

# AVALIAÇÃO: CONSIDERAÇÕES TEÓRICAS E POSICIONAMENTOS

HERALDO MARELIM VIANNA<sup>1</sup>

*Escuta, irmão,  
O homem é verdade maior.  
Nada há além dele.*  
Cânticos de Amor de Chandidas<sup>2</sup>

## Introdução

A avaliação educacional no atual contexto brasileiro vive momento eufórico e afeta diferentes segmentos da comunidade educacional em diversos níveis administrativos. A avaliação não se limita apenas à verificação do rendimento escolar, atividade rotineira (e burocrática) no âmbito institucional da escola. A avaliação atual concentra-se em um nível maior, segundo uma perspectiva integrada a programas de qualidade. Subitamente, a avaliação do ensino básico passou a ser uma exigência, ainda que seu impacto seja mínimo, na maioria das vezes. Os estados, dentro de suas limitações, procuram avaliar seus sistemas de ensino, abrangendo muitas vezes a totalidade do seu alunado. As avaliações de programas (cursos) e até mesmo de instituições passaram a dominar os centros educacionais federativos

---

<sup>1</sup> Pesquisador do Departamento de Pesquisas Educacionais, da Fundação Carlos Chagas.

<sup>2</sup> In: MEHTA, G. *O monge endinheirado, a mulher do bandido e outras histórias de um rio indiano*. Trad. de Hildegard Feist. 1ª impressão. Companhia das Letras. São Paulo, 1996.

via programação especial à distância. O MEC, associado à Universidade Nacional de Brasília (UnB), procura desenvolver **expertise** em nível de 3º grau, afetado, no momento, pela propositura de diferentes formas de acesso à universidade que fogem à ortodoxia e à tradição. Assim, antes de apresentarmos a nossa percepção sobre alguns teóricos da avaliação que nos influenciaram, parece-nos importante considerar diferentes questões ligadas à avaliação e apresentar o nosso posicionamento, antecipando, desse modo, alguns pontos que serão mais tarde explorados com maior amplitude, ao enfocarmos nos próximos estudos as figuras de **Tyler, Cronbach, Scriven, Stufflebeam** e, especialmente, a grande figura humana de **Robert E. Stake**.

### 1. Avaliação: seu interesse

Antes de prosseguir, tentaremos responder a uma pergunta que é, sem dúvida, de relevância para toda a sociedade: por que o atual interesse pela avaliação? sempre não houve, desde os primeiros tempos, alguma forma de avaliação? Sem dúvida, desde o início do processo civilizatório houve alguma forma de avaliação. Ousaríamos dizer que a avaliação surgiu com o próprio homem, se entendermos por avaliação a visão apresentada por Stake — **o homem observa; o homem julga**, isto é, avalia. A avaliação, entretanto, no seu devenir sofreu transformações e gerou novas construções. É um constante **vir-a-ser** na área das ciências do homem, tornando-se uma atividade complexa, fundamentada no pensamento descritivo, analítico e crítico. O seu enfoque não está circunscrito ao aluno e seu rendimento, ao desenvolvimento de atitudes e de interesses, que constituem o produto do processo instrucional que ocorre na escola. O seu interesse se ampliou, mas não ficou no âmbito da microavaliação (Payne, 1974), passou a se interessar por grupos de indivíduos (alunos, professores, administradores, técnicos etc); projetos, produtos e materiais; instituições e sistemas educacionais nos seus diversos níveis e competências administrativas, evoluindo, assim, para uma área bem mais ampla, que constitui o campo de macroavaliação.

É evidente que para essa modificação bastante ampla houve a influência de um concurso de fatores, em que se destacam,

inicialmente, a tomada de consciência dos educadores face à complexidade do seu campo de atuação e à necessidade de definir e avaliar a prioridade de alguns problemas no mundo moderno, como, por exemplo, a educação de crianças de 0 a 6 anos, a extensão da educação a populações marginais e/ou carentes, que exigem o emprego de novas metodologias e às vezes de formas heterodoxas de avaliação, a questão crítica das altas taxas de reprovação — muito possivelmente em decorrência de avaliações precárias, hipotetizamos — e a necessidade, entre outros problemas usualmente relevantes, de selecionar e treinar professores capazes para sistemas educacionais de massa. Tudo isso envolveu, naturalmente, posicionamentos e ações a partir da avaliação desses problemas, que passaram a ser encarados sob diferentes perspectivas.

## 2. Avaliação, transformação e educação

A transformação tecno-científica, iniciada nos tempos modernos e ativada a partir da primeira década do século XX, sofreu um processo de aceleração que provocou o rápido envelhecimento dos currículos e programas, que se transformaram em obsolescências incompatíveis com a vida moderna, o que se refletiu, naturalmente, nas práticas de ensino e nos materiais pedagógicos, que se tornaram inúteis para um ensino de qualidade e uma orientação pedagógica eficiente. A avaliação vai ter um papel importante na crítica para a transformação da escola, de seus currículos e de seus programas, o que ocorreu mais visivelmente nos países do Primeiro Mundo (vide, por exemplo, os relatórios da **Fundação Nuffield**, na Inglaterra), e que vem ocorrendo em ritmo retardado nos chamados países emergentes (Brasil, por exemplo).

Ao que foi anteriormente apresentado, precisamos acrescentar a questão da qualidade em educação<sup>3</sup>, preocupação generalizada de todos os segmentos sociais e não apenas uma preocupação corporativista de professores, por razões de ofício. A educação, como um valor social, extrapola os interesses circunstanciais de grupos

---

<sup>3</sup> Vide *Estudos em Avaliação Educacional*, nº 5, 1992, dedicado inteiramente ao problema. Fundação Carlos Chagas, São Paulo.

isolados e passa a ser uma preocupação da sociedade global. Por que a decadência do ensino? Por que a perda do **status** do professor? Por que, aos poucos, a educação deixou de ser prioridade nacional, apesar do discurso político? Há toda uma rede de fatores confluentes e que se interpenetram gerando uma rede de causas, fatos e efeitos que somente a avaliação associada à pesquisa (ou vice-versa) conseguirá deslindar essa realidade e talvez de forma não plenamente satisfatória.

As grandes agências financiadoras nacionais e internacionais são em parte responsáveis pelo desenvolvimento que a avaliação vem tendo, assim como o desenvolvimento de cursos de pós-graduação nos seus vários níveis, com a exigência de trabalhos de pesquisa. O assunto será retomado em Tyler e na discussão sobre **accountability**, quando discutiremos a avaliação nos Estados Unidos, especialmente a partir dos anos 60. Os altos investimentos financeiros exigem, necessariamente, um projeto de avaliação que demonstre os resultados em termos de custos-benefícios, apesar de muitas vezes essa demanda refletir uma concepção econometrista de educação, sem considerar a componente humanista. A exigência tem contribuído para que a avaliação se torne uma constante no processo educacional, envolvendo não apenas a avaliação do aluno (rendimento escolar), mas considerando outros elementos (currículos, programas, professores, rede física etc.), igualmente válidos para expressar a qualidade de um sistema educacional<sup>4</sup>.

Uma ressalva, entretanto, precisa ser feita — impõe-se uma análise crítica da avaliação, cujos produtos nem sempre são satisfatórios, inclusive nos países que já fizeram a sua revolução educacional e dispõem entre seus educadores de uma massa crítica extremamente capaz. As avaliações, admitindo-se que sejam feitas com suficiente **expertise**, ainda assim apresentam problemas, que decorrem, em parte, como assinalou Stufflebeam et al. (1971), da falta de uma tipologia que permita estabelecer o máximo possível de informações necessárias ao processo decisório; da carência de instrumentos válidos (e fidedignos) que possam ser utilizados para

---

<sup>4</sup> VIANNA, H. M. Avaliação da Qualidade em Educação - apresentação de um modelo. *Estudos em Avaliação Educacional*, nº 5, 1992, São Paulo: Fundação Carlos Chagas. (ver Anexo 2).

captar a complexidade dos fenômenos que são objeto da avaliação; da inexistência de um sistema informatizado que possibilite organizar, processar e relatar as diversas variáveis envolvidas na complexidade das múltiplas dimensões intervenientes nos diferentes processos educacionais e, finalmente, a falta de mão-de-obra qualificada para atender à complexidade daquilo que constitui o objeto de nosso maior interesse — avaliar e tomar decisões fundamentadas.

