

Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema

João Luiz Horta Neto

Resumo

Apresenta e discute alguns dos principais conceitos presentes no debate acadêmico sobre a avaliação externa de escolas e sistemas, e sua relação com as políticas públicas. Conceitua a avaliação da educação e analisa seus diversos usos, evidenciando o caráter central que ela passa a ocupar. Expõe a estrutura de um sistema de avaliação externa e, para situar o leitor na realidade brasileira, apresenta um breve histórico do desenvolvimento das avaliações externas no Brasil, com ênfase no Sistema de Avaliação da Educação Básica (Saeb). Baseia-se em uma revisão da bibliografia sobre avaliação educacional externa e da legislação educacional brasileira.

Palavras-chave: avaliação externa; educação; política pública.

Abstract

External assessment of schools and systems: issues on debate

This article presents and discusses some key concepts in the academic debate on the external assessment of schools and educational systems, and their relationship to public policies. Conceptualizes education assessment and analyzes its varied uses, highlighting its crucial role. Exposes the structure of the Brazilian external assessment system and places the reader in the Brazilian reality, presenting a brief history of the development of external assessments in Brazil, with emphasis on the Sistema de Avaliação da Educação Básica (Saeb). This article was based on a review of the educational evaluation literature and on the Brazilian educational legislation.

Keywords: external assessment; education; public policy.

A avaliação externa da educação básica que vem acontecendo no Brasil nos últimos 15 anos, envolvendo as escolas e os sistemas de ensino, tem sido objeto de muito debate entre dirigentes educacionais, professores e pesquisadores. No bojo dessas discussões, estão presentes questões ligadas à qualidade, às medidas em educação, à responsabilização por resultados e à prestação de contas. Essas questões adquirem centralidade quando se discute a própria relevância da avaliação externa para alcançar melhorias no sistema educacional.

Dias Sobrinho (2001) indica que a avaliação aparece como um mecanismo de seleção e distribuição de indivíduos nos lugares sociais e nas hierarquias de poder e prestígio em diferentes localidades do mundo. Os gregos a utilizavam como mecanismo de seleção de indivíduos para o serviço público e os chineses para a seleção de indivíduos para a guarda dos mandarins. A Revolução Francesa amplia o acesso à educação básica e, com isso, a avaliação foi utilizada para distribuir os alunos conforme a sua capacidade individual. Com a Revolução Industrial, a avaliação passou a se desenvolver tecnicamente e foi utilizada principalmente para selecionar os indivíduos para o serviço público e para outros postos de trabalho. As escolas passam a utilizar o teste escrito e o sistema de notação para dar maior credibilidade pública, transparência e rigor no julgamento dos resultados, com isso eles acabam interferindo nos currículos e nas propostas pedagógicas.

A avaliação tem estado presente no dia a dia das escolas, dos professores e dos alunos há muito tempo. Os professores avaliam seus

alunos para saber sobre suas dificuldades e o que foi aprendido, de forma a auxiliá-los no seu processo de aprendizagem. O acúmulo de informações, obtidas pelo docente nas mais diferentes formas de avaliação ao longo do ano, ajudam-no a decidir sobre a progressão dos alunos. Os professores também passam por processos de avaliação. No caso brasileiro, por determinação constitucional, para ingressar na carreira de docente da rede pública de ensino, os professores precisam se submeter a uma avaliação realizada a partir de um concurso de provas e títulos. Em alguns países, os professores também são avaliados pelos sistemas que emitem uma certificação profissional para que possam atuar na função docente.

Nos últimos anos, muitos países vêm investindo significativos recursos na realização de avaliações externas constituídas, entre outros instrumentos, por provas aplicadas aos alunos com o objetivo de verificar se eles estão adquirindo os conhecimentos e as competências necessárias para a conclusão de um nível ou ciclo do sistema educativo. Poderíamos identificar dois pressupostos que orientariam a realização dessas avaliações. Primeiramente, os ministérios de educação desses países estariam indicando de forma objetiva os referenciais de qualidade que deveriam ser seguidos pelas escolas. O outro seria uma preocupação com as provas aplicadas pelos professores, que não permitiriam avaliar com objetividade os conhecimentos e as competências adquiridos pelos alunos. Qualquer que seja a premissa que oriente essas avaliações, é inevitável identificar nessas concepções alguma intenção de desenvolver mecanismos de controle sobre os processos educacionais.

