

# AVALIAÇÃO EDUCACIONAL: UMA PERSPECTIVA HISTÓRICA<sup>1</sup>

## 1. INTRODUÇÃO

A pesquisa e a avaliação têm um significado especial no delineamento do processo decisório em educação. A pesquisa, inicialmente, e, depois, a avaliação, especialmente a avaliação de programas, projetos e produtos, passaram a ter uma dimensão maior a partir do século XX, sofrendo ambas – pesquisa e avaliação – a influência de diferentes ciências, como por exemplo, a psicologia, a psicomетria, a sociologia, a antropologia, a etnografia e a economia, entre outras, de determinaram novos enfoques metodológicos com base em vários posicionamentos teóricos.

A avaliação, após o trauma provocado pela constatação da deficiência tecnológica associada à carência educacional, no mundo ocidental, especialmente nos Estados Unidos, com o lançamento do Sputnik, no dia 4 de outubro de 1957, tornou-se impositiva. Há todo um esforço no sentido de recuperar o tempo educacional que fora perdido, sendo criados novos currículos, e a avaliação passou a ter papel de relevância no desenvolvimento de novas estratégias de ensino<sup>2</sup>.

A literatura sobre avaliação educacional, a partir dos anos 60, cresce enormemente. Surgem novos posicionamentos

<sup>1</sup> Artigo publicado na revista *Estudos em Avaliação Educacional*, São Paulo, n. 12, p. 7-24, jul./dez. 1995.

<sup>2</sup> Alguns desses novos currículos em Matemática (*Mathematics Study Group* - MSG), em Física (*Physical Science Study Committee* - PSSC), em Química (*Chemical Bond Approach* - CBA) e em Biologia (*Biological Study Committee* - BSC), entre outros, foram traduzidos e adaptados ao contexto nacional pela Fundação Brasileira para o Ensino de Ciências (FUNBEC) e no Centro de Treinamento de Professores de Ciências (CECISP), graças aos esforços dos professores Isaías Raw, Myrian Krasilchik e Norma Maria Cleffi, com a colaboração de professores brasileiros, no início da década de 60.

teóricos e novas propostas de atividades práticas, que enriquecem e tornam complexo o campo da avaliação. Algumas questões fundamentais são levantadas, com profundas implicações na sua teoria e na sua prática. É justo, portanto, que, inicialmente, seja discutida a sua evolução histórica a partir da perspectiva norte-americana, em cujo contexto se desenvolveram numerosos e importantes trabalhos, em particular em decorrência da ação seminal de Ralph W. Tyler. Outras fontes, entretanto, também serão indicadas, especialmente inglesas e, subsidiariamente, trabalhos de autores nacionais, ainda que a produção teórica destes últimos seja bastante reduzida no seu aspecto quantitativo.

#### 1.1. AVALIAÇÃO EDUCACIONAL NOS ESTADOS UNIDOS: UM ESBOÇO

A sociedade norte-americana sofreu grandes transformações estruturais em decorrência do impacto advindo da Revolução Industrial e, a exemplo do que ocorreu na Inglaterra, o poder público passou a discutir diferentes programas sociais, inclusive de natureza educacional. A avaliação nos Estados Unidos possui uma tradição de quase dois séculos, ainda que o seu momento mais intenso tenha ocorrido a partir da década de 1960. É também nesse período que a avaliação começa a se definir como uma profissão estruturada. A fim de chegar a essa situação, a avaliação atravessou diferentes momentos, conforme Madaus, Stufflebeam e Scriven (1993).

A avaliação associada ao processo educativo sempre existiu sob diferentes formas (WORTHEN; SANDERS, 1987), mas somente adquire uma natureza formal quando Horace Mann (1845) inicia a prática da coleta de dados para a fundamentação de decisões de políticas públicas que afetam a educação. A avaliação, aos poucos, começa a desenvolver procedimentos ainda utilizados nos dias de hoje, como o *survey*, e instrumentos objetivos padronizados. Além disso, determinadas práticas ainda persistem, porque baseadas na obtenção de escores dos alunos como principal elemento para a avaliação da eficiência de programas educacionais, aliás, traço característico das avaliações realizadas no atual contexto brasileiro (1995).

Houve, nessa época, como mostram Madaus *et al.* (1993), uma politização dos dados obtidos através de provas escritas, que foram usados para comparar escolas e promover o afastamento de diretores que se opunham ao programa de eliminação de castigos físicos que então existiam, e cuja supressão era defendida por Samuel G. Howe e Horace Mann. A primeira avaliação propriamente dita somente teve lugar bem mais tarde, no final do século XIX, entre 1887 e 1898, quando Joseph Rice procurou verificar a influência do tempo dedicado a exercícios (*drill*) no processo de alfabetização (*spelling*) em diferentes unidades escolares. Ainda que esse trabalho de Rice tenha tido uma grande repercussão na época e levado muitos professores a mudarem seus procedimentos de alfabetização, na verdade, o maior impacto desta avaliação estava no seu caráter experimentalista e quantitativo, procedimento até então inédito ainda segundo o destaque de Madaus *et al.* (1993). Rice, de certa forma, antecipou-se a ideias sobre pesquisa experimental em educação que somente seriam defendidas na segunda metade do século XX, a partir dos anos 50, por Lindquist (1953) e Campbell e Stanley (1963).

Os trinta primeiros anos da vida social norte-americana no século XX sofreram a influência de três elementos desenvolvidos em princípio para o gerenciamento industrial: sistematização, padronização e eficiência, que acabaram por afetar a totalidade da sociedade, inclusive na área educacional. A comunidade educacional como um todo passa, então a preocupar-se com o desenvolvimento de uma metodologia que permitisse medir a eficiência dos seus professores, a construir instrumentos e a definir padrões que possibilitassem a mensuração do grau de eficiência das suas escolas e dos diversos sistemas educacionais, seguindo, assim, no campo da educação, os procedimentos que empresários procuravam implantar no mundo da indústria. Aproximadamente quase cem anos depois, a educação brasileira, muitas vezes por influência de agências financiadoras externas, começa a se preocupar com os mesmos problemas ligados diretamente ao processo de gerenciamento das instituições educacionais, conforme poder-se-á observar no item final do presente trabalho.

O grande passo na evolução da avaliação educacional foi dado por E. L. Thorndike, nos princípios do século XX, ao desenvolver toda uma fundamentação teórica sobre a possibilidade de

medir mudanças nos seres humanos (WORTHEN; SANDERS, 1987). A partir desse momento, todo um aparato tecnológico é desenvolvido para a medida de capacidades humanas e a avaliação, em consequência, passa a ter o significado de medida (*testing*), concepção que ainda prevalece em amplos setores educacionais, inclusive fora dos Estados Unidos, como no Brasil.