### 3. Avaliação, mensuração e valor

A avaliação educacional com a abrangência que possui nos dias fluentes somente se iniciou na década de 40, graças à intensa atuação de Tyler, como detalharemos na discussão de suas idéias, e, durante os anos 60, à produção teórica de Cronbach, Scriven e Stake que abordou aspectos fundamentais dos procedimentos de avaliação, influenciando a geração que começou a trabalhar nessa área. Ao apresentarmos diferentes definições de avaliação, antecipamos a idéia de que, no processo de formação do avaliador, a definição que formulamos, ou aceitamos, condiciona os procedimentos do avaliador.

Antes de enfrentarmos a questão, faz-se necessário esclarecer que, ainda hoje em dia, apesar do desenvolvimento da avaliação, prevalece a injustificada confusão entre avaliação e mensuração. Uma rápida reflexão histórica mostra que a avaliação, inicialmente, confunde-se com medida, surge, assim, como uma disciplina psicométrica, ou mais exatamente docimológica. Por outro lado, os estudos sobre diferenças individuais iniciados no princípio do século XIX contribuem para que a identidade entre avaliação e medida se acentue, mesmo no espírito dos mais esclarecidos; por sua vez, nessa mesma época, a avaliação é associada à mensuração do rendimento escolar, confusão que ainda persiste nos dias fluentes, em que **medir**, quantificação de um atributo, segundo determinadas regras, é visto como **avaliar**. A medida pode ser um momento inicial de uma avaliação, mas não é condição essencial para que se tenha uma avaliação, que se concretiza quando ocorre um julgamento de valor, na visão de Scriven (1967), ao analisar os desdobramentos metodológicos da avaliação.



Seguindo a lógica de Scriven, a avaliação está ligada à determinação do **valor** de alguma coisa para uma certa destinação; desta forma, para avaliar cumpre determinar um universo de informações que permita ajuizar uma manifestação de valor em relação ao que pode ser um programa, um produto, um material, e não exclusivamente o rendimento escolar (Worthen e Sanders, 1973). A fim de que a avaliação se concretize há necessidade de um conjunto de atividades formais, sistemáticas, no dizer de Popham (1975), que nos levam a ter condições de apresentar um juízo de valor sobre determinado aspecto de proceder educacional. A avaliação para outros não precisa ser um processo sistemático, pode ser assistemático, baseando-se na **expertise** de alguém, ou seja, na vivência de um profissional qualificado e experiente no campo educacional. As avaliações informais são bastante freqüentes em educação e perfeitamente válidas. Nem sempre há necessidade de criar todo um complexo tecnológico, apoiado em sofisticados procedimentos estatísticos, para a avaliação de um livro, por exemplo, ou mesmo de um programa. Avaliações impressionistas, desde que bem feitas com base no julgamento de pessoas qualificadas, podem ser extremamente úteis no dia-a-dia da educação, ainda que possam se revestir de um caráter por muitos considerado não-científico, porque eivado de subjetividade. A subjetividade, veremos, ao analisarmos as colocações de Stake, tem o seu lugar e o seu significado na avaliação educacional, apesar da reação daqueles que não se tornaram autônomos em relação ao pensamento positivista, que teve o seu momento em uma fase da evolução do pensamento, inclusive do educacional. Aos poucos, a área da avaliação se vai dando conta de que percepções e intuições podem conduzir à geração de conhecimentos; entretanto, simultaneamente, a avaliação procura sistematizar suas atividades, na medida do possível, para, aos poucos, criar uma estrutura que permita a coleta de diferentes tipos de informação que possibilitem verificar a congruência das intencionalidades em educação (Tyler, Stake), apresentar juízos de valor (Scriven) e tomar decisões face às alternativas que nos são oferecidas (Stufflebeam). [1].

#### **4. Avaliação — dimensão histórica, definições**

Se colocarmos o problema da avaliação educacional em sua dimensão histórica, veremos que Tyler (1942), há mais de meio século, apresentou a avaliação como um processo de estabelecimento da comparação entre os desempenhos e a concretização de objetivos instrucionais pré-definidos. A definição de Tyler, atacada por uns, aceita incondicionalmente por outros, teve amplas repercussões, inclusive influenciando o pensamento de Stake (1967) na construção de um modelo neo-tyleriano — **countenance** em avaliação — em que os **intents** (intenções, objetivos) desempenham importante papel, como será analisado posteriormente. Outros também se preocupam em definir avaliação, como foi o caso de Scriven (1967), que com sua intensa atuação marcou o campo da avaliação educacional. Scriven destaca a avaliação como um processo de levantamento de dados para análise e posterior determinação do valor de um certo fenômeno, posicionamento que vai influenciar grande número de teóricos da avaliação, inclusive Stufflebeam et al. (1971). Estes desenvolveram todo um plano teórico baseado na idéia de que a avaliação visa a tomada de decisões, ou seja, é um processo de identificar e coletar informações que permitam decidir entre várias alternativas. Ora, ainda que Stufflebeam e seus colaboradores não tenham explicitado e incluído um julgamento de valor na sua conceituação de avaliação, ele está implícito na mesma, porque escolher entre diferentes alternativas para tomar uma decisão exige, como condição necessária e indispensável, a realização de julgamentos de valor [2a] [2b].

#### **5. Avaliação e pesquisa**

Antes de considerarmos outros aspectos ligados à avaliação educacional e ao avaliador, precisamos nos posicionar sobre uma questão que não é unicamente teórica, mas que tem suas implicações práticas: as relações entre a avaliação e a pesquisa. Este problema pode ser considerado sob dois aspectos: 1) a avaliação é uma aplicação dos métodos da pesquisa na solução de problemas específicos de uma área; e 2) a avaliação e a pesquisa são campos diferentes, possuindo cada um deles características específicas. Uma questão se nos

apresenta de imediato: — haveria necessidade de discutir se esses dois campos da atividade científica em educação apresentam aspectos convergentes ou se, ao contrário, são áreas singulares, cada uma com as suas próprias especificidades? Não se trata de uma questão especulativa, tendo em vista que, em função do posicionamento adotado, o comportamento do avaliador e/ou do pesquisador, no exercício das suas atividades empíricas, vai variar na definição, no planejamento e no desenvolvimento do seu programa de trabalho, assim como na elaboração do relatório final da avaliação e/ou da pesquisa (Norris, 1990).

À primeira vista observa-se que há uma transvariância entre pesquisa e avaliação<sup>5</sup>: entretanto, na prática efetiva da avaliação, é possível perceber, como o fazem Glass e Worthen (1971) e Worthen e Sanders (1973), que existem características de uma e de outra que as distinguem, bastando, no caso, que se considere, entre outros aspectos, as diferenças de motivação do pesquisador (**curiosidade científica**) e do avaliador (**interesse na solução de um problema**), os objetivos da pesquisa (**estabelecimento de conclusões**) e os da avaliação (**possibilitar a tomada de decisões**), o caráter da pesquisa (**nomotético — estabelecimento de leis/princípios**) e o da avaliação (**ideográfico — descrição de um fenômeno**), o grau de autonomia do pesquisador (**maior**) e o do avaliador (**menor**), a utilidade social (**direta, na avaliação; indireta, na pesquisa**); o nível de universalidade dos fenômenos estudados (**amplo, na pesquisa; restrito, na avaliação**), a questão do valor, que é **determinante** na avaliação, orientando, inclusive, a procura das informações necessárias; as técnicas investigativas apresentam grandes **similaridades** em relação à determinação e ao julgamento da validade; a pesquisa é julgada em função de sua **validade interna e externa**, enquanto na avaliação o critério de julgamento se fundamenta na sua **utilidade** e na sua **credibilidade**, como destacaram Glass e Worthen (1971) [3].

---

<sup>5</sup> Veremos mais adiante a posição de Stufflebeam, mostrando que existe um traço comum entre pesquisa e avaliação, que é a obtenção de conhecimento para a compreensão dos fenômenos educacionais e a solução de seus problemas.