Este preâmbulo é importante para poder iluminar a compreensão sobre o termo avaliação, demonstrando a complexidade de algumas das relações em que está envolvido.

Este conceito é em si mesmo polissêmico. Para uns, a avaliação resume-se a medir resultados educacionais geralmente por meio de provas. Para outros, a simples medição reduz o significado da avaliação. Avaliar significaria ir além das medições ou apresentação de resultados e envolveria a definição de políticas e estratégias governamentais que levariam ao aperfeiçoamento institucional e do próprio processo de ensino-aprendizagem em seus diferentes graus e modalidades.

Neste ponto, é necessário diferenciar o que é uma prova aplicada por um professor aos seus alunos do que é um sistema de avaliação. Uma prova aplicada pelo professor é um evento único e independente, com ênfase somente em um conjunto de objetivos cognitivos que incide sobre o futuro imediato do aluno e que é interpretado independentemente do seu contexto socioeconômico ou educacional. Ele também pode lançar mão de diferentes instrumentos, os quais permitam identificar outras características que indiquem como está ocorrendo a aprendizagem do aluno. Já um sistema de avaliação – que tem como principais usuários, mas não únicos e exclusivos, os sistemas educacionais – é um processo que envolve um levantamento periódico de dados com o objetivo de captar a evolução do quadro educacional (Tomassia, 2004). Para isso, utilizam-se tanto testes de rendimento, com o objetivo de verificar o desempenho

dos alunos, quanto questionários que procuram levantar os fatores que influem no desempenho desses discentes. É importante destacar que esses fatores levantados pelos questionários não estabelecem uma relação de causalidade entre o fator levantado e o desempenho do aluno, mas podem indicar quais as áreas que deveriam ser prioritárias para a intervenção dos governos e quais as iniciativas políticas que deveriam ser tomadas, com o objetivo de promover melhorias no ensino.

O aumento do interesse pela avaliação do desempenho dos alunos coincide com o momento em que houve um forte acréscimo da demanda por educação nos países centrais, levando a questionamentos com relação a se o conteúdo aprendido na escola seria aquele necessário para a vida em sociedade. Em uma entrevista,¹ o americano Ralph Tyler (1981), considerado por muitos como o iniciador da avaliação educacional externa, destaca que com a crise de 1930 nos Estados Unidos houve um aumento muito grande de jovens que procuravam o ensino secundário, motivados principalmente pela falta de emprego. Com essa forte demanda, no Estado americano de Ohio, surgiu a preocupação com o desempenho desses novos alunos que anteriormente à crise não estavam cursando esse nível de ensino. Foram então selecionadas 30 escolas e dado para elas um período de oito anos para que desenvolvessem e aplicassem novos métodos de ensino, com a contrapartida de que seriam avaliadas durante esse período. Este programa teve início em 1934 e foi chamado de Eight Year Study. Os alunos dessas escolas foram avaliados a partir de uma variedade de instrumentos desenvolvidos por Tyler. Mesmo depois de concluída a escola secundária, os egressos ainda eram acompanhados com o objetivo de identificar de que forma os conhecimentos adquiridos na escola estavam sendo utilizados tanto no seu trabalho, se ele interrompia seus estudos, quanto na universidade, se ele continuava seus estudos.

Na proposta de Tyler, a avaliação da aprendizagem estava integrada ao modelo de elaboração de currículo, e este currículo deveria assumir características de planejamento das atividades educacionais. Essa abordagem recebeu inúmeras críticas, pois considerava-se que a avaliação se tornava sinônimo de simples medida. Nesse sentido, tinha uma característica inflexível e limitada, pois levava em consideração apenas alguns dos conhecimentos que eram escolhidos para serem avaliados nas provas. Além disso, as avaliações estariam condicionando os currículos e fazendo com que a escola e principalmente os professores perdessem sua autonomia para definir quais os conteúdos e as abordagens que seriam mais adequados aos seus alunos, dentro do contexto social em que viviam.