Madaus, Stufflebeam e Scriven (1993) mostram claramente que essa é a época do *survey*, que emprega diferentes critérios (taxa de evasão, taxa de aprovação) para medir a eficiência da escola e/ou professor. O uso do *survey* leva, naturalmente, ao desenvolvimento de diversos tipos de **testes objetivos** nas várias áreas curriculares. Esses testes, por sua vez, eram **referenciados a objetivos** e deram origem aos **testes referenciados a normas**, sendo, nesse caso, a porcentagem de alunos aprovados nos testes o critério adotado para julgar a eficiência da escola/professor. Além disso, é nesta mesma época, há quase um século, que começam a surgir testes normativos, elaborados por importantes cientistas educacionais, como foi o caso de Edward Thorndike, com vistas ao desenvolvimento de um instrumental que possibilitasse a comparação entre sistemas. Os **testes padronizados**, bastante característicos da cultura educacional norte-americana, mas que, felizmente, ainda não chegaram ao meio educacional brasileiro, surgiram logo após o conflito mundial de 1914/18 e passaram a ser usados, equivocadamente, na determinação da eficiência de programas educacionais e na apresentação de diagnósticos relativos a currículos e sistemas educacionais, além de servirem, também, para a tomada de decisões sobre o desempenho escolar dos alunos.

Avaliação e medida por meio de testes confundem-se nessa época – e ainda continuam a se confundir nos dias atuais; por outro lado, começam a surgir instituições especializadas na realização de *surveys* na área educacional. Estas instituições constituíram o embrião que geraria futuros e importantes centros educacionais, criados nos anos 60 e 70, nas universidades, com o objetivo de realizarem trabalhos avaliativos em diferentes áreas, inclusive na educacional. Os *surveys*, inicialmente, eram locais, limitados aos *school districts*, ainda segundo Madaus *et al.* (1993). No decorrer dos anos 50 e 60 é que começam a surgir efetivamente nos Estados Unidos os primeiros estudos de currículo a nível

nacional. Os problemas passam a ser considerados não apenas em função do interesse local, mas levando em conta o contexto nacional, com uma audiência bem mais ampla do que aquela dos *school districts*. E, assim, por via de consequência, complexos problemas ligados à epistemologia da pesquisa/avaliação educacional começam a ser discutidos por educadores e avaliadores, que, entre outros aspectos, passam a se preocupar com a questão da **generalizabilidade** das suas conclusões, como mostram Madaus *et al.* (1993) e Norris (1993).

A influência de Ralph W. Tyler durante o período de 1930 a 1945 foi considerável e, assim, com justa razão, passou a ser considerado o verdadeiro iniciador da avaliação educacional. A sua ação foi bastante ampla, influenciando na educação em geral, especialmente em assuntos ligados à teoria, à construção e à implementação de currículos, que procurou conceituar como um conjunto de experiências educacionais diversificadas que deveriam ser planejadas de forma a levar os alunos à concretização de determinados objetivos. A avaliação educacional, cujo termo foi por ele criado, objetivaria que professores aprimorassem seus cursos e os instrumentos de medida que construísem pudessem verificar a congruência entre os conteúdos curriculares e as capacidades desenvolvidas.

A sociedade norte-americana depois do término da Segunda Guerra Mundial (1939-45) atravessou um período bastante crítico, envolvendo problemas econômicos e sociais (racismo/segregacionismo). Houve, por outro lado, como decorrência da chamada “guerra fria”, resultante do conflito ideológico-maniqueísta entre o capitalismo ocidental e o socialismo real do leste europeu, a criação de um grande complexo militar-industrial que atuaria intensamente na vida americana por largo tempo até o início dos anos 90, com a crise que levaria à desagregação da União Soviética. Apesar da crise socioeconômica, especialmente no período inicial, entre 1946 e 1957, importantes transformações ocorreram na educação americana.

O problema da avaliação educacional foi bastante discutido, somas consideráveis de dados foram levantadas, mas quase não foram utilizadas para a solução de problemas, como ressaltaram Madaus *et al.* (1993). É nesse período que ocorre o desenvolvimento da teoria clássica dos testes e o surgimento de novas abordagens sobre avaliação educacional, como uma rea-

ção ao modelo proposto por Tyler. É importante destacar, nesse período, o aparecimento de uma instituição atuante até os dias fluentes e que serviu de modelo para a criação de outros órgãos semelhantes nos seus objetivos: – o **Educational Testing Service** – ETS (1947), por inspiração de vários educadores, destacando-se E. F. Lindquist e Ralph W. Tyler. O ETS, a partir de sua criação, passaria a ter influência decisiva no desenvolvimento de testes (padronizados) e em amplos programas de avaliação, como o **National Assessment of Educational Progress** (NAEP) e o **International Assessment of Educational Progress** (IAEP). A proliferação dos testes, nem sempre construídos segundo princípios claros e bem definidos cientificamente, levou algumas instituições representativas (*American Psychological Association* – APA, *American Educational Research Association* – AERA, *National Committee on Measurement in Education* – NCME, sobretudo) a estabelecerem recomendações para a construção de testes psicológicos e técnicas de diagnóstico (1954) e para a elaboração de testes de rendimento escolar (1955), – o que nem sempre é seguido, inclusive no Brasil, que desconhece essas recomendações, salvo as exceções de sempre; mais tarde, após sucessivas revisões, essas mesmas recomendações se transformaram em padrões (*standards*) a serem seguidos na área de construção dos vários instrumentos psicométricos.

A avaliação educacional nos princípios dos anos 60, nos Estados Unidos, começa a incidir sobre grandes projetos de currículos financiados com o apoio federal, surgindo, nesse momento histórico, a figura do avaliador como um profissional com atividades específicas até então exercidas por educadores com formação generalista. É o início de uma época de especialização – que ainda não chegou ao contexto brasileiro – em que as universidades começam a se preocupar com a formação específica de recursos humanos efetivamente qualificados para os trabalhos de avaliação.

Apesar do envolvimento das melhores cabeças pensantes na área educacional e dos grandes fundos alocados a múltiplos projetos, os resultados das avaliações durante a década de 60, ainda que realizados com grande rigor técnico, não se revelaram plenamente satisfatórios e surgiram diversas reações. É nesse contexto que o **Teachers College Record** publica importante artigo de Cronbach (1963) criticando a situação e pro-

pondo novos caminhos para a prática da avaliação, conforme veremos. Apesar da advertência de Cronbach de que a avaliação da eficácia de um currículo se deveria processar ao longo de sua estruturação e não depois de sua conclusão, o impacto das suas ideias não foi imediato, mas levou especialistas a discutirem detidamente seus problemas profissionais, como mostraram Worthen e Sanders (1987).