---

Apesar da argumentação apresentada por muitos teóricos, parece-nos que a distinção entre avaliação e pesquisa não seja tarefa simples de ser concretizada. Há entre pesquisa e avaliação uma comunalidade de aspectos, bastando lembrar que a avaliação está relacionada ao desenvolvimento geral da pesquisa científica e que os interesses dos avaliadores e dos pesquisadores na área educacional nem sempre se revestem de significado, e toda discussão sobre similaridades e diferenças entre avaliação e pesquisa deve ser feita com extremo cuidado, a fim de não partirmos de abstrações que acabem levando a comparações entre fases do processo de pesquisa e da avaliação apresentadas nos livros, textos e manuais didáticos. As conclusões não partem de uma análise comparativa do desenvolvimento da pesquisa e da avaliação, mas de posicionamentos idealistas e prescritivos de como deve ser o processo de cada uma delas, não podendo, assim, oferecer uma descrição efetiva das atividades de uma e de outra. É necessário destacar que, ao compararmos pesquisa e avaliação, não podemos ignorar que suas metodologias foram influenciadas por metodologias de outras áreas do conhecimento, em particular da psicologia, da sociologia e, sobretudo, da antropologia, o que faz com que a avaliação e a pesquisa tenham inúmeros pontos de confluência e, assim, apresentem pontos de identidade e abordem aspectos que podem diferir quanto à ênfase que lhes é dada, mas não em relação à relevância na criação do conhecimento e na solução de problemas [4].

### **6. Abordagens conceituais e orientações teóricas**

Ao tratarmos de algumas questões teóricas, selecionadas entre as que julgamos de maior relevância, no caso do presente trabalho, não podemos deixar de analisar diferentes abordagens conceituais e orientações teóricas. A existência de diversos tipos de avaliação, em decorrência de pressupostos metodológicos ou epistemológicos, levamos à discussão dos paradigmas e modelos em avaliação, que será precedida de considerações sobre Popper e o positivismo. Algumas idéias de Kuhn sobre paradigmas antecedem nossas colocações referentes a modelos e a nosso posicionamento face à avaliação responsiva e à avaliação iluminativa.

Os anos 90 assistiram ao desaparecimento de duas grandes figuras do pensamento científico moderno: Karl Popper (1902-94) e Otto Kuhn (1922-96), que, com suas idéias provocativas e polêmicas contribuíram para gerar novas perspectivas no campo da filosofia da ciência e de forma indireta acabaram influenciando as várias metodologias de pesquisa e de avaliação educacional. Não nos ocuparemos dos grandes problemas epistemológicos, o que foi feito com bastante brilho pelos autores do livro editado por Egon G. Guba (1990)<sup>6</sup>, mas tão somente procuraremos apresentar algumas das idéias de Popper e Kuhn, que, entre outras, precisam ser consideradas por quem se dedica à avaliação educacional, e que são necessárias para a compreensão do nosso posicionamento.

A partir de sua obra mais conhecida<sup>7</sup>, acreditamos seja o momento para algumas reflexões baseadas nas posições teóricas de Popper que dizem respeito tanto à pesquisa como à avaliação. Segundo Popper, as ciências, inclusive as do homem, baseiam-se em construtos teóricos que necessitam ser testados em sua veracidade, a fim de que sejam integrados à constelação conceitual que constitui o **corpus** científico. Os construtos — formulações hipotéticas sobre a natureza de um fenômeno —, como todos os produtos da concepção humana, devem ser analisados, inclusive experimentalmente, se possível, para que se possa determinar a consistência de sua estrutura lógica. A preocupação é comum a todos os avaliadores em educação, que se vêem na contingência de avaliar e pesquisar grande número de construtos, como rendimento escolar, compreensão de textos, atitudes, habilidade verbal, aptidão numérica etc, que nem sempre são fáceis de definir operacionalmente e de avaliar, diante da precariedade e limitação dos recursos instrumentais.

Precisamos esclarecer, dentro da perspectiva de Popper, que, quando um construto é estabelecido, a preocupação do cientista (e também a do avaliador) não se concentra na explicação de sua gênese, mas no conhecimento de sua lógica interna (validade) e de suas

<sup>6</sup> GUBA, E. G. (Ed.) *The paradigm dialog*. Sage Publications, Newbury Park, CA. 1990.

<sup>7</sup> POPPER, K. R. *A lógica da pesquisa científica*. In: Moritz Schlick, Rudolf Carnap e Karl R. Popper. *Coletânea de Textos. Os Pensadores*. Abril Cultural. Ed. Victor Civita. São Paulo, 1975.

relações com outros construtos, em um processo de confirmação ou contradição. O problema para o avaliador, na nossa concepção, não está na procura de como esse construto surgiu, mas sim no modo de analisá-lo de forma a permitir um pensamento crítico sobre a sua natureza.

A atividade do avaliador torna-se bastante complexa, exigindo o emprego de técnicas sofisticadas, quando se considera que um construto precisa ser comparado a outros para que se possa determinar se o novo construto representa efetivamente algum avanço em relação ao anterior; por outro lado, é preciso testar a aplicabilidade de suas conclusões na explicação dos fenômenos e comparar o construto a outros anteriormente comprovados, confirmando sua veracidade ou positivando sua falsidade. O problema implícito no campo da avaliação estaria na comprovação da validade dos instrumentos que supostamente estariam medindo os traços ou características definidas nos construtos. É necessária a construção de uma rede nomológica<sup>8</sup>, como colocam Cronbach e Meehl<sup>9</sup>. Popper, por sua vez, lembra com bastante propriedade que um construto nunca é permanente; outros, com o progresso do conhecimento, podem anulá-lo. Assim, os construtos devem ser reiteradamente verificados em suas diferentes dimensões para que sua validade seja ou não corroborada, o que implica trabalho de grande responsabilidade para o avaliador.

As idéias de Popper são bastante provocadoras e não podem ser examinadas em sua totalidade no âmbito do presente trabalho, nem é esse o seu objetivo. Gostaríamos, entretanto, de fazer referência a suas colocações sobre objetividade e indução, por suas implicações no âmbito da pesquisa e da avaliação educacional. A ciência, no dizer de Popper, preocupa-se com o objetivo e a objetividade seria a característica de um construto, no sentido de que é objetivo tudo aquilo que é possível de ser testado e permite repetições controladas para constatação de sua ocorrência efetiva. As teorias deveriam, assim, ser testadas nos seus fundamentos básicos, e não apenas verificadas,

---

<sup>8</sup> Nomológica, que é relativo à nomologia, estudo das leis que presidem os fenômenos naturais.

<sup>9</sup> CRONBACH, L. J. e MEEHL, P. E. Construct validity and psychological tests. *Psychological Bulletin*, 52, 1955.

antes da sua aceitação. Parece-nos que Popper fica preso ao cientifismo e ao positivismo, ignorando as limitações do mesmo, e que muitas teorias nem sempre podem ser testadas em seus fundamentos, sendo válidas por sua lógica interna e por explicarem, ainda que transitoriamente, um determinado fenômeno na área educacional e, em geral, no campo das ciências do homem. As teorias são basicamente construtos, um conjunto de idéias logicamente estruturadas sobre as quais hipotetizamos e procuramos aplicá-las na compreensão do mundo e de seus fenômenos.

É preciso não esquecer que Popper, inicialmente, pertenceu ao Círculo de Viena e que seu pensamento foi receptivo às idéias de David Hume (1711-1776) e Stuart Mill (1806-73); desse modo, pode-se compreender a propositura de sua questão sobre até que ponto a indução poderia ter um papel significativo para a compreensão da lógica do conhecimento e da sua veracidade. A indução, segundo a colocação de Popper, pouco representaria como princípio orientador da pesquisa do conhecimento científico, e deveria ser evitada, pois sua suposta lógica nem sempre seria consistente, permitindo, quando muito, inferências probabilísticas, que não possibilitariam assegurar a veracidade ou falsidade de um construto, mas apenas se seria provável ou não, o que não constituiria uma efetiva contribuição para o desenvolvimento do pensamento científico.

O problema do positivismo e especialmente a questão da importância da indução na avaliação educacional será objeto de consideração na discussão das idéias de Stake e da **avaliação responsiva**. Podemos, entretanto, antecipar que a discussão **dedução x indução** parece-nos hoje inteiramente superada tendo em vista o desenvolvimento da avaliação qualitativa, seu amplo emprego nos dias fluentes e o desenvolvimento de novas metodologias com uma abordagem naturalista, bastando considerar a fundamentação apresentada por Guba e Lincoln (1981, 1985 e, especialmente, a obra de 1989), que oferece novas perspectivas para a avaliação, partindo de uma abordagem que situa a avaliação em um amplo contexto humano, político, social e cultural, em que o pensamento indutivo vai desempenhar um papel relevante, dando um novo sentido à avaliação educacional.