Muitos autores referem-se a esse tipo de avaliação desenvolvida por Tyler como meritocrática, somativa, tecnicista ou positivista. Barreto (2001, p. 55) identifica seus defensores como sendo pertencentes ao “núcleo duro da avaliação da aprendizagem”. Eles partiriam de um paradigma que a autora chama de positivista, pois teriam a intenção de analisar a complexa realidade social utilizando-se de ferramentas estatísticas.

As avaliações somativas estariam considerando apenas os resultados da aprendizagem e não os processos que estariam sendo construídos

¹ Entrevista concedida a Jeri Ridings Nowakowski, da Universidade de Michigan.

durante ela, portanto, não teriam nenhum significado importante, já que a educação seria um fenômeno essencialmente qualitativo. Além disso, não respeitariam a diversidade de aprendizagens que se apresenta em diferentes contextos, nem mesmo a diversidade dos alunos. Assim, elas sinalizariam que todos os discentes deveriam aprender o mesmo e que todos os professores deveriam ensinar o mesmo. Esse tipo de avaliação estaria, portanto, fortalecendo o poder nas mãos do Estado e, com isso, limitando a atuação dos agentes locais.

Como oposição a esse modelo, Barreto (2001) indica que começaram a ser construídas novas abordagens em que eram levadas em conta não apenas a dimensão cognitiva do aluno, mas as dimensões social e afetiva, seus valores, motivações e até mesmo a sua própria história de vida. Nessas novas abordagens, conhecidas como avaliações formativas ou diagnósticas, a avaliação deixaria de girar exclusivamente em torno do aluno e da preocupação de medir tecnicamente seu rendimento. Ela cita alguns autores que desenvolveram pesquisas dentro dessa abordagem, como: Luckesi, que centra seu trabalho no que a autora chama de pedagogia do exame; Franco, que estabelece como unidade de análise o vínculo entre o indivíduo e a sociedade numa dimensão histórica; e Saul, que desenvolve em seu trabalho as bases de uma avaliação chamada de emancipatória.²

Lüdke (1995, *apud* Barreto, 2001) defende que a avaliação esteja fortemente integrada ao processo de ensino-aprendizagem, de forma a atender as necessidades dos alunos e as dos professores, com uma característica de auxiliar no desenvolvimento desses atores. A autora admite que uma avaliação desse tipo necessitaria de um longo período para a sua execução e destaca que, ainda assim, não seria eliminada a necessidade das avaliações baseadas em testagens, que possibilitariam a identificação de novos problemas não detectados antes.

Existem, portanto, diversos tipos de avaliações que são realizadas dentro do sistema educacional: aquelas que acontecem em sala de aula, passando por outras que ocorrem na própria escola, ou até as que perpassam todo o sistema educacional. Com relação às avaliações externas, existem aquelas que têm consequências diretas importantes sobre indivíduos e instituições e que apresentam resultados numéricos (somativas) e aquelas que têm como propósito aprender mais sobre o processo educacional com o objetivo de procurar melhorias, e que não têm o interesse em dar consequência imediata ao seu resultado (formativas).