O trabalho de Cronbach – *Course Improvement through Evaluation* (1963) – levantou importantes questões metodológicas, inclusive a relacionada com a análise direta dos resultados dos itens em substituição à análise concentrada em escores globais, além de abordar outros problemas de natureza conceitual que continuam válidos na atualidade e mostrar as limitações das avaliações *post-hoc* no desenvolvimento de currículos, conforme foi referido anteriormente. Esse artigo mostra os fundamentos básicos da prática da avaliação e abre novas e estimulantes perspectivas para o trabalho do avaliador no seu dia a dia como profissional.

O contexto político decorrente, inicialmente, da administração John Kennedy e, a seguir, após 1963, da gestão Lyndon Johnson, em que houve uma grande preocupação com os reflexos das desigualdades sociais na diferenciação das oportunidades educacionais, passa a determinar um grande esforço no sentido de criar novas condições na área educacional, surgindo, por influência de Robert Kennedy, o conceito de responsabilidade em educação (*accountability*), a fim de evitar possíveis desperdícios dos recursos financeiros concedidos aos programas curriculares e a suas avaliações, na área da educação compensatória. A avaliação deixa, nesse novo contexto, de ser apenas um trabalho teórico de alguns educadores, transformando-se numa prática constante, que, em muitos casos, assume um caráter quase obsessivo na cultura educacional norte-americana.

A aprovação de importante lei sobre educação elementar e secundária – *Elementary and Secondary Education Act (ESEA)*, em 1965, vai gerar um novo quadro que favorecerá a pesquisa e a avaliação educacional, tendo em vista a imposição legal de que todos os projetos financiados deveriam ser obrigatoriamente avaliados (WORTHEN; SANDERS, 1987), o que obrigou os educadores a dedicarem grandes esforços para a avaliação de suas atividades e a elaboração de relatórios passou a ser uma

prática habitual, no controle da qualidade do ensino e dos investimentos feitos em educação.

A situação criada após 1965 mostrou que os educadores norte-americanos não estavam preparados para os grandes desafios da avaliação, como, aliás, ainda não estão suficientemente preparados os educadores brasileiros para o enfrentamento dos grandes problemas da avaliação de hoje (1995), havendo muito teorização abstrata sobre o assunto, mas pouca realização prática. Professores norte-americanos, sob o impacto do ESEA em 1965, improvisaram-se em avaliadores. Worthen e Sanders (1987) mostraram que mesmo aqueles que tinham alguma *expertise* técnica não tinham treinamento suficiente em planejamento, medidas e estatísticas, quadro que, no momento, ocorre no Brasil. Isso produziu avaliações bastante comprometidas, que pouco serviram para avaliar a efetiva eficiência dos sistemas educacionais, seus currículos e programas.

A prática generalizada da avaliação mostrou que muitos instrumentos e estratégias utilizadas em avaliação não eram adequados aos propósitos definidos, ficando, inclusive, caracterizada a inadequação do emprego de testes padronizados na avaliação de programas, conforme a discussão de Madaus *et al.* (1993). Essa insatisfação acabou tendo efeitos positivos. Novas teorias sobre avaliação começam a despontar e todo um trabalho de reformulação da prática da avaliação é empreendido, destacando-se as figuras exponenciais de M. Scriven (1967), R. Stake (1967) e D. Stufflebeam (1971), cujos trabalhos vão dar uma nova dimensão metodológica à avaliação educacional. A avaliação de alguns programas de impacto, como o **Head Start** e o **Sesame Street**, criados com objetivo de eliminar desequilíbrios educacionais decorrentes da origem social e racial, geram, na verdade, acentuadas controvérsias, conforme a colocação de Madaus, Airasian e Kellaghan (1980), sobretudo face ao baixo desempenho das crianças urbanas das áreas de periferia. A avaliação desses e de outros programas sociais teve um grande impacto social, ao apresentar resultados pouco promissores, e serviu para aprofundar as críticas a muitos procedimentos então em uso e possibilitou o desenvolvimento de novas metodologias.

A partir dos anos 70, a avaliação educacional torna-se um campo profissional definido, exigindo especialização aprofundada, com a exclusão de improvisações supostamente técnicas, mas

pouco sólidas conceitualmente, num campo em que atuavam diferentes profissionais, inclusive administradores. A avaliação, pelo menos no contexto norte-americano, deixa de ser uma “terra de ninguém”. Simultaneamente, surgem importantes revistas especializadas, algumas de alta qualidade técnica, associando avaliação e políticas públicas, avaliação e planejamento, entre outras, permitindo, assim, a disseminação de novas ideias, a formulação de teorias e modelos, e sobretudo, a divulgação de importantes estudos.

A literatura sobre avaliação educacional – distinta da bibliografia sobre testes e medidas – torna-se copiosa e algumas universidades, como por exemplo, a de Illinois (Urbana), a de Stanford (Palo Alto, Califórnia) e a *Western Michigan University* (Kalamazoo, MI), entre outras, desenvolvem intensos programas de formação de avaliadores, ressaltando-se, ainda, o trabalho de Ben S. Bloom em Chicago e o de W. James Popham, na *University of Califórnia* (UCLA). Além disso, numerosos centros de avaliação foram disseminados por todo o território norte-americano, conforme o longo elenco de instituições apresentado por Madaus *et al.* (1993).

Apesar dos avanços e de todo aparato técnico desenvolvido, a área da avaliação ainda não está exaurida em todas as suas potencialidades, novas perspectivas de ação estão sendo abertas e a radicalização positivismo/quantitativo em relação ao fenomenológico/qualitativo é menos radical, sendo as divergências porventura ainda subsistentes na verdade um reflexo de diferenças ideológicas e menos conflitos metodológicos, conforme ressaltam Madaus *et al.* (1993).

Finalmente, após a longa exposição sobre a avaliação na perspectiva norte-americana, queremos ressaltar que, hoje, nos Estados Unidos, há uma grande preocupação com a qualidade dos trabalhos de avaliação, tendo sido fixados padrões a serem seguidos conforme o posicionamento do **Joint Committee on Standards for Educational Evaluation**, coordenado por D. Stufflebeam. Ainda que a avaliação educacional, suas teorias e modelos deem margem a grandes controvérsias, é, sem dúvida, um campo dinâmico, em constante transformação para atendimento das múltiplas exigências da qualidade em educação.