Antes de um maior detalhamento sobre paradigmas em avaliação, apresentaremos algumas das idéias de Kuhn que tiveram grande repercussão a partir da publicação, no início da década de 60, do seu livro, hoje um clássico, — **A Estrutura das Revoluções Científicas**, que tem como idéia nuclear o conceito de paradigma. Kuhn apresentou a tese de que cada disciplina científica resolve seus problemas dentro de uma estrutura pré-fixada com base em pressupostos metodológicos, convenções lingüísticas e experimentos. A ciência assim desenvolvida, na argumentação kuhniana, acaba por entrar em choque com situações em que se vê impossibilitada de resolver um número cada vez maior de problemas com fundamento no paradigma então vigente — é uma situação de crise, que somente será superada com o surgimento de um **novo** paradigma.

A turbulência da crise gera novos paradigmas, o que traduz um avanço, apresentando inovações que superam os elementos geradores da crise e os problemas que a provocaram, segundo a concepção de Kuhn. Surge, assim, um novo paradigma. São estas crises, com a configuração de novos paradigmas, que possibilitam o progresso da ciência, na visão de Kuhn. O paradigma ptolomaico se exauriu e não conseguiu explicar os problemas da cosmologia no início dos Tempos Modernos. Surgiu, então, superando-o, o paradigma de Copérnico, depois o de Newton, mais tarde o de Einstein .... e assim caminhou e se desenvolveu a ciência, numa apresentação extremamente esquematizada.

As concepções de Kuhn abriram novas perspectivas, superando o positivismo em ciência, que supunha ser o saber científico resultado de um processo linear de identificação de verdades objetivas e da estruturação da ciência em torno dessas verdades. A concepção de Kuhn é bem diversa dos positivismos e cientificismos. É uma proposta que mostra o poder criativo da ciência ao passar de um paradigma para outro. Assim, na visão kuhniana, a função do paradigma seria constituinte e criadora da ciência, possibilitando o crescimento do saber.

Nem tudo foi aceitação e muitas das idéias de Kuhn foram contestadas, a partir do próprio uso da palavra paradigma, que, a acreditar em Masterman, citado por Norris (1990), é empregada em seu texto básico com 21 sentidos diferentes; por outro lado, há um



argumento maior que precisaria ser discutido nas suas implicações: — até que ponto as mudanças de paradigma na metodologia das ciências do homem teriam força suficiente para alterar a estrutura dos problemas na avaliação educacional? Ter-se-ia que pensar também na relação que existe entre o paradigma e a prática da avaliação. A maior preocupação dos avaliadores, como é o caso de Parlett e Hamilton (1973), centra-se nos paradigmas que englobariam visão de mundo, conjunto de crenças, novas maneiras de ver e outros mais que se revestem de um caráter que costuma ser categorizado como “metafísico”.

## 7. Paradigmas

A questão dos paradigmas merece um maior aprofundamento e passamos, então, a examiná-la em função do paradigma naturalista, segundo a linha de pensamento de Guba e Lincoln, cuja argumentação nos parece bastante coerente em termos da fase atual de desenvolvimento da avaliação; por outro lado, a apresentação constituirá o elo de ligação relativamente aos próximos assuntos, apresentados no presente capítulo, sobre **modelos, avaliação responsiva e avaliação iluminativa**.

Ao discutirem as bases epistemológicas e metodológicas da pesquisa naturalista, Guba e Lincoln (In: Madaus et al., 1993) tecem importantes considerações sobre os paradigmas em geral, o paradigma naturalista e o paradigma científico, que chamam de paradigma racionalista. Inicialmente, destacam a pesquisa naturalista como um paradigma de investigação, ou seja, um padrão ou modelo de como uma pesquisa pode ser conduzida. Mostram que a investigação naturalista é caracterizada por ocorrer em um contexto natural, utilizar o estudo de caso e apoiar-se principalmente em métodos qualitativos, que, entretanto, são aspectos que não a caracterizam devidamente. Os autores, desse modo, chamam a atenção para outras características, tendo em vista que o conjunto anterior não traduz realmente o que seja um paradigma. O local (contexto), o formato da pesquisa (estudos de caso) e a natureza do método não são suficientes para configurar um paradigma, na concepção de Guba e Lincoln, os quais deixam claro que os paradigmas são sistemas axiomáticos caracterizados por

diferentes pressupostos relacionados com fenômenos a investigar ou avaliar.

Guba e Lincoln (In: Madaus et al., 1993) mostram que existem diferentes paradigmas de pesquisa sobre os mais diversos aspectos da vida humana, os paradigmas científicos ou racionalistas. Outros paradigmas — naturalistas — estão surgindo em **oposição** aos anteriores. Argumentam que o suposto sucesso do paradigma racionalista nas ciências sociais e comportamentais é equivocado e perguntam-se onde estaria o resíduo de mais de 100 anos de pesquisas em psicologia e educação? Os pesquisadores enfrentam dificuldade em aplicar o próprio paradigma segundo seus pressupostos, citando o caso da amostra aleatória, como sendo virtualmente impossível por razões políticas e éticas. A questão do impacto das pesquisas científicas é posta de uma forma bem agressiva: **“The impact of research on practice is conspicuous by its absence...”** (p. 312), haja vista o caso das avaliações educacionais, cujos dados quase nunca são usados e a prática da escola baseia-se na experiência dos professores e dirigentes.

Guba e Lincoln (1993) questionam o valor do paradigma científico por refletir uma “epistemologia da ciência desacreditada — o positivismo” (p. 312), ainda que muitos continuem a agir como se o positivismo ainda fosse válido e persistem aceitando um posicionamento que é essencialmente analítico, reducionista, empiricista, associacionista, relativista, nomológico e monístico, acentuam os referidos autores. A alternativa para Guba e Lincoln (In: Madaus et al., 1993) seria a investigação/avaliação naturalista, por sua relevância contextual e riqueza, sua sensibilidade ao processo, e por apresentar uma teoria fundamentada em dados, uma teoria que explica os dados (e não uma teoria que se ajusta aos dados); finalmente, os autores argumentam com o poder do “homem como instrumento” da própria pesquisa/avaliação.

As conclusões de Guba e Lincoln são no sentido de que estamos no meio de uma revolução paradigmática, tendo em vista que o paradigma científico nem sempre se revela capaz de responder às questões das ciências sociais e comportamentais, chegando o momento de um novo paradigma, que, na opinião dos autores, é o naturalista, com uma nova visão da natureza da realidade, da natureza da relação

investigador-objeto, da natureza das sentenças verdadeiras, dos pressupostos de relações causais e, finalmente, do papel dos valores<sup>10</sup>.

## 8. Modelos

O uso da palavra **modelo** é de grande frequência em avaliação, apesar de não haver um conceito bastante preciso do seu significado, como tentaremos mostrar mais adiante. Se atentarmos para o seu emprego, verificaremos que pode se referir a uma determinada **concepção de avaliação**, mas em outros casos a mesma palavra passa a significar uma certa **forma de abordagem** e, numa terceira situação, o vocábulo é usado para traduzir um **método** empregado em avaliação. Acreditamos não ser simples encontrar uma definição que não seja passível de discussão, dada a diversidade de significados com que os teóricos da avaliação a empregam.

Os modelos em avaliação educacional, na verdade, descrevem o que os avaliadores fazem ou prescrevem o que devem fazer. A palavra modelo é usada em dois sentidos: 1) **prescritivo**, que é o mais comum, apresenta um conjunto de regras, prescrições, proibições, quadros referenciais que especificam o que é bom e próprio para a avaliação, e como a avaliação deve ser feita; 2) **descritivo**, conjunto de afirmações e generalizações que descrevem, prevêm e explicam as atividades da avaliação, conforme a colocação de Alken e Ellett (In: Walberg e Haertel, 1990).

Os modelos quase sempre se concentram na avaliação do ensino, aprendizagem e eficiência do currículo; outros, consideram atividades, práticas e políticas que facilitam o ensino, a aprendizagem e a eficiência do currículo; outros, ainda, preocupam-se com aspectos administrativos, financeiros e com as implicações sócio-políticas da educação. Tudo isso reflete a complexidade da avaliação, que abrange um amplo espectro de problemas, desde situações específicas ligadas à epistemologia e à natureza da pesquisa científica até problemas de relevância ligados à política educacional e à ética na avaliação.

---

<sup>10</sup> Ao final do estudo — **Avaliação Responsiva: Robert Stake** —, a ser divulgado em futuro próximo, apresentaremos um comentário relativo à avaliação naturalista.