Considero que todos os tipos de avaliação são importantes e que tratam de responder a diferentes necessidades dos sistemas educativos e da vida em sociedade; elas fazem parte do dia a dia do processo educativo. Com relação a avaliações externas, não se pode desqualificar a avaliação somativa com o argumento de que elas levam em conta os resultados, mas não os processos, pois a elaboração de uma avaliação envolve um minucioso trabalho de análise das condições em que a aprendizagem ocorre. Da mesma forma, também não se pode desqualificar a avaliação formativa sob o argumento de que ela, para ser aplicada em larga escala, demandaria

² Para aprofundar esses assuntos, sugere-se a leitura de Luckesi (1991, 1992, 1996), Franco (1990) e Saul (1992).

mais tempo e recursos, envolveria um processo mais demorado e estaria voltada à realidade das instituições avaliadas, e não ao conjunto delas. Cada processo avaliativo tem um objetivo a atingir e deve ser analisado à luz desses objetivos. Além disso, da mesma forma que acontece na sala de aula, são diversos os instrumentos que podem ser utilizados para avaliar um determinado processo educacional. Mas uma coisa é certa: qualquer que seja a abordagem utilizada, não existe sentido dar às avaliações um caráter punitivo. A avaliação em qualquer de suas acepções deve ter um forte caráter formativo, levando em consideração que seu objetivo é compreender melhor a realidade e, dessa forma, atuar sobre ela, dentro dos limites impostos pelos instrumentos utilizados. É importante destacar que o limite de uma avaliação é definido previamente quando se determinam quais os instrumentos que serão usados no processo.

Vianna (2003) chama a atenção para o fato de que a avaliação deva ter uma consequência, que não se limite a apenas informar os resultados. Atitudes devem ser tomadas no sentido de buscar superar os problemas apontados pelos resultados obtidos.

Outros autores reforçam essa posição, destacando que “é insuficiente a avaliação que apenas localiza o aluno na escola. A avaliação tem de dizer o que fazer” (Belloni, Belloni, 2003, p.16).

Além disso, é preciso estar alerta para o fato de que a avaliação é um processo complexo não somente pelas suas características técnicas, mas, sobretudo, pelo impacto dos questionamentos que o processo avaliativo deve provocar (Dias Sobrinho, 2000). Assim, é preciso ter em conta que outros atributos devem ser acrescentados à avaliação, pois ela “precisa ser um processo de construção, e não uma mera mediação de padrões estabelecidos por iluminados” (Ristoff, 1995, p. 47). Este mesmo autor salienta que a prática educacional no Brasil coloca ênfase na exclusão e não na construção de aprendizagens.

O fato de medir uma determinada realidade não significa que a estamos avaliando. A medida é a etapa inicial que enceta uma avaliação. Assim, se o objetivo é conhecer determinada realidade, ela deve ser identificada a partir da realização de um conjunto de medidas que indicam, trazem informações sobre a realidade que se procura conhecer. Essas medidas sozinhas, obtidas a partir de observação empírica, têm pouco significado. Para dar significado às medidas, é preciso que elas sejam comparadas a uma situação anterior que já tenha sido objeto de medição ou então que elas sejam comparadas com uma situação ideal, um padrão definido por um determinado conjunto de atores, com base em critérios para atingir determinados objetivos. Da comparação entre o resultado obtido da medição e a situação ideal ou os resultados das medidas anteriores, é possível fazer um julgamento, emitir um juízo de valor, sobre a situação atual. Ao emitir o juízo de valor estaremos avaliando essa realidade.

Avaliar como crítico um sistema educacional que tem um desempenho abaixo de um patamar julgado ideal, *per se* não fará com que ele melhore seu desempenho. É preciso identificar as causas desse baixo desempenho e atuar para superá-las.

A instituição de um sistema de avaliação não pode ser fruto de uma improvisação. Vianna (2001) destaca que para a instituição de um sistema de avaliação em larga escala é necessária uma grande capacidade de processamento de informação, além de uma competente equipe multidisciplinar para gerenciar tanto a sua aplicação como seus resultados e, não menos importante, significativos recursos financeiros. Alguns autores, como é o caso de Barreto (2001), alegam que, diante da escassez de recursos, o alto custo desses sistemas tem consumido os recursos que poderiam ser utilizados para a realização de estudos e pesquisas menos custosos, que também teriam potencial para melhorar o sistema escolar. Além disso, essa mesma escassez, que impede até mesmo as medidas educacionais mais básicas de serem utilizadas, impossibilita que os sistemas invistam outros recursos financeiros para solucionar os problemas apontados pelos resultados da avaliação. Assim, os problemas só estariam sendo apontados e nada de efetivo estaria sendo feito. Portanto, a opção pelo desenvolvimento de um sistema de avaliação pressupõe o efetivo enfrentamento dos problemas que vierem a ser apontados, mobilizando para tanto os instrumentos e os recursos necessários.