## 1.2. AVALIAÇÃO EDUCACIONAL NA INGLATERRA:

### ALGUMAS EXPERIÊNCIAS

A avaliação educacional na Inglaterra surge em meados do século XIX, integrando um programa social maior com vistas à eficiência nacional. A educação, na concepção inglesa, é considerada como elemento indiscutivelmente essencial para alcançar essa eficiência. As pesquisas (e a avaliação) na área educacional, inicialmente, desenvolvem uma linha ligada a métodos estatísticos e à técnica de *survey*, como ocorreu também nos Estados Unidos, contribuindo, assim, para o desenvolvimento de uma tecnologia: – a psicometria, que mais tarde, seria duramente criticada por muitos educadores ingleses.

Os trabalhos realizados por F. Galton, K. Pearson, C. Spearman e C. Burt, entre outros, na Inglaterra, contribuíram para que a psicometria tivesse influência considerável na avaliação educacional, em especial na construção de instrumentos de medidas psicológicas e do rendimento escolar. A utilização da “teoria dos erros” e a aplicação da noção de “distribuição normal” no estudo da variabilidade das diferenças individuais concorreu para gerar reações justificadas ao uso da psicometria da avaliação<sup>3</sup>. Isso contribuiu para que novos caminhos se abrissem para a exploração de diversos aspectos da avaliação educacional. Entretanto, apesar dessas reações adversas à tradição psicométrica, a Inglaterra, a partir da década de 50, realizou importantes trabalhos sobre a alfabetização e a aprendizagem da leitura, contribuindo, desse modo, para o delineamento de novas metodologias experimentais e o desenvolvimento da teoria dos testes (NORRIS, 1993).

A teoria do desenvolvimento econômico, pleno emprego e segurança nacional, inspirada em concepções apresentadas por Keynes, teve amplas repercussões na definição das políticas públicas da educação inglesa. Houve, naquele momento histórico, uma preocupação especial com os problemas relativos à qualidade dos estudantes que, em grande número, procuravam cursos de ciências aplicadas e tecnologias, e, concomitantemente, um profundo interesse na formação de professores de ciências destinados às escolas secundárias. A partir do ideário econômico de Keynes, que em grande parte fundamenta o capitalismo posterior à guerra de 1939-45, inclusive nos Estados Unidos, os educadores ingleses passaram a considerar a relação entre mão de obra tecnicamente qualificada pelo processo educacional e o crescimento

<sup>3</sup> Para um estudo detalhado da influência da Psicometria na evolução das medidas educacionais ver J. Rust e S. Golombok. *Modern Psychometrics - The Science of Psychological Assessment*. London: Routledge, 1992.

econômico/segurança nacional/e desenvolvimento tecnológico. A Inglaterra, por outro lado, constatou a inteira obsolescência dos currículos de suas escolas, o que tinha grande impacto nos programas que visavam à formação de mão de obra qualificada. A mudança desse quadro se impunha, a fim de proporcionar recursos humanos qualificados capazes de garantir o crescimento econômico. À educação caberia essa tarefa, ainda que as transformações no sistema educacional não pudessem ser imediatas, tendo em vista a impossibilidade de o governo inglês influir diretamente em uma estrutura altamente descentralizada.

Havia, desde meados da década de 50, uma consciência de que a Inglaterra não estava desenvolvendo capacidades e formando cientistas da mesma forma como outros países industrializados. A partir de 1962, graças à associação do Governo com a Fundação Nuffield, instituição de caráter privado, iniciaram-se atividades para a reformulação dos currículos e a formação de professores de ciências. A avaliação, no início dos anos 60, estava mais ligada à pesquisa educacional e os responsáveis pelo Projeto Nuffield revelaram uma certa relutância em aceitar a ideia de uma avaliação do Projeto, conforme Norris (1993); porquanto, argumentavam, o interesse maior estava na mudança de atitudes dos estudantes em relação à ciência e não na medida de possíveis ganhos de conhecimento fatural. A qualidade do Projeto, na visão da Fundação Nuffield, era bastante clara e para uma conclusão a respeito do valor desse empreendimento o Projeto dependeria, apenas, de **informações** das escolas que o utilizavam, das **críticas** que seriam oferecidas e do conhecimento dos **problemas** constatados durante sua execução.

Ainda que uma avaliação externa tenha sido afastada, em virtude da impossibilidade de estabelecer comparações entre o Projeto Nuffield e outros tipos de abordagem curricular, uma avaliação estava sendo feita na realidade, via informações, críticas e análise dos problemas, conforme foi apontado. Uma avaliação do tipo que se tornaria bastante utilizado por amplos segmentos educacionais ingleses: – a avaliação qualitativa. Ou seja, uma avaliação não-quantitativa, na tradição psicométrica, mas uma avaliação igualmente válida do ponto de vista científico. Isso ocorreu na parte de ciências do Projeto, na de línguas (francês).

*A National Foudation for Educational Research* (1964) reali-

zou uma avaliação baseada em testes e questionários, a fim de obter dados psicométricos e sociométricos, com estudos comparativos de vários grupos. Apesar da pouca repercussão dos resultados na comunidade educacional – o que ocorre com bastante frequência, inclusive, ou sobretudo, no Brasil –, ficou evidenciada a coexistência do quantitativo e do qualitativo no contexto educacional inglês, tendência que ainda hoje existe, apesar da falsa imagem divulgada por alguns de que haveria uma total dominância da pesquisa e da avaliação qualitativa, simplesmente.

A partir de 1966, ainda segundo o trabalho desenvolvido por Norris (1993), iniciaram-se estudos avaliativos sobre os Projetos Nuffield (ciências e matemática) por intermédio dos **Schools Council**, agência autônoma fundada pelo governo central e governos locais. Algumas colocações do projeto avaliativo são extremamente importantes do ponto de vista metodológico. Assim, a avaliação era vista não apenas como uma medida dos resultados do projeto, mas como parte integrante do próprio projeto de construção do currículo, segundo a perspectiva apresentada por Cronbach (1963). A avaliação, por outro lado, em relação ao currículo não deveria ser vista como uma simples apresentação de meros exercícios para que um produto final fosse avaliado, mas, na verdade, a avaliação consistiria na coleta e no uso de informações que possibilitariam decisões sobre as diferentes fases do desenvolvimento de um programa educacional.

Ao destacarem as relações entre avaliação e tomada de decisões, as diretrizes estabelecidas para os **Schools Council** deixaram perfeitamente claro os momentos em que a avaliação educacional poderia emprestar sua contribuição: – 1) nas decisões sobre o conteúdo dos cursos e relativas aos métodos a empregar para o seu ensino; – 2) nas decisões sobre as necessidades dos alunos; e – 3) nas decisões concernentes ao treinamento de professores.