Alken e Ellett descrevem em detalhes as características dos modelos prescritivos, suas preocupações com a metodologia, valores e uso dos resultados, elementos que dão margem a controvérsias e discussões entre os teóricos, especialmente no que diz respeito à ênfase a ser dada a cada uma das categorias. O suposto “modelo” de Scriven é utilizado para uma discussão aprofundada dos vários aspectos de um modelo que seria prescritivo. Quase todos os teóricos da avaliação acabam por apresentar, em certa medida, modelos que são prescritivos, pelo destaque que dão à metodologia, a valores e ao uso dos dados. As diferenças entre os modelos prescritivos estão na ênfase que cada um deles dá às diferentes categorias (aspectos), ora concentrando na metodologia, outros no aspecto valorativo e outros ainda dando maior ênfase ao uso; mas, qualquer que seja o destaque, os três elementos coexistem nos modelos prescritivos.

Alken e Ellet (In: Walberg e Haertel, 1990) chamam a atenção para as diferentes percepções de alguns modelos, que fazem da avaliação sinônimo de pesquisa experimental; outros destacam os dados quantitativos — condição **sine Qua non** da avaliação, nesta visão. Vários teóricos e avaliadores afirmam que a essência da avaliação estaria na quantificação (medida). Muitos, também, acreditam que somente os teste referenciados a critério são os instrumentos adequados para a avaliação educacional, havendo ainda quem advogue modelos causais, com procedimentos quase-experimentais. Enfim, as variações nos modelos prescritivos são as mais diversas, contemplando uma grande diversidade de situações e usando diferentes metodologias.

Quanto à questão dos valores, há quem advogue que ao avaliador cabe fazer julgamentos de valor, a partir dos dados coletados de diferentes fontes, ou com base nas suas vivências pessoais, no seu conhecimento e na sua experiência, mas o assunto não é pacífico; há discordâncias. Ao usuário é que caberia os julgamentos de valor. Relativamente à categoria uso, há aqueles que destacam, no contexto da avaliação, a preocupação com os responsáveis pela tomada de decisão, enquanto outros centram seu enfoque nos usuários. Não deixam de considerar a metodologia e os valores, mas seu objetivo centra-se, especialmente, em coletar dados que sejam relevantes para os responsáveis pelas decisões, tendo em vista o relacionamento

dinâmico que existiria entre usuários e avaliadores. Uma observação deve ser feita — os que advogam a ênfase no uso, consideram que há uma grande identidade entre avaliação e pesquisa.

Algumas considerações devem ser feitas sobre os modelos descritivos em avaliação, ou seja, modelos que seriam uma “teoria” e corresponderiam a um conjunto de afirmações empíricas com a finalidade de descrever, prever ou explicar atividades de avaliação. Observe-se, como fazem Alken e Daillak (In: Walberg e Haertel, 1990), que nos últimos anos, a partir da década de 70, não há estudos que se direcionem primordialmente para o desenvolvimento de uma teoria geral da avaliação. A maior parte dos estudos se concentra em aspectos particulares da avaliação, como os estudos de caso, cujas formulações teóricas se baseiam em dados coletados em campo.

Stake (In: Madaus et al., 1993), ao discutir avaliação de programas, mostrou que, na década de 60, havia dois conjuntos básicos de modelos: um, informal, baseado nas percepções (**insights**) de profissionais e autoridades; outro, preferido da maioria dos avaliadores (Tyler, Husen, entre outros), o do pré-teste/pós-teste. Stake, ainda que de forma simplificada, analisa nove diferentes abordagens de atividades de avaliação, entre as quais destacaremos cinco delas, por serem de interesse imediato, conforme a apresentação a seguir.

Antes de discutir a sua proposta de avaliação responsiva, Stake, com justa razão, procura destacar a diferença entre abordagem preordenada (**preordinate**) e abordagem responsiva, cujos conceitos seriam mais importantes que os de avaliação formativa e avaliação somativa, apresentados por Scriven. Antes, porém, Stake chama a atenção para o fato de que existem diferentes maneiras de avaliar um programa educacional e que nenhuma é inteiramente correta: não apresentam uma descrição de seus méritos e deficiências, ignoram importantes questões que deveriam ser levantadas durante a avaliação; mostram-se pouco sensíveis às singularidades das condições locais e ao clima proporcionado para a aprendizagem. Stake, assim, declara que prefere uma abordagem de avaliação que desempenhe um serviço e que seja útil a pessoas específicas.



**ALGUNS MODELOS DE AVALIAÇÃO, segundo STAKE  
(In: Madaus et al., 1993)**

ABORDAGEM	OBJETIVOS	PRINCIPAIS ELEMENTOS	COMPETÊNCIA (ênfase)	PROTAGONISTAS	RISCOS	VANTAGENS
Medida de ganhos do estudante	Medir o desempenho e o progresso do estudante	Definição de objetivo, análise de escores, discrepância entre escores e a realidade	Psicólogos educacionais	R. TYLER B. Bloom W. Popham M. Provas	Super-simplificação dos objetivos educacionais, não consideração do processo	Destaca, identifica o progresso do estudante
Interação / Observação	Compreender atividades e valores	Questões educacionais, observação em sala de aula, estudo de caso, visão pluralista	Usuários do sistema educacional	L. Smith Parlett-Hamilton R. Rippey R. STAKE	Grande apoio em percepções subjetivas; ignora causas	Mostra amplo quadro do programa; identifica conflitos de valores
Análise gerencial	Aumentar a racionalidade e nas decisões do dia-a-dia	Relação de opções, estimativa, retroinformação contínua, custos, eficiência	Administradores Economistas	L. Lessinger D. STUFFLEBEAM M. Alkin	Super-valorização da eficiência, subestimação de elementos implícitos	Retro-alimentação para tomada de decisão
Pesquisa Educacional	Desenvolver explicações e técnicas de instrução	Condições controladas, análise multivariada, base para generalização	Experimentalistas	L. CRONBACH J. Stanley D. Campbell	Condições artificiais; ignora aspectos humanísticos	Novos princípios instrucionais e desenvolvimento de materiais
Avaliação sem objetivos pré-estabelecidos (goal-free)	Avaliar efeitos de um programa	Ignora objetivos propostos; segue uma check-list	Consumidores em potencial	M. SCRIVEN	Supervaloriza documentos e relatórios	Informa os efeitos, minimiza fatores subjetivos

Os modelos ou planejamentos preordenados de avaliação destacam objetivos instrucionais, usam testes objetivos, estabelecem padrões e fazem relatórios típicos da pesquisa educacional. O modelo responsivo de avaliação propõe-se a fazer tudo aquilo que qualquer pessoa faz quando avalia, ou seja, observa e reage. É subjetivo, mas essa subjetividade — principal argumento dos que a ele se opõem — pode ser reduzida, na opinião de Stake, pela definição operacional de alguns termos, entre outros aspectos, ainda que seja um trabalho que se baseie, fundamentalmente, nos **insights** da observação pessoal.

## 9. Avaliação responsiva

Uma avaliação é responsiva, na proposta de Stake, se sua orientação está voltada para as atividades do programa e menos para os seus objetivos; se ela corresponde à necessidade de informações propostas por diferentes públicos, e se diferentes perspectivas de valor são apresentadas no relatório sobre o sucesso ou fracasso do programa. O ponto central numa avaliação é o plano de observação para a coleta e o registro dos dados, que devem ser de qualidade para que um retrato fiel das ações e reações possa ser apresentado e discutido.

O grande problema é que as observações estão direta ou indiretamente relacionadas a problemas de valor e que para cada um dos propósitos do programa há necessidade de coletar dados, pois um único conjunto de informações não é suficiente para uma avaliação. Na avaliação responsiva, dentro da proposta de Stake, cada avaliador, em cada situação, deve decidir o que espera obter. Uma das suas decisões está certamente presa à convicção de que o valor de um programa pode não estar relacionado à mensuração dos resultados. Testes e outros instrumentos nem sempre são fundamentais, mas não devem ser aprioristicamente desprezados. A observação sistemática, diz Stake, e as entrevistas devem contribuir para a compreensão do problema identificado. O importante é que o avaliador disponha de um número considerável de fontes independentes e confiáveis de informação a fim de que possa ter uma percepção do desenvolvimento de um programa e do seu êxito.