Num extenso e detalhado trabalho, Ravela (2005) fornece uma série de reflexões sobre os aspectos envolvidos na construção de um sistema de avaliação. O autor indica que é preciso definir os modelos conceituais que serão utilizados, quais as questões que se quer responder e quais os usos e os usuários da avaliação. Cumpridas essas etapas, é preciso produzir a evidência empírica do que se quer medir desenvolvendo instrumentos que auxiliem no levantamento dos dados, levando-se em consideração o tempo necessário para produzi-los, os recursos que se dispõe e as consequências que terá a avaliação. Assim, muito pouca influência terá uma avaliação realizada por um determinado governo que produza resultados três anos após ter sido iniciada e que tenha consumido uma parcela significativa dos recursos orçamentários disponíveis.

Continuando na análise da construção de sistemas de avaliação, Ravela (2005) destaca que muitas vezes se constroem instrumentos que capturam informações que não se adequam àquelas definidas inicialmente, criando-se assim problemas que podem mascarar os dados obtidos, influenciando as análises que se farão. Vianna (2001) também aponta a tendência de querer que uma única avaliação teste o máximo de atributos, utilizando-se para isso um número limitado de itens e de instrumentos. Este autor ainda destaca que qualquer que seja o instrumento que venha a ser utilizado, não deve ser relegada a segundo plano a necessidade de realizar uma pré-testagem dos instrumentos para adequá-los aos sujeitos integrantes do conjunto avaliado.

Como o essencial em todo o processo não são os dados em si, Ravela (2005) aponta a importância da valoração do que está sendo avaliado. Dessa forma, muito mais importante do que saber quantas perguntas determinado aluno respondeu em uma prova de determinada área, é indicar se nessa área ele teve, ou não, um resultado considerado adequado, comparando-se o desempenho esperado – baseado nos referenciais utilizados – e o desempenho alcançado.

Os juízos de valor podem ser expressos basicamente de três formas. Quando a comparação é realizada a partir de referências definidas previamente, diz-se que o juízo de valor é referido a critério. Assim, podem ser estabelecidos distintos níveis de desempenho, indicando níveis de aprendizagem que se espera para cada aluno, possibilitando construir uma escala que contenha diferentes níveis de aprendizagem de uma determinada área, cada um deles apontando o que se espera do aluno em diferentes fases da sua vida escolar.

Quando a comparação é entre indivíduos ou instituições que fazem parte de um mesmo grupo, o juízo de valor está baseado na comparação entre as unidades de análise utilizadas (no caso indivíduos ou instituições), e diz-se que o juízo de valor é normativo ou que está referido à norma, numa referência à curva normal da estatística. Dessa forma, quando a referência é a norma, o juízo de valor produzido é com relação à classificação do indivíduo em relação aos demais.

A terceira forma de expressar um juízo de valor é conhecida como a do crescimento da aprendizagem, em que se procura analisar qual a variação do indivíduo em relação a um ponto de partida ou linha de base anterior. Essa forma pode ser caracterizada como uma avaliação criterial, se antecipadamente forem definidas as referências que serão utilizadas, ou pode ter características normativas, se a preocupação é com a posição relativa entre os indivíduos. De qualquer forma, o objetivo aqui é comparar o desempenho do próprio indivíduo em diversas fases, produzindo juízos de valor que evidenciem as variações sofridas entre os diversos períodos da avaliação.

No caso brasileiro, cada uma das avaliações de desempenho dos alunos desenvolvida satisfaz a uma forma de expressar o juízo de valor. O Sistema de Avaliação da Educação Básica (Saeb) é um misto de avaliação normativa, já que primeiro compara os indivíduos entre si, e avaliação criterial, já que procura descrever para cada ponto da escala de proficiência as habilidades demonstradas pelos alunos a partir de suas respostas aos itens da prova.