O projeto do **Schools Council**, ainda segundo Norris (1993), destaca a participação dos avaliadores, muitas vezes apenas professores, e não necessariamente especialistas no campo da avaliação, desde o início dos trabalhos avaliativos e acentua a concepção de uma avaliação para tomada de decisões, aspecto que é expresso na maioria das diferentes abordagens sobre avaliação desenvolvidas na Inglaterra, e que merece especial destaque no modelo de Stufflebeam, nos Estados Unidos. A

avaliação, na visão do *Schools Council*, é responsabilidade de um grupo de avaliadores, que nem sempre são avaliadores profissionais, e não de apenas um único indivíduo.

Avaliação educacional (*evaluation* – avaliação de programa/*assessment* – avaliação do rendimento escolar) é um assunto altamente controvertido na Inglaterra, apresentando tendências várias que quase sempre se opõem à avaliação somativa, apegando-se mais à do tipo individualizada. A avaliação também entra em choque com um dos mitos caros à educação inglesa – a autonomia dos professores. Observa-se, ainda, que várias abordagens têm um conteúdo sociológico, refletindo uma posição aceita por expressivo número de educadores. Os que seguem essa linha mostram a existência de uma correlação entre o socialmente carente e o rendimento escolar, advogando que os responsáveis pela educação devem apresentar uma discriminação positiva a favor daqueles que se acham nessa situação. Essa tese parece-nos de certa forma enviesada, pois, na realidade, a correlação não se limita ao carente, ela simplesmente existe, independentemente da condição social, tornando-se, entretanto, evidente entre os mais desfavorecidos economicamente<sup>4</sup>.

<sup>4</sup> Uma apresentação detalhada da evolução histórica da avaliação na Inglaterra e suas mais recentes experiências de avaliação de programas acha-se em Norris (1993), capítulos 2, 4, 5 e 6.

### 1.3. AVALIAÇÃO EDUCACIONAL NO BRASIL (1960-95)

As atividades de avaliação educacional no Brasil são bastante escassas, ainda que, no momento atual (1995), a temática venha sendo verbalizada pelo Ministério da Educação, que tem apresentado ideias, às vezes bastante discutíveis, sobre a seleção para o acesso ao 3º grau de ensino e relativas à avaliação institucional por intermédio de uma avaliação do rendimento acadêmico ao término dos cursos de graduação.

A avaliação educacional, conforme será discutido, mesmo quando a nível de sistema, no contexto brasileiro, baseia-se, fundamentalmente, no rendimento escolar, ainda que haja uma coleta simultânea de dados socioeconômicos e de variáveis ligadas ao ensino, ao professor e à escola, em alguns casos. As pesquisas e estudos de avaliação apresentados nesse item em discussão não visam a uma revisão da literatura e nem incluem trabalhos acadêmicos, quase sempre relacionados a pré-requisitos da pós-graduação universitária, como, por exemplo, os mencionados por Luckesi (1991);

na verdade, objetivam mostrar alguns projetos que tiveram, ou ainda estão tendo, certo impacto nas escolas, especialmente nas de 1º grau, sem, entretanto, a pretensão de exaurir o assunto, que, sem dúvida, merece ser submetido a uma meta-avaliação.

A evolução da avaliação educacional no Brasil com o objetivo de verificar a eficiência de professores, currículos, programas e sistemas, além de possibilitar a identificação de diferentes tendências, sobretudo quanto ao desempenho educacional, entre outros aspectos, como seria desejável (VIANNA, 1992), ainda está para ser pesquisada e analisada. A partir da década de 60, e ao longo dos anos seguintes, pode-se constatar que alguma coisa importante começou a ser realizada, ainda que de forma algo incipiente, mas revelando um esforço para proceder de acordo com orientação metodológica, especialmente com base em fontes norte-americanas.

A avaliação no contexto educacional brasileiro é quase sempre promovida por órgãos governamentais a nível federal – Ministério da Educação – ou a nível estadual, através das Secretarias de Estado, que, por falta de estrutura, muitas vezes solicitam a colaboração de outras instituições, universidades ou fundações públicas e privadas. Ao contrário do que ocorre nos Estados Unidos, em que as universidades assumem a iniciativa de projetos de avaliação, ainda que com financiamento externo, ou o que se passa na Inglaterra, em que fundações privadas, contratadas para prestação de serviços, realizam pesquisas e avaliação, a nossa situação como será apresentado, é bastante diversa.

A **FUNBEC – Fundação Brasileira para o Ensino de Ciências**, ao iniciar um programa de novos currículos em Física, Matemática, Química, Biologia e Geociências, nos anos 60 e 70, começou, igualmente, uma avaliação de seus programas, contando, para esse fim, com a *expertise* de Hulda Grobman, na área da Biologia. Essa atividade pioneira, entretanto, não teve continuidade em outras instituições, perdendo-se essa rara oportunidade para o desenvolvimento de *know-how* e a formação de capacitações na área de avaliação.

A **Fundação Getúlio Vargas**, igualmente em meados da década de 60, iniciou importante programa de avaliação somativa no Rio de Janeiro, desenvolvendo um instrumento para avaliar a capacitação de crianças ao término do 1º grau na rede oficial, inspirando-se no teste **Iowa Basic Skills**. O projeto obteve a cola-

boração de Anne Anastasi, Frederick Davis e Robert L. Ebel, que contribuíram para a formação de especialistas brasileiros, ministrando cursos de treinamento em 1965. Razões diversas, inclusive uma radical e abrupta alteração curricular, no antigo Estado da Guanabara (Rio de Janeiro), sem uma prévia avaliação da nova proposta de currículo, provocou a descontinuidade do programa e a consequente dispersão dos grupos de trabalho, que passaram a atuar em outras atividades, muitas vezes estranhas à avaliação.

Ao longo dos anos 70 e na década de 80, face ao processo de massificação das instituições de ensino, especialmente no 3º grau, houve certa intensificação de estudos ligados ao acesso ao ensino superior, sobretudo os relacionados a aspectos psicométricos dos instrumentos de medida e à análise de dados socioeconômicos, como pode ser observado em algumas publicações, especialmente em **Educação e Seleção** (1980-89) e a partir de 1990, em **Estudos em Avaliação Educacional**, ainda que esta revista tenha modificado a orientação editorial da anterior, passando a privilegiar problemas ligados à avaliação em geral, sem se deter, prioritariamente, na problemática da seleção para a Universidade. A década de 70 apresentou, igualmente, interesse, ainda que teórico, na área da avaliação de programas, com a tentativa de disseminação do modelo **CIPP** – contexto, input, processo e produto – desenvolvido por Daniel Stufflebeam e Egon Guba. Alguns poucos trabalhos foram realizados nessa linha metodológica, mas, também, não tiveram prosseguimento, sendo um momento transitório, como habitualmente ocorre na área educacional brasileira.