Como proceder para realizar uma avaliação responsiva? Stake (1993) detalha os vários procedimentos e apresenta os diferentes momentos de uma avaliação responsiva, que podem ser resumidos da forma seguinte: 1) discutir com os responsáveis pelo programa; 2) identificar o escopo do programa; 3) analisar as atividades do programa; 4) identificar propósitos e preocupações; 5) conceituar questões e problemas; 6) identificar necessidades de dados; 7) selecionar observadores, juízes e até mesmo instrumentos, se necessários; 8) observar antecedentes, interações e resultados, como no modelo **countenance**; 9) preparar descrições, estudos de caso; 10) validar — confirmar ou “desconfirmar”; 11) elaborar informações para vários públicos; 12) apresentar relatório formal, se for o caso.

O que importa é o que está ocorrendo no programa, inclusive a reação dos alunos e discussões posteriores. O avaliador deve procurar padrões, estruturas, fatos significativos, como ocorreram em condições não controladas ou modificadas pelo avaliador, e que não se devem basear apenas no seu poder de avaliação, julgamento e na sua capacidade de reagir a acontecimentos, mas valer-se da colaboração de estudantes, professores e outros mais que possam oferecer informações para o problema em estudo e que sejam de interesse para as audiências.

A capacidade de apresentar um relatório decodificado é importante, tendo em vista que é uma oportunidade para apresentar uma experiência vicária, que é o melhor substituto de uma experiência direta, ou seja, o avaliador, numa avaliação responsiva, deve ser um contador de histórias, dominar a **art of storytelling**, a fim de retratar múltiplas realidades, em toda a sua complexidade e refletir o mistério da sua experiência.

Apesar da importância da avaliação responsiva e muitas vezes ser mais crucial avaliar o valor intrínseco de uma experiência e não exclusivamente os seus resultados. Stake de forma bastante transparente mostra que, em certas circunstâncias, a investigação necessita da tecnologia do **survey** e da especificação de objetivos e, textualmente, afirma que os trabalhos de John Tukey, Torsten Husen, Ralph Tyler, Benjamin Bloom e James Popham continuarão servindo para estudos dessa natureza, considerando que diferentes modos de avaliar servem a diferentes propósitos (In: Madaus et al., 1993).

A Avaliação Responsiva, em conclusão, é particularmente útil durante a avaliação formativa, quando há necessidade de um monitoramento do programa, e quando ainda não se tem certeza absoluta dos problemas que poderão surgir. Ao mesmo tempo, pode ser de grande utilidade numa avaliação somativa, quando diferentes grupos de pessoas desejam uma compreensão mais aprofundada das atividades dos programas, seus méritos, suas prováveis deficiências, e, também, quando o avaliador sente que é sua responsabilidade proporcionar uma experiência vicária, como bem coloca Stake no seu magnífico ensaio (In: Madaus et al., 1993).

Tudo isso nos conduz a refletir sobre um novo posicionamento, que está ligado à avaliação responsiva, à avaliação que, no fundo, é uma reação aos excessos das tecnalidades: — a avaliação iluminativa.

### 10. Avaliação iluminativa

A avaliação iluminativa, na concepção de Parlett (In: Walberg e Haertel, 1990), procura investigar e interpretar as práticas educacionais, as experiências dos participantes do processo educacional, os procedimentos institucionais e os problemas gerenciais de uma forma que seja útil para aqueles a que o estudo investigativo se destina. O avaliador vai contribuir para a tomada de decisão proporcionando informações, comentários e análises que aumentam o conhecimento e a compreensão dos problemas educacionais. A sua metodologia é flexível e usa técnicas diversas conforme a natureza do problema investigado<sup>11</sup>.

A avaliação iluminativa pode ser caracterizada como sistêmica, ou seja, é uma abordagem holística, preocupando-se, assim, com a totalidade das inter-relações existentes nos fenômenos educacionais, não se fixando em aspectos limitados de um programa ou nas correlações entre variáveis individuais. A avaliação iluminativa, como abordagem sistêmica, tem como preocupação analisar a estrutura complexa dos fenômenos. Admite que os acontecimentos devam ser vistos nas suas interrelações; aceita a possibilidade de múltipla causalidade; e tem como objetivo explicar a complexa rede de causa e efeito.

Devemos considerar a avaliação iluminativa como um esforço integrado de investigação em que cada elemento analisado contribui para a compreensão do todo; é uma abordagem que apresenta diferentes características, sendo **responsiva, naturalista, heurística e interpretativa**. É responsiva na medida em que procura oferecer elementos aos interessados que satisfaçam às suas exigências de informação, aos seus interesses, aos seus questionamentos, às suas necessidades, tudo isso de uma forma que não afronte o bom senso, usando uma linguagem sem jargão tecnológico, simples, clara e

---

<sup>11</sup> Ver a discussão em Hamilton et al. *Beyond the Numbers Game* (1978).

---

precisa, sem maiores divagações elucubrativas. Além disso, a avaliação iluminativa procura apresentar seus resultados sem maiores digressões, com uma certa economia verbal e sem maiores delongas, apesar de algumas vezes usar técnicas antropológicas, que demandam tempo para sua execução. A avaliação iluminativa é **naturalista** na medida em que estuda os fenômenos ao ocorrerem “naturalmente” em um contexto real, sem a preocupação de criar condições artificiais, para, supostamente, melhor realizar o estudo. É naturalista na medida em que se fixa no que é inesperado e atípico, sem ficar presa exclusivamente ao rotineiro e ao óbvio, no dizer de Parlett (In: Walberg e Haertel, 1990). A avaliação iluminativa, no seu aspecto naturalista, apresenta fatos, material estatístico, opiniões, observações, elementos históricos, enfim, uma grande diversidade de elementos com vistas a aumentar a compreensão relativa às características do complexo analisado. A avaliação iluminativa é heurística porque **demand** um planejamento que traduza uma estratégia de investigação continuamente atualizada à medida em que o investigador tem que se ajustar às novas emergências de situações do sistema como um todo, ao mesmo tempo em que se adapta às mudanças do sistema que geram um novo fluxo de eventos no decorrer dos trabalhos e que precisam ser estudados e analisados. A avaliação iluminativa tem também um caráter **interpretativo** tendo em vista que não se limita a descrever e a fazer um relatório, vai bem mais além, promove uma discussão dos resultados sob diferentes perspectivas, procura desvendar as complexidades que se apresentaram, separa o que é realmente significativo do trivial, no dizer de Parlett e Hamilton (1972), e procura apresentar uma discussão daquilo que é efetivamente relevante.

A avaliação iluminativa, na visão de Parlett (In: Walberg e Haertel, 1990), está intimamente relacionada à avaliação responsiva, compartilhando com esta a preocupação de informar àqueles que devem tomar decisões; desse modo, procura apresentar os detalhes e os aspectos mais complexos dos programas que investiga, usando para esse fim diferentes técnicas de investigação para que possa oferecer um retrato detalhado do objeto que estudou, seja ele um sistema ou um caso isolado. A avaliação iluminativa preocupa-se com quem irá ler o relatório, o que esse leitor necessita efetivamente conhecer e o que



significa uma análise detalhada para atendimento das várias clientelas. A avaliação iluminativa reveste-se, desse modo, de fundamental importância, porque procura condensar o máximo possível de experiências válidas e de comentários informativos em um relatório que estimule a leitura, a discussão e a análise de diferentes questões que precisam ser resolvidas e, possivelmente, a formulação de políticas educacionais ligadas às práticas da educação, não ficando restrita a altas especulações de aspectos nem sempre relevantes, como muitas vezes acontece em avaliações de grande impacto na **mídia** sensacionalista, mas de repercussão mínima na comunidade. O caminho da avaliação no contexto nacional, segundo o nosso juízo pessoal, está na estruturação de um sistema que seja responsivo e iluminista, para que seja compreensível às diversas clientelas e possa solucionar os nossos problemas, que são graves e exigem pronta ação.

