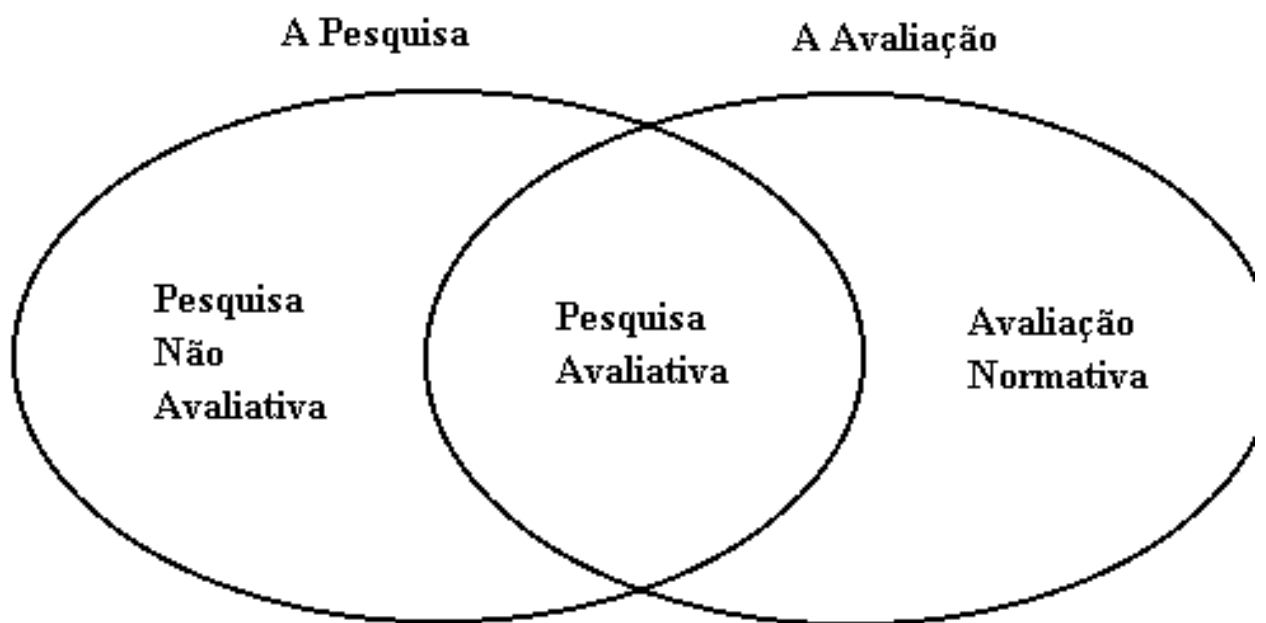
ANEXO XII

O EIXO INDIVIDUAL/COLETIVO NA DISTINÇÃO ENTRE A AVALIAÇÃO DE QUALIDADE E A AVALIAÇÃO DE PROGRAMAS NO CAMPO DA SAÚDE

	AVALIAÇÃO DE QUALIDADE	AVALIAÇÃO DE PROGRAMAS
DONABEDIAN(1982) (Cuidado em Saúde)	Eficiência Clínica	Eficiência de Produção
CLEMENHAGEN & CHAMPAGNE (1986) (Âmbito Hospitalar)	Aspectos Clínicos Nível Individual Garantia Qualidade	Aspectos não Clínicos Nível Agregado Avaliação de Programa
BATTISTA et al. (1989)	Pesquisa das Práticas (Epidemiologia Clínica)	Pesquisa de Programas e Políticas (Health Services Research)
SILVER (1992)	Aspectos Clínicos (assistência curativa)	Aspec.Organizacionais (assist. preventiva)
COMITÊ DE APLICAÇÃO DA TQT* NA SAÚDE (1992)	Serviços Clínicos	Serviços Operacionais

*TQT: Teoria de Qualidade Total

Pesquisa e Avaliação



ANEXO XIV

UMA TIPOLOGIA INTEGRADA DOS TIPOS DE AVALIAÇÃO

AUTOR	CLASSIFICAÇÃO	
Educação	AVALIAÇÃO FORMATIVA	AVALIAÇÃO SOMATIVA
Donabedian	ESTRUTURA E PROCESSOS	RESULTADOS
Ander-Egg Conhen&Franco	CONTROLE OPERACIONAL/ AVALIAÇÃO CONTÍNUA	AVALIAÇÃO FINAL
Aguilar&Ander-Egg Cohen&Franco	AVALIAÇÃO FORMAL -coerência interna - (racionalidade funcional)	AVALIAÇÃO SUBSTANTIVA -coerência externa - (racionalidade substantiva)
Contandriopoulos e colaboradores	AVALIAÇÃO NORMATIVA (componentes)	PESQUISA AVALIATIVA (relações entre componentes)

ANEXO XV

QUADRO COMPARATIVO DA RELAÇÃO ENTRE A FINALIDADE DA AVALIAÇÃO, O ASPECTO AVALIADO NA INTERVENÇÃO E OS COMPONENTES DO PROGRAMA