O **Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro – EDURURAL** –, planejado em 1977, na parte referente a estudos e avaliação, esteve a cargo da **Fundação Cearense de Pesquisa**, que se ocupou de aspectos institucionais, e da **Fundação Carlos Chagas**, que centrou suas atividades na avaliação do rendimento escolar (GATTI; VIANNA; DAVIS, 1991). O projeto, com financiamento do Banco Mundial, coletou dados nos anos de 1981, 1983 e 1985, nos Estados do Ceará, Piauí e Pernambuco, por intermédio de provas de Português e Matemática, aplicadas a crianças de 2ª e 4ª séries do Ensino Fundamental, em 603 escolas rurais. Além da qualidade do ensino da escola rural, o projeto **EDURURAL** realizou seis estudos de caso sobre a atuação dessas escolas nos limites das

relações socioeconomicoculturais locais.

O EDURURAL não se limitou a coletar dados sobre o rendimento escolar em Português e Matemática nas 2<sup>as</sup> e 4<sup>as</sup> séries de escolas rurais, que, aliás, mostraram que as crianças daqueles Estados do Nordeste apresentavam uma aprendizagem dos conceitos básicos visivelmente prejudicada, nessas escolas de ensino multisseriado bastante precário (GATTI, 1993). Incluiu, também, no seu projeto, a avaliação de diversas variáveis, como as condições das escolas, perfil dos professores, impacto do treinamento e condições da família, entre outras.

Os seis estudos etnográficos realizados nos três Estados possibilitaram identificar algumas razões explicativas para o baixo rendimento das crianças nas duas áreas curriculares: – rotatividade dos professores, influência política na designação de professores, inconstância na distribuição de livros, material e merenda, que é feita, igualmente, tendo em vista considerações políticas, baixos salários, condições precárias das escolas multisseriadas, infraestrutura curricular deficiente, pouco tempo dedicado ao ensino durante o dia (2.00 a 2.30 horas/dia), frequência irregular dos alunos, doenças das crianças, condições familiares, dificuldade de acesso à escola, ensino baseado na memorização sem significado e passividade induzida do aluno, entre outros fatores, conforme foi caracterizado com realismo esse estado de coisas por Gatti (1994). Um quadro bastante dramático do Nordeste rural, refletido pela situação igualmente trágica de seu ensino, que a avaliação revelou.

Quase no final da década de 80, foi iniciado um amplo programa de avaliação do rendimento de alunos de escolas de 1<sup>o</sup> grau da Rede Pública em todo o País, por iniciativa do **Instituto Nacional de Estudos e Pesquisas Educacionais – INEP**. Inicialmente, em 1987, essa avaliação (VIANNA; GATTI, 1988; VIANNA, 1989a; VIANNA, 1989b; GATTI; VIANNA; DAVIS, 1991; GATTI, 1993) objetivou identificar, na diversidade do quadro educacional brasileiro, pontos curriculares críticos; verificar o desempenho em aspectos cognitivos básicos de alunos de 1<sup>a</sup>, 2<sup>a</sup>, 3<sup>a</sup>, 5<sup>a</sup> e 7<sup>a</sup> séries; e subsidiar os professores para uma recuperação de seus alunos em aspectos básicos do currículo escolar. Essa avaliação, ao final, abrangeu uma amostra de 27.455 alunos de 238 escolas em 69 cidades de vários Estados da Federação existentes à época, inclusive no então Território do Amapá. O projeto teve seus desdobra-

mentos, com novas avaliações, a pedido da Secretaria de Estado da Educação do Paraná, envolvendo, inicialmente, alunos de 2ª e 4ª séries de escolas de sete cidades (VIANNA; GATTI, 1988) e depois, uma amostra de alunos de escolas oficiais de 22 outras cidades (VIANNA, 1991), com os mesmos objetivos dos trabalhos propostos pelo INEP.

As avaliações a nível estadual promovidas pelo **Instituto Nacional de Estudos e Pesquisas Educacionais – INEP** – visavam a fornecer às Secretarias de Estado da Educação um conjunto de informações sobre as deficiências da aprendizagem escolar; por outro lado, os projetos procuravam criar, também, condições para que as próprias Secretarias tivessem uma efetiva participação nos assuntos pertinentes à avaliação do rendimento, assim como se envolvessem, ainda, em projetos outros relacionados a programas, sistemas e materiais didáticos (GATTI, 1994). As conclusões do projeto, que mostraram aspectos críticos do ensino de 1º grau em quase 70 cidades das diversas regiões geográficas do país, não podem ser, entretanto, generalizadas, tendo em vista a natureza da amostra, um segmento estatisticamente pouco representativo do universo escolar a nível de 1º grau, e a falta de equalização (*equating*) dos instrumentos. O interessante, e digno de ser destacado, é que esse projeto, iniciado em 1987, mais tarde, em 1991, realizaria pela primeira vez uma **avaliação de escolas do sistema privado de ensino**, mostrando a relação entre condição social e rendimento escolar. O estudo, entretanto, ao comparar o desempenho da escola privada com o dos alunos das escolas públicas, revelou que nem sempre a escola privada é um mar de excelência e a escola pública também nem sempre é tão ruim quanto se julga, aprioristicamente.

Ainda na década de 80, um novo projeto de avaliação a nível nacional foi realizado por intermédio da Secretaria de Ensino do 2º Grau, do Ministério da Educação, com apoio do Banco Mundial e a colaboração científica da **Fundação Carlos Chagas**, sobre o desempenho escolar de alunos da 3ª série do Ensino Médio (VIANNA, 1991). A partir de uma amostra de 3.972 sujeitos de escolas técnicas federais, escolas estaduais, escolas particulares, escolas oficiais e particulares com habilitação magistério e escolas de formação profissional industrial (SENAI), nas cidades de Fortaleza, Salvador, São Paulo e Curitiba, o projeto identificou variáveis sobre escolaridade e a influência de fatores socioeconômicos.

micos. Esse estudo, que no fundo é de avaliação do rendimento escolar, apresentou resultados surpreendentes, que eram conhecidos mas não constatados, sobre o alto desempenho das escolas técnicas federais em relação às demais escolas, sem, entretanto, pesquisar as razões desse desempenho.