---

## 1.1 NOTAS E COMENTÁRIOS

[1] O problema da subjetividade em avaliação educacional é recorrente, sempre surge como tema para discussão entre especialistas e o mundo laico. A subjetividade parece-nos inevitável na apreciação de qualquer fenômeno. Diferentes pessoas diante de uma mesma situação têm percepções diferentes, daí, por conseguinte, a inevitabilidade do subjetivismo. Stake (1983a), ao abordar a questão, coloca uma pergunta que nos parece bastante relevante: *“por que, então, não utilizar a subjetividade de forma a que possamos melhorar nossos trabalhos de pesquisa e avaliação?”*. Stake, em outro de seus numerosos trabalhos (1983b) argumenta: *“A subjetividade precisa ser reconsiderada. Penso que a subjetividade do observador de campo pode ser controlada, mas não eliminada. Controlada e usada, entretanto, de um modo vigoroso a fim de fazer com que as conclusões das pessoas sejam mais relevantes e úteis.”* Qualquer que seja o “*approach*” seguido pelo avaliador, a questão da subjetividade está presente, e é vista por Stake (1983b) como uma característica da avaliação: *“A avaliação é subjetiva. São necessários dados consistentes em estudos de avaliação, mas os avaliadores com sensibilidade acham que a utilidade da subjetividade está sendo cada vez mais valorizada. O estudo de caso é um modo de fazê-lo.”*

[2a] A palavra *objetivo* costuma apresentar diferentes significados, o que fica claro em língua inglesa, que apresenta uma sinonímia bastante representativa (**objective, goal, outcome, aim, intention** etc). Se considerarmos do ponto de vista estritamente técnico, como o faz Eraut (In: Walberg e Haertel, 1990), objetivo deve ser considerado como fazendo referência a algo que foi pré-

definido no planejamento de um programa de ensino e que expressa aquilo que se espera que o aluno tenha aprendido após um período de orientação da aprendizagem. Às vezes, em inglês, o seu significado se complica, dado que as palavras usadas como sinônimos e que foram anteriormente apresentadas nem sempre têm a mesma conceituação, como é o caso de **goal**, que se refere especificamente a objetivos gerais da educação, e **outcome** traduz mais um objetivo específico, resultado da aprendizagem.

A questão dos objetivos não é tranqüila, sendo objeto de críticas do ponto de vista teórico e prático. O emprego dessa palavra surge em discussões sobre currículo, em obra de F. Bobbit (1918), por influência das idéias de Herbert Spencer (1860). Ainda que na década de 20 a questão dos objetivos tivesse tido grande importância na área de currículo, sobretudo graças à atuação de Chartes (1924), passa, no início dos anos 30, a sofrer um certo colapso, mas volta a ter algum ímpeto na década de 40, por influência de Ralph Tyler, ao colocar a questão da avaliação na época do **Eight -Year Study**. Um antigo discípulo de Tyler, Benjamin Bloom, juntamente com outros educadores, publica suas célebres **Taxonomias** (1956), abordando objetivos cognitivos e, depois, objetivos não-cognitivos. Deve-se lembrar que Tyler e Bloom apresentaram os objetivos educacionais em termos gerais. Foi Mager (1962), ao escrever **Preparing Instructional Objectives**, obra traduzida para o português e que teve alguma influência (nefasta) no Brasil, o responsável pela volta às posições **behavioristas** defendidas por Bobbit, em 1918. A sua influência foi grande ao preparar objetivos para instrução programada, sendo o grande iniciador da reação a certos verbos, como **saber, conhecer, compreender, sentir, apreciar**, entre outros, que passaram a ser **taboo** em alguns segmentos da avaliação educacional. O apoio de Gagné (1965) à definição operacional de Mager (1962) contribuiu grandemente para a sua divulgação no meio educacional, como mostra Eraut (In: Walberg e Haertel, 1990), entretanto, é bom lembrar que Tyler tinha como preocupação orientar os professores e os responsáveis pela elaboração de questões (itens), enquanto Mager e Gagné se concentraram no planejamento instrucional, especialmente em relação ao ensino programado.

Eraut, no seu artigo sobre objetivos educacionais, (In: Walberg e Haertel, 1990), apresenta uma ampla discussão sobre níveis de especificação dos objetivos, analisa a controvérsia sobre o uso de objetivos e suas limitações, seu emprego no planejamento de currículos e do ensino, e, particularmente, no caso da avaliação de programas, destacando sobretudo a sua importância no exame da concretização de intenções.

Ao falarmos em objetivos sempre nos surgem à mente a questão das **Taxonomias**, especialmente pela importância e o mau uso da Taxonomia de Bloom no contexto brasileiro. O artigo de De Landsheere (1990), na enciclopédia organizada por Walberg e Haertel, é bastante elucidativo, analisando, na área cognitiva, a Taxonomia de Bloom (1956), o modelo da

estrutura do intelecto de Guilford (1967), a Taxonomia de Gagné-Merrill (1971), a Taxonomia de Gerlach e Sullivan (1967) e a de De Block (1975). Os objetivos na área não-cognitiva (afetiva) também são vistos, concentrando-se na Taxonomia de Krathworl (1964), sendo que De Landsheere oferece várias sugestões no sentido do seu aprimoramento, além de chamar a atenção para a importância de estudos na área afetiva. A partir de uma discussão sobre a importância dos objetivos na área psico-motora, De Landsheere apresenta uma visão sumária das Taxonomias que foram delineadas por Rapsdale (1950), Guilford (1958), Dave (1969) e Kliber (1970), detendo-se mais longamente nas Taxonomias de Simpson (1966) e Harrow (1972), detalhando sobretudo esta última que parece ser a que melhor se aplica ao estudo do domínio psicomotor. Ressalta, ainda, que certas categorias descritivas não são plenamente satisfatórias. Apesar da importância desse movimento de elaboração de taxonomias, as duas mais significativas, e que merecem ser exploradas são, na verdade, a de Bloom (1956) e a de Harrow (1972), na nossa apreciação. Queremos acentuar que nem Bloom e nem Harrow devem ser consideradas bíblias a serem seguidas com absoluta fidelidade; são obras importantes, mas são apenas manuais de referência que merecem ser consultados na elaboração de um instrumento ou no caso do esclarecimento de uma dúvida.

- [2b] A questão dos objetivos instrucionais, ainda hoje objeto de controvérsias, inclusive no contexto brasileiro, passou a ser considerada mais amplamente na área da avaliação a partir dos anos 60. Se observados do ponto de vista histórico, os objetivos instrucionais atraíram inicialmente a atenção dos psicólogos da educação interessados em problemas da aprendizagem e no desenvolvimento de tecnologias instrucionais. A partir de Skinner (1958), com o seu interesse pela instrução programada e a possibilidade de "máquinas de ensinar", que tanta controvérsia suscitaram, foi que os objetivos instrucionais passaram a merecer uma atenção mais acurada dos educadores. O primeiro a ocupar-se em escrever sobre o assunto foi, sem dúvida, Robert Mager (1962), com o objetivo de familiarizar os educadores com a construção de objetivos. Originalmente, o seu trabalho se chamava **Preparing Objectives for Programmed Instruction**, pois seu interesse estava concentrado em instrução programada, mas diante do sucesso alcançado pelo livro, o título foi modificado para **Preparing Instructional Objectives**, tendo, inclusive, uma tradução brasileira, divulgada pelo SENAI. O êxito alcançado pelo modelo de Tyler, no **Eight -Year Project**, e a publicação de seu livro sobre currículo (1949) igualmente bem sucedido, associados à bem sucedida obra de Mager e às imposições do **US Congress of Elementary and Secondary Education Act — ESEA** (1965), que estabeleciam a obrigatoriedade de avaliações para todos os projetos que recebessem verbas federais, foram fatores que contribuíram intensamente para a divulgação da avaliação orientada por

objetivos no contexto norte-americano e se expandissem para outros contextos, inclusive o brasileiro.

O artigo de Popham, (In: Walberg e Haertel, 1990), oferece uma boa análise da história dos objetivos educacionais em avaliação, apresentando, entre outros aspectos, as discussões sobre objetivos comportamentais; a elaboração de taxonomias, especialmente a de Bloom (1956) sobre o domínio cognitivo; o emprego de objetivos nas medidas, inclusive apresentando a discussão de Robert Glaser (1963), que foi o primeiro a caracterizar os instrumentos de medida como referenciados a norma e a critério; a questão dos padrões de desempenho que passou a ser seriamente considerada a partir dos anos 70; finalmente, Popham oferece uma série de orientações bastante pertinentes aos avaliadores que usam objetivos em suas atividades profissionais.

- [3] O problema da validade interna e da validade externa está relacionado a possíveis distorções do significado das conclusões em um **design** experimental. A **validade interna** pode estar ameaçada porque a conclusão sobre a diferença de tratamentos pode resultar de outras explicações que não a esperada para o caso sob estudo. Assim, o objetivo primeiro de uma distribuição aleatória dos sujeitos, em um **experimental design**, é o de controlar a ameaça do fator **seleção** sobre a validade interna; ou seja, o resultado final pode refletir diferenças que simplesmente existam no grupo inicial, que, depois, sofreu um processo de **randomização** no experimento. A **validade externa** refere-se à validade da generalização dos resultados de um estudo para os sujeitos da população. A validade externa da generalização vai depender do grau com que a amostra é efetivamente representativa da população. Existem procedimentos formais de amostragem aleatória que, associados a técnicas de estatística inferencial, permitem afirmações sobre o grau de **não-representatividade** em decorrência do acaso.