FINALIDADE DA AVALIAÇÃO	ASPECTO AVALIADO	COMPONENTE DO PROGRAMA
Estratégica	Necessidade	Pré-requisitos (Problemas e Objetivos)
Formativa	Adequação	Implementação (Recursos e Serviços)
Somativa	Efetividade	Efeitos (Efeitos)

ANEXO XVI

- A AVALIAÇÃO DAS INTERVENÇÕES DE SAÚDE

QUADRO COMPARATIVO DA RELAÇÃO ENTRE O OBJETIVO DA AVALIAÇÃO AVALIADO NA INTERVENÇÃO, O ENFOQUE DA AVALIAÇÃO E OS TIPOS DE

OBJETIVO DA AVALIAÇÃO	ASPECTO AVALIADO	FASE OU MOMENTO DO PROGRAMA	ENFOQUE DA AVALIAÇÃO	
			COMPONENTES DO PROGRAMA	RELAÇÕES ENTRE OS COMPONENTES
Estratégico	Necessidade ou Pertinência	Pré-Requisitos	<i>Problemas</i>	
				Problemas e Objetivos
Formativo	Adequação ou Suficiência	Implementação	<i>Objetivos</i>	Objetivos e Recursos
			<i>Recursos</i>	Recursos e Serviços
Somativo	Efetividade	Efeitos	<i>Serviços</i>	Serviços e Efeitos
			<i>Efeitos</i>	

Análise de Eficiência: relaciona os Efeitos e os Recursos
 Análise de Implantação: relaciona a dinâmica interna do programa (relação Serviços, Recursos e Efeitos) com o Contexto d

ANEXO XVII

QUADRO COMPARATIVO ENTRE O TIPO DO DE ANÁLISE, AS QUESTÕES DE ESTUDO NA PESQUISA AVALIATIVA

TIPO DE ANÁLISE	QUESTÕES DA PESQUISA AVALIATIVA	TIPO
Análise Estratégica	Justifica-se a intervenção em relação ao problemas da população? É pertinente que seja este interventor que a assumam?	- Análise de ne - Métodos de d - Análises de n - Análise de ca
Análise da Intervenção	A intervenção se propõe a alcançar os objetivos de maneira satisfatória?	Extrapolações apreciar a qual instrumento (confiabilidade métodos que pe qualidade de u (veracidade e
Análise da Produtividade	Os recursos são alocados de maneira que otimizem os serviços produzidos?	Modelos deriva econômicos e n contabilidade a
Análise dos Efeitos (Eficácia e Efetividade)	Quais os efeitos que podem ser atribuídos à intervenção?	Muitos método qualitativos; p pesquisa sintét
Análise do Rendimento (Eficiência)	Com que custo são obtidos os efeitos da intervenção?	- Custo/Benefício - Custo/Eficácia - Custo/Utilidade
Análise da Implantação	Como e por quê os efeitos variam entre os meios nos quais a intervenção é introduzida?	Estudos de Cas

Fonte dos dados: CONTANDRIOPOULOS et al. 1992, 1994.

ANEXO XVIII

AS SETE CARACTERÍSTICAS DE UM BOM INDICADOR

Jordan Filho, 1974 Cf. Ilara H. Moraes, p.21)	Jair Ferreira, p. 44-45)
<i>Simplicidade:</i> deve ser fácil de ser calculado	<i>Simplicidade:</i> facilidade de cálculo a partir das informações básicas
<i>Validade:</i> deve ser função da característica que se deseja medir	<i>Especificidade:</i> deve ser função das características do fenômeno que se quer ou se necessita medir
<i>Disponibilidade:</i> deve usar dados habitualmente disponíveis ou de fácil obtenção	<i>Disponibilidade:</i> os dados básicos para a construção do indicador devem ser de fácil obtenção para diferentes áreas e épocas
<i>Robustez:</i> pouco sensível às deficiências dos dados necessários a sua construção	<i>Confiabilidade:</i> pouco influenciado pelas variações intrínsecas dos dados
<i>Sinteticidade:</i> deve refletir o efeito do maior número possível de fatores	<i>Abrangência:</i> deve sintetizar o maior número possível das condições ou fatores diferentes que afetam a situação que se quer descrever
<i>Discriminatoriedade:</i> deve possuir um alto poder discriminatório para vários níveis de saúde e indicar alterações que ocorram com o tempo	<i>Sensibilidade:</i> deve ter o poder de distinguir e apontar os vários níveis de saúde, reconhecer sensivelmente e expor à vista as variações ocasionais da tendência do problema de saúde pública em uma determinada área.
<i>Cobertura:</i> deve referir-se, tanto quanto possível, se necessário for, a cada país ou território como um todo e não somente a determinada área selecionada ou a um grupo da população	<i>Discriminatoriedade:</i> deve ter o poder de refletir os diferentes níveis de saúde mesmo entre áreas com particularidades distintas

ANEXO XIX

O PENSAMENTO ESTRATÉGICO DE MÁRIO TESTA

RACIONALIDADE DO PLANEJAMENTO	TIPOS DE PODER	POSTULADO DE COERÊNCIA
Política	Poder Político	Propósitos de Governo (propostas)
Administrativa	Poder Administrativo	Organização (instituições)
Técnica	Poder Técnico	Método