O Ministério da Educação, utilizando a competência técnica do INEP, no início de 1990, iniciou a implantação de um **Sistema Nacional de Avaliação da Educação Básica – SAEB**, com o objetivo de qualificar os resultados obtidos pelo sistema educacional de ensino público, criar e consolidar competências para a avaliação do sistema educacional (PESTANA, 1992), realizando um trabalho cooperativo entre o MEC e as Secretarias de Estado da Educação. A proposta do SAEB adotou um modelo de estudo de fluxo e de produtividade da UNESCO, com vistas a estudar questões relacionadas com a gestão escolar, competência docente, custo-aluno direto e indireto e rendimento escolar, com base em uma metodologia de amostras relacionadas (WAISELFISZ, 1991), estando em vias de sofrer radical transformação em sua metodologia e nos procedimentos de amostragem, na avaliação do segundo semestre de 1995.

A **Avaliação da Jornada Única em São Paulo**, realizada no segundo semestre de 1992, visou a verificar os efeitos da implantação da Jornada Única no Ciclo Básico, a partir de 1988, em escolas da Grande São Paulo. Ou seja, a avaliação procurou saber, conforme o destaque de Gatti (1992), se o Ciclo Básico/Jornada Única causou algum impacto, algum diferencial no desempenho e no desenvolvimento intelectual das crianças; se houve diferenciais no desenvolvimento de habilidades básicas que devem ser adquiridas, e se ocorreram alterações na relação aprovação/reprovação/ e evasão.

A partir de 1992, a Secretaria de Educação do Estado de Minas Gerais iniciou a **Avaliação do Sistema Estadual de Ensino de Minas Gerais**<sup>5</sup>, envolvendo nesse ano a população de estudantes do Ciclo Básico de Alfabetização e os da 8ª série. Os trabalhos prosseguiram em 1993 com a avaliação da 5ª série do Ensino Fundamental, da 2ª série do Ensino Médio e das 3ª/4ª séries da Habilitação Magistério. Ao todo, cerca de 930.000 estudantes foram avaliados externamente nesses dois anos, sendo levantados dados sobre o desempenho escolar em áreas curriculares, informações socioeconômicas dos alunos, atitudes em relação à ciência (2ª série do

<sup>5</sup> O artigo de Maria Alba de Souza, na revista *Estudos em Avaliação Educacional*, São Paulo, n. 12, 1995, oferece uma visão bastante detalhada do programa de avaliação de Minas Gerais no período de 1992-95.

Ensino Médio), dados relacionados à escola, incluindo elementos sobre os vários currículos, e informações relativas à formação das futuras professoras da escola de Ensino Fundamental (alunas das 3<sup>a</sup>/4<sup>a</sup> séries da Habilitação Magistério). O programa de avaliação do sistema de ensino de Minas Gerais, integrante de um programa de qualidade da escola, com apoio financeiro do Banco Mundial, teve prosseguimento em 1994, com a repetição do processo nas escolas avaliadas em 1992 e a inclusão de escolas da rede municipal de ensino, abrangendo, no caso, o sistema em 414 municípios. O processo de avaliação em Minas Gerais já começou a produzir seus primeiros resultados, sobretudo em modificações curriculares e na disseminação de centros de atualização de professores em 53 diferentes pontos do Estado (VIANNA, 1992), além da autonomia administrativa e financeira, estando em vias a implantação da autonomia pedagógica, em um processo de qualidade da escola que tem como um de seus pontos centrais a avaliação.

Um programa de avaliação foi realizado em 1991 sobre o **Desempenho da Rede Pública Escolar do Estado de Pernambuco na Área da Linguagem**, em crianças de 1<sup>a</sup> a 4<sup>a</sup> série, o que possibilitou à Secretaria de Educação do Estado, segundo Buarque *et al.* (1992), a revisão dos conteúdos programáticos das Propostas Curriculares, redirecionamento dos conteúdos nos programas de capacitação de docentes, acompanhamento das classes de alfabetização e o desenvolvimento de um programa de pesquisas.

A nível internacional, no período de 1990/92, houve uma avaliação do desempenho de crianças de 13 anos em Matemática e Ciências, nas cidades de São Paulo e Fortaleza, como parte integrante do projeto da **International Assessment of Educational Progress – IAEP**, com a participação de 20 países, sob a coordenação do **Educational Testing Service** (Princeton, New Jersey) e o financiamento da **National Science Foundation – NSF**. A parte relativa ao Brasil ficou sob a responsabilidade da **Fundação Carlos Chagas**, visando o projeto a analisar o desempenho dos estudantes nos vários países participantes e a identificar tipos de ambientes culturais e práticas educacionais associadas a um alto desempenho (VIANNA, 1992). Além das provas de Matemática e Ciências foram desenvolvidos três outros instrumentos – questionário do aluno, da escola e dos pais –, a fim de explorar variáveis associadas ao desempenho educacional

e proporcionar um contexto que possibilitasse compreender os resultados da avaliação. O desempenho das crianças brasileiras, lamentavelmente, foi bastante comprometido, refletindo, assim, a conhecida crise do nosso sistema de ensino da escola de 1º grau. Esse tipo de avaliação costuma, no entanto, com justa razão, ser criticado por vários educadores, sobretudo na Inglaterra, tendo em vista a diversidade curricular, os objetivos do ensino, as metodologias empregadas, entre outros fatores, e inclusive o próprio contexto cultural, em termos sociológicos.

A análise dos vários projetos de avaliação, segundo demonstrou Gatti (1994), evidencia que nem sempre é possível a adoção de um delineamento experimental, que possibilite comparar diferenças de tratamentos; as amostras, muitas vezes, são bastante comprometidas, tendo em vista a carência de estatísticas básicas confiáveis, os procedimentos de aplicação dos instrumentos e coleta de dados nem sempre são merecedores de confiança, apesar das orientações e dos treinamentos a que são submetidos os aplicadores; e a adoção de modelos de avaliação que nem sempre se ajustam ao nosso contexto socioeducacional, havendo necessidade, portanto, de uma adaptação à realidade brasileira. Esses e outros problemas – como, por exemplo, a falta de elementos experientes em avaliação educacional – fazem com que, em muitos casos, o avaliador responsável tome decisões que, em princípio, fogem à ortodoxia doutrinária, mas que resultam da sua experiência e do seu bom senso, face ao imprevisto das situações que surgem, sobretudo em países que não possuem uma tradição de avaliação no seu sistema educacional, como é o caso do Brasil, e em que a realidade é bem diferente da teoria.