O artigo de R. Tate sobre **experimental design**, na Enciclopédia de Avaliação, editada por Walberg e Haertel (1990), é bastante explicativo ao analisar as formas de inferências sobre validade causal em experimentos aleatórios. O assunto é muito desenvolvido por Cronbach (1982), especialmente no capítulo 4, páginas 112 a 132, em que parte de uma citação de Campbell e Stanley, na sua obra igualmente clássica (**Experimental and Quasi-experimental Designs for Research on Teaching**), sobre a importância fundamental em estabelecer distinção entre **validade interna**, o mínimo básico sem o qual qualquer experimento não pode ser interpretado, ou melhor, nas suas palavras, é **uninterpretable**, e **validade externa**, que parte de perguntas sobre a generalização dos resultados de um experimento. Uma análise clara do problema acha-se apresentada por Fraenkel, J. R. e Wallen, N. E., em trabalho recente (3ª edição) sobre **How to Design and Evaluate Research in Education**. McGraw-Hill, Inc. New York, 1996, pp. 241 a 255 (validade interna) e pp. 105-109 (validade externa).

- [4] Há uma discussão na área educacional que possivelmente nunca será resolvida, mas que tem implicações na avaliação: — a educação é uma arte ou uma ciência? A avaliação qualitativa, com sua visão holística, aproxima-se da primeira posição; a avaliação quantitativa, com a sua concepção de uma realidade tangível, convergente e fragmentável, posiciona-se favorável à educação como ciência. Kopperman (1993) mostrou que o modelo de Tyler, apesar de admirável nos diversos aspectos de seu planejamento, não previu e nem tem condições de lidar com os objetivos inesperados do processo educacional, malgrado sua importância.

A abordagem antropológica da avaliação, como mostra Kopperman (1993), oferece uma solução para os dois problemas anteriormente apresentados. A antropologia, nas suas diversas linhas de ação, consegue realizar uma síntese das perspectivas artística e científica, combinando os **insights** da observação subjetiva com o rigor científico dos processos de análise. A etnografia, por sua vez, nos estudos realizados em uma comunidade, pode incorporar aspectos sutis do ambiente educacional, como os objetivos não-previstos, o que é importante, apesar de o estudo etnográfico possuir a limitação de não poder avaliar comportamentos mensuráveis; contudo, esse aspecto não lhe diminui o valor, como metodologia de ação, cujo objetivo é compreender um fenômeno e cuja preocupação maior é fenomenológica.

## 2. Bibliografia

- ALKIN, M. C. e DAILLAK, R. H. Impact of Evaluation Studies. In: WALBERG, H. J. et al. **The International Encyclopedia of Educational Evaluation**. Oxford, Pergamon Press. 1990.
- ALKIN, M. C. e ELLETT Jr., F. S. Development of Evaluation Models. In: WALBERG, H. L. et al. **The International Encyclopedia of Educational Evaluation**. Oxford. Pergamon Press. 1990.
- CRONBACH, L. J. **Designing Evaluation of Educational and Social Programs**. San Francisco, Jossey-Bass. 1982.
- DE LANDSHEERE, D. V. Taxonomies of Educational Objectives. In: WALBERG, H. J. et al. **The International Encyclopedia of Educational Evaluation**. Oxford, Pergamon Press. 1990.



- ERAUT, M. R. Educational Objectives. In: WALBERG, H. J. et al. **The International Encyclopedia of Educational Evaluation**. Oxford. Pergamon Press. 1990.
- GLASS, G. V. e WORTHEN, B. R. Evaluation and research: similarities and differences. **Curriculum Theory Network**, Fall. 1971.
- GUBA, E. G. e LINCOLN, Y. S. **Effective Evaluation**. San Francisco: Jossey-Bass. 1981.
- GUBA, E. G. e LINCOLN, Y. S. **Fourth Generation Evaluation**. Newbury Park, California. Sage Publications. 1989.
- GUBA, E. G. e LINCOLN, Y. S. Epistemological and Methodological Bases of Naturalist Inquiry. **Educational Communications and Technology Journal**, nº 4, 30, Winter, 1982. In: MADAUS, G. F. et al. **Evaluation Models**. Boston. Kluwer-Nijhoff Publishing. 1993.
- HAMILTON, D. et al. **Beyond the Numbers Game: a reader in educational evaluation**. London: MacMillan Education. 1978.
- KOPPELMAN, K. L. The Explication Model: An Anthropological Approach to Program Evaluation. **Educational Evaluation and Policy Analysis**, nº 3, 1. July-August, 1999, p. 59-64. In: MADAUS, G. F. et al. **Evaluation Models**. Boston. Kluwer-Nijhoff Publishing, 1993.
- LINCOLN, Y. S. e GUBA, E. G. **Naturalistic Inquiry**. Newbury Park, California. Sage Publications. 1985.
- NORRIS, N. **Understanding Educational Evaluation**. Kogan Page. London. 1993. Published in Association with CARE — School of Education, University of East Anglia.

- PARLETT, M. R. Illuminative Evaluation. In: WALBERG, H.L. et al. **The International Encyclopedia of Educational Evaluation**. Oxford. Pergamon Press. 1990.
- PARLETT, M. e HAMILTON, D. Evaluation as illumination: a new approach to the study of innovatory programs. In: GLASS, G.V. (Ed.) **Evaluation studies review annual**. Vol. 1. Beverly Hills, California. Sage. 1973.
- PAYNE, D. A. (Ed.) **Curriculum Evaluation: Commentaries on Purpose, Process, Product**. Lexington, Mass. D.C. Heath and Company. 1974.
- POPHAM, W. J. **Educational Evaluation**. Englewood Cliffs, New Jersey. Prentice-Hall Inc. 1975.
- SCRIVEN, M. **The Methodology of Evaluation**. Perspectives of Curriculum Evaluation. AERA. Monograph 1, Rand McNally and Co., Chicago. 1967.
- STAKE, R. E. The countenance of educational evaluation. **Teachers College Record**, 68, nº 7, p.523-540. 1967.
- STAKE, R. E. Program Evaluation, Particularly Responsive Evaluation. (Paper) Conference on New Trends in Evaluation, Göteborg, Suécia, 1973. In: MADDAUS, G. F. et al. (Ed.) **Evaluation Models**. Boston. Kluwer-Nijhoff Publishing. 1993.
- STAKE, R. E. Pesquisa qualitativa/naturalista — problemas epistemológicos. **Educação e Seleção**, jan/jun., nº 7. 1983a. Trad. de H. M. Vianna. São Paulo. Fundação Carlos Chagas.
- STAKE, R. E. Estudos de caso em pesquisa e avaliação educacional. **Educação e Seleção**, jan/jun., nº 7, 1983b. Trad. de H. M. Vianna. São Paulo. Fundação Carlos Chagas.

- STAKE, R. E. Responsive Evaluation. In: WALBERG, H.J. et al. **The International Encyclopedia of Educational Evaluation**. Oxford. Pergamon Press. 1990.
- STAKE, R. E. Case Studies. In: DENZIN, N. e LINCOLN, Y. **Handbook of Qualitative Research**. Sage Publications. Thousand Oaks, California, 1994.
- STUFFLEBEAM, D. L. The relevance of the CIPP Evaluation Model for Educational Accountability. **Journal of Research and Development in Education**, 5 (1), p. 19-25. 1971.
- TYLER, R. W. General Statement on Evaluation. **Journal of Educational Research**, 35, p. 492-501. 1942.
- TYLER, R. W. Considerations in Selecting Objectives. In: PAYNE, D. (Ed.) **Curriculum Evaluation: Commentaries on Purpose, Process, Product**. Lexington, Mass. D.C. Heath and Company. 1974.
- WALBERG, H. J. e HAERTEL, G. D. **The International Encyclopedia of Educational Evaluation**. Oxford. Pergamon Press. 1990.
- WORTHEN, B. R.; SANDERS, J.R. **Educational Evaluation: theory and practice**. Belmont, California. Wadsworth. 1973.