Em estudo anterior, Horta Neto (2005) aborda os desafios da avaliação institucional no contexto do Sistema Nacional de Avaliação da Educação Superior (Sinaes), indicando como as modalidades de provas utilizadas no Brasil, a partir de 1996, fazem parte de avaliações que ora são referenciadas à norma, ora procuram medir o crescimento da aprendizagem. O Exame Nacional de Desempenho de Estudantes (Enade), uma das avaliações que fazem parte do Sinaes, é um exemplo de avaliação de crescimento da aprendizagem, visto que são avaliados os alunos do primeiro e do último ano da graduação. Já o Exame Nacional de Cursos (ENC), conhecido também como Provão, utilizado como um dos instrumentos que faziam parte do antigo sistema usado para avaliar a educação superior antes da instituição do Sinaes, pode ser classificado como uma avaliação referenciada à norma. Isso porque os conceitos atribuídos a cada curso tinham como referência a nota média obtida, que recebia um conceito equivalente a "C". Com isso, aqueles cursos que obtivessem notas inferiores a um

ou a meio desvio padrão com relação à média eram classificados com os conceitos "E" ou "D", respectivamente, e aqueles que obtivessem notas superiores a um ou a meio desvio padrão com relação à média seriam classificados com os conceitos "A" ou "B", respectivamente.

A definição de qual a forma mais adequada para expressar um juízo de valor irá depender dos objetivos que se quer alcançar com a avaliação. Por isso, não é possível dizer que uma forma seja superior à outra, e sim que uma seja mais adequada aos objetivos propostos que a outra.

As considerações sobre a forma de expressar o juízo de valor reforçam a necessidade de se estar atento às consequências dadas aos resultados da avaliação, para que ela não seja usada para objetivos que desvirtuem completamente o seu sentido. Vianna (2001) chama a atenção para exemplos como o que ocorreu na Inglaterra durante a era Thatcher, quando os resultados obtidos nas avaliações eram utilizados como justificativas para fechar instituições, atribuir prêmios ou bônus e fazer *ranking* das instituições, ao invés de serem usados como uma forma de conhecer a realidade e atuar sobre ela.

Outro ponto que é importante destacar refere-se à forma como devem ser divulgados os resultados das avaliações. Os responsáveis pela avaliação devem cuidar para que os resultados sejam apresentados para as autoridades/decisores com clareza, para que permitam oferecer uma visão ampla dos problemas levantados. É importante salientar que os resultados apurados não são suficientes, *per se*, para a tomada de decisões e que é necessário aprofundar os aspectos mais problemáticos apontados pela avaliação.

Deve se destacar, também, a referência que é feita à suposta imparcialidade do processo avaliativo, pelo fato de ele utilizar técnicas objetivas. Uma avaliação sofre influência dos valores, das visões de mundo e da percepção que os avaliadores possuem da realidade e das interferências dos governos. Além disso, é importante levar em conta que os dados e as percepções sobre a realidade não conseguem captá-la em sua totalidade. Eles são aproximações dessa realidade.

Tendo em conta os aspectos levantados até aqui, é importante salientar que, mesmo tendo seus limites, a avaliação é capaz de fornecer informações para a tomada de decisões para a implantação de políticas públicas e, nesse sentido, é muito melhor do que um voo cego, no qual não existem dados em que os formuladores de políticas possam se embasar.

Avaliação externa e qualidade educacional: problematizando o tema

Um dos propósitos da avaliação é que ela possa contribuir para a melhoria da qualidade da educação. Para isso é importante discorrer sobre o significado desse termo.

Conceituar qualidade não é uma tarefa simples nem existem consensos sobre o melhor caminho a seguir. Como é um conceito

polissêmico, primeiramente é preciso identificar qual é a definição de qualidade a que estamos nos referindo.