A análise da produção científica sobre avaliação educacional publicada em *Cadernos de Pesquisa*, no período de 20 anos (VIANNA, 1992) mostra que o tema é de interesse dos educadores, sendo essa produção bastante diversificada e revelando uma grande preocupação metodológica na abordagem dos vários assuntos, sobretudo na área de rendimento escolar, educação de adultos, treinamento e formação de professores, temas dominantes no contexto da educação brasileira, pelo menos no momento atual. Observa-se, também, nesses artigos e ensaios, um destaque em relação ao emprego de metodologias qualitativas, especialmente estudo de casos, ainda que sejam bastante reduzidos os trabalhos teóricos sobre essas metodologias por autores

nacionais, que se apoiam quase sempre na fundamentação de teóricos estrangeiros, sobretudo norte-americanos e ingleses. A partir dessa análise (VIANNA, 1992), percebe-se, entretanto, que já há uma consciência da importância da avaliação e que a sua prática, apesar de inspirada em outros contextos, e partindo de modelos adaptados à nossa realidade, mas nem sempre de forma cientificamente rigorosa, está começando a criar uma **cultura da avaliação** na sociedade brasileira.

## REFERÊNCIAS BIBLIOGRÁFICAS

BUARQUE, Lair L. *et al.* Avaliação do desempenho da rede pública escolar do estado de Pernambuco na área de linguagem. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 95-106, 1992.

CAMPBELL, D. T.; STANLEY, J. C. Experimental and quasi-experimental designs for research on teaching. In: GAGE, N. L. (Ed.). *Handbook of research on teaching*. Chicago, IL: Rand McNally, 1963. p. 171-246.

CRONBACH, L. J. Course improvement through evaluation. *Teachers College Record*, n. 64, p. 672-683, 1963.

GATTI, Bernardete A. O rendimento escolar em distintos setores da sociedade. *Estudos em Avaliação Educacional*, São Paulo, n. 7, p. 95-112, 1993.

\_\_\_\_\_. Avaliação da jornada única em São Paulo. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 85-89, 1992.

\_\_\_\_\_. Avaliação educacional no Brasil: - experiências, problemas, recomendações. *Estudos em Avaliação Educacional*, São Paulo, n. 10, p. 67-80, 1994.

GATTI, Bernardete A.; VIANNA, Heraldo M.; DAVIS, Claudia. Problemas e impasses da avaliação de projetos e sistemas educacionais: dois casos brasileiros. *Estudos em Avaliação Educacional*, São Paulo, n. 4, p. 07-26, 1991.

LINDQUIST, E. F. *Design and analysis of experiments in Psychology and Education*. Boston, Houghton Mifflin, 1953.

LUCKESI, Cipriano C. *Avaliação dos resultados da aprendizagem escolar: a prática atual e sua compreensão nas concepções pedagógicas*. 1991. Tese (Doutorado) – Pontifícia Universidade Católica de São Paulo, São Paulo, 1991.

MADAUS, G. F.; AIRASIAN, P. W.; KELLAGHAN, T. *School effectiveness: a reassessment of the evidence*. New York: McGraw-Hill Book, 1980.

MADAUS, G. F.; STUFFLEBEAM, D. L.; SCRIVEN, M. S. Program evaluation: a historical overview. In: MADAUS, G. F. *et al.* (Ed.). *Evaluation Models: viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff, 1993.

NORRIS, N. *Understanding educational evaluation*. London: Kogan Page, 1993.

PESTANA, Maria Inês G. de S. O Sistema Nacional de Avaliação da Educação Básica. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 81-84, 1992.

SCRIVEN, Michael. *The methodology of evaluation: perspectives of curriculum evaluation*. Chicago: Rand McNally, 1967. (AERA. Monograph, n.1).

STAKE, Robert E. The countenance of educational evaluation. *Teachers College Record*, v. 68, n. 7, p. 523-540. 1967.

STUFFLEBEAM, Daniel L. The relevance of the CIPP Evaluation Model for Educational Accountability. *Journal of Research and Development in Education*, v. 5, n. 1, p. 19-25. 1971.

VIANNA, Heraldo M. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: um estudo em 20 cidades. *Educação e Seleção*, São Paulo, n. 19, p. 33-98, 1989a.

\_\_\_\_\_. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: um estudo em 39 cidades. *Educação e Seleção*, São Paulo, n. 20, p. 5-56, 1989b.

\_\_\_\_\_. *Rendimento de alunos das 2ª e 4ª séries das escolas oficiais do estado do Paraná: um estudo avaliativo*. São Paulo: Fundação Carlos Chagas, 1991. Relatório de pesquisa.

\_\_\_\_\_. Avaliação do rendimento escolar de alunos da 3ª série do 2º grau: subsídios para uma discussão. *Estudos em Avaliação Educacional*, São Paulo, n. 3, 1991.

\_\_\_\_\_. Avaliação do desempenho em matemática e ciências: uma experiência em São Paulo e em Fortaleza. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 9-14, 1992a.

\_\_\_\_\_. Avaliando a avaliação: da prática à pesquisa. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 55-64, 1992b.

\_\_\_\_\_. Avaliação do ciclo básico de alfabetização em Minas Gerais. *Estudos em Avaliação Educacional*, São Paulo, n. 5, p. 91-94, 1992c.

\_\_\_\_\_. Avaliação Educacional nos Cadernos de Pesquisa. *Cadernos de Pesquisa*, São Paulo, n. 80, p. 100-105, 1992d.

VIANNA, Heraldo M.; ANTUNES, Ana Lúcia; SOUZA, Maria Alba de. Desenvolvimento de um programa de avaliação do sistema estadual de ensino: o exemplo de Minas Gerais. *Estudos em Avaliação Educacional*, São Paulo, n. 8, p. 5-37, 1993.

VIANNA, H. M.; FRANCO, Gláucia T. Avaliação do rendimento de alunos de escolas do 1º grau da rede privada: pontos críticos e convergência. *Estudos em Avaliação Educacional*, São Paulo, n. 7, p. 113-132, 1993.

VIANNA, Heraldo M.; GATTI, Bernardete A. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: uma aplicação experimental em 10 cidades. *Educação e Seleção*, São Paulo, n. 17, p. 5-52, jan./jun. 1988a.

\_\_\_\_\_. Avaliação do rendimento de alunos de 2ª e 4ª séries de escolas oficiais do estado do Paraná. *Educação e Seleção*, São Paulo, n. 18, p. 5-62, 1988b.

WASELFSZ, Júlio J. O Sistema Nacional de Avaliação Educacional de 1º Grau. *Estudos em Avaliação Educacional*, São Paulo, n. 4, p. 65-72, 1991.

WORTHEN, B. R.; SANDERS, J. R. *Educational Evaluation: alternative approaches and practical guidelines*. New York, London: Longman, 1987.