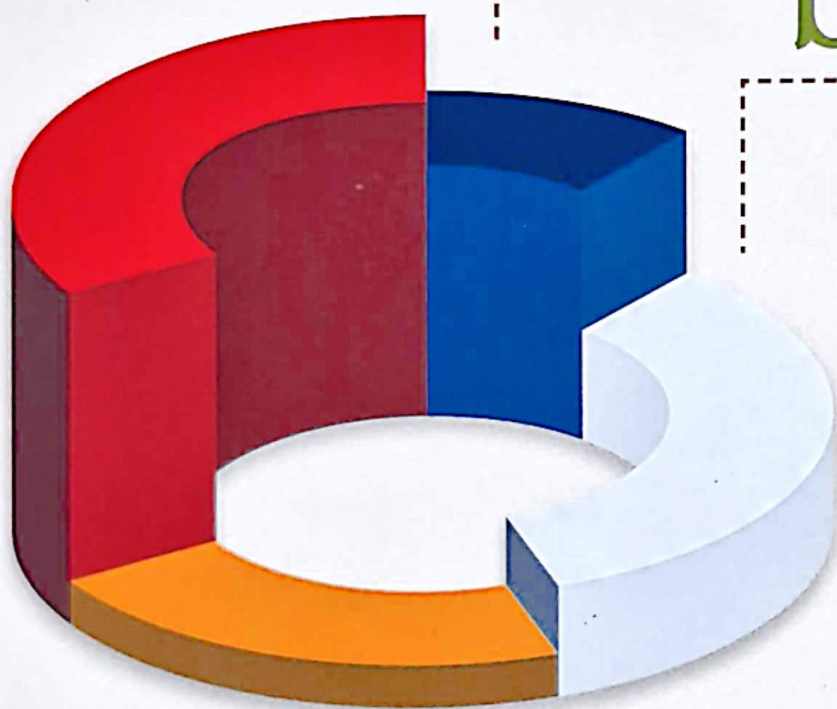


A AVALIAÇÃO DA EDUCAÇÃO BÁSICA:

a
experiência
brasileira



Nigel Brooke

Maria Teresa Gonzaga Alves

Lina Kátia Mesquita de Oliveira

(Orgs.)

FINO TRAÇO

FT
EDITORA

A Avaliação da Educação Básica: a experiência brasileira

ORGANIZAÇÃO

Nigel Brooke

Maria Teresa Gonzaga Alves

Lina Kátia Mesquita de Oliveira

FINO TRAÇO



EDITORA

Todos os direitos reservados à Fino Traço Editora Ltda.

© Nigel Brooke, Maria Teresa Gonzaga Alves, Lina Kátia Mesquita de Oliveira

Este livro ou parte dele não pode ser reproduzido por qualquer meio sem a autorização da editora.

As ideias contidas neste livro são de responsabilidade de seus organizadores e autores e não expressam necessariamente a posição da editora.

CIP-Brasil. Catalogação na Publicação | Sindicato Nacional dos Editores de Livros, RJ

A464

A Avaliação da Educação Básica: a experiência brasileira / organização Nigel Brooke, Maria Teresa Gonzaga Alves, Lina Kátia Mesquita de Oliveira. - 1. ed. - Belo Horizonte, MG: Fino Traço, 2015.

604 p. : il. ; 24 cm. (EDVCERE ; 36)

ISBN 978-85-8054-257-8

1. Avaliação educacional. I. Brooke, Nigel. II. Alves, Maria Teresa Gonzaga. III. Oliveira, Lina Kátia Mesquita de. IV. Série.

15-24957

CDD: 371.26

CDU: 37.091.26

CONSELHO EDITORIAL COLEÇÃO EDVCERE

Alfredo Macedo Gomes | UFPE

Álvaro Luiz Moreira Hypolito | UFPEL

Dagmar Elizabeth Estermann Meyer | UFRS

Dalila Andrade Oliveira | UFMG

Diana Gonçalves Vidal | USP

Elizeu Clementino de Souza | UNEB

Luiz Fernandes Dourado | UFG

Wivian Weller | UNB

FINO TRAÇO EDITORA LTDA.

Av. do Contorno, 9317 A | 2º andar | Barro Preto | CEP 30110-063

Belo Horizonte. MG. Brasil | Telefone: (31) 3212-9444

finotracoeditora.com.br

Leitura 1

Avaliação Educacional: uma perspectiva histórica¹

Heraldo Marelím Vianna

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 – a pesquisa e a avaliação – a influência de diferentes ciências, como, por exemplo, a psicologia, a psicometria, a sociologia, a antropologia, a etnografia e a economia, entre outras, que 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, principalmente 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.²

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

1. VIANNA, Heroldo Marelím. Avaliação Educacional: uma perspectiva histórica. *Estudos em Avaliação Educacional*, nº 12, p. 07-24, 1995. Reproduzido com permissão da editora.

2. Alguns desses novos currículos em Matemática (*Mathematics Study Group* - MSG), em Física (*Physical Science Study Committee* - PSSC), em Química (*Chemical Bound Approach* - CBA) e em Biologia (*Biological Study Committee* - BSC), entre outros, foram traduzidos e adaptados no 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, Myriam K. e Norma Maria Cleffi, com a colaboração de professores brasileiros, no início da década de 60.

perspectiva norte-americana, em cujo contexto se desenvolveram numerosos e importantes trabalhos, em particular como 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 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 e Sander, 1987), mas somente adquire uma natureza formal quando Horace Mann (1845) inicia a prática de 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 o 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 estão 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 dedicados ao 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 caráter experimentalista e quantitativo, procedimento até então inédito 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 e 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. Estes 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 começaram 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 1917/18 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. Os *surveys*, inicialmente, eram locais, limitados aos *school districts*, ainda segundo Madaus et al (1993). No decorrer dos anos 1950 e 1960 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 de 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 construíssem 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 longo tempo até início dos anos 1990, 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 levantados, mas quase não foram utilizados para a solução de problemas, como ressaltaram Madaus et al (1993). É nesse período que ocorre o de-

envolvimento da teoria clássica dos testes e o surgimento de novas abordagens sobre avaliação educacional, como uma reação 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 Tyler. O ETS, a partir da 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 da construção dos vários instrumentos psicométricos.

A avaliação educacional nos princípios dos anos 60, nos Estados Unidos, começa a incidir sobre grandes projetos de currículo 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 propondo novos caminhos para a prática da avaliação, como 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 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 a 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 e 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 estava 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 hoje (1995), havendo muita 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 começaram 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 1970, 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 California (UCLA). Além disso, numerosos centros de avaliação foram disseminados por todo o território norte-americano, conforme o longo elenco de instituições apresentados 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 e 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.

[...]