O termo qualidade tem assumido diversos significados ao longo do tempo, e, segundo Reeves e Bednar (1994, *apud* Turchi, 1997, p. 9), eles remontam à Grécia antiga, onde Aristóteles, Sócrates e Platão referiam-se ao termo como sendo um “padrão de excelência moral que deveria ser buscado pela sociedade grega. O critério para definir esse padrão de excelência variava de acordo com as circunstâncias em análise e era baseado na percepção de quem avaliava”. Portanto, apesar de ser referenciado a um padrão, este era totalmente subjetivo, dependia exclusivamente do avaliador.

Com a evolução da sociedade estruturada dentro do modo de produção capitalista, naturalmente o conceito foi gradualmente sendo associado à produção de bens. Por exemplo, no período anterior à Revolução Industrial, “qualidade de um produto era definida pela reputação da habilidade e do talento do artesão que o produzia” (Turchi, 1997, p. 9). Aos poucos, além da excelência, outras dimensões foram sendo incorporadas, buscando-se uma quantificação. Assim, na época da Revolução Industrial, um produto de qualidade estava associado à ideia de menor preço. Nos anos 1930 do século 20, “qualidade incorpora o significado de produção em conformidade com o projeto” (p.10). Com isso, ganharam importância as técnicas de controle dos processos produtivos com base no uso intensivo de técnicas estatísticas. A partir dos anos 1950, outra dimensão foi associada ao conceito: a de que um produto, além de estar em conformidade com o projeto, deveria atender às necessidades dos usuários. Incorporou-se, assim, a ideia de que a qualidade deveria ser muito mais que a simples inspeção do produto, mas que deve ser construída ainda durante a fase de elaboração do projeto. Qualidade então passa a ser entendida como aquela adequada ao uso. Essa teria sido a base dos conceitos que fundamentariam a proposta de qualidade total que se afirmaria no final do século 20.

Tarso Genro (2005, p. 12), ministro da Educação, em uma entrevista concedida à revista *Nova Escola*, referiu-se à qualidade da seguinte forma:

- Nova Escola: O que é um ensino de qualidade?
- Tarso Genro: É uma educação com *boa estrutura escolar e métodos pedagógicos adequados* à nossa sociedade atual para todos os alunos da rede pública. Envolve também *professores bem remunerados*. Para conseguir isso a educação precisa ser considerada uma área estratégica, o que significa *receber investimentos*. (grifos do autor).

No excerto acima, o autor associou a qualidade do ensino a uma boa estrutura escolar e a métodos pedagógicos adequados, direcionados a um determinado público (alunos da rede pública). Estabeleceu também uma relação entre qualidade do ensino e os investimentos que a área deve receber. Como se percebe, todas as questões levantadas pelo ex-ministro envolvem decisões políticas relativas ao financiamento da educação, à revisão curricular e metodológica e à gestão escolar.

Ainda assim, apesar de uma grande amplitude dada ao tema qualidade, ela envolve ainda outros aspectos até aqui não abordados.

De forma mais ampla, autores, como Dias Sobrinho, definem qualidade como sendo uma construção social que varia segundo os interesses dos grupos de dentro e de fora da instituição educativa e que reflete as características da sociedade que se deseja para hoje e que se projeta para o futuro. Não é um conceito unívoco e fixo, devendo ser construído por meio de consensos e negociações entre os atores (Dias Sobrinho, 1995). Na mesma linha, Schwartzman (1990) indica que o termo qualidade depende dos pontos de vista dos atores envolvidos, sendo que a qualidade de uma instituição de educação superior, na visão dos alunos e de suas famílias, vincula-se ao prestígio da futura profissão ou à garantia de trabalho bem remunerado; os professores universitários relacionam a qualidade às possibilidades da investigação e às condições para a realização das atividades científicas; as empresas referem-se à competência esperada dos egressos da universidade. Radicalizando na definição de qualidade, Pirsig (1984) afirma que ela não deve ser definida e que devemos apenas compreendê-la, já que é fruto de uma experiência direta que é independente e anterior a qualquer abstração intelectual. Gentili (1995), por um lado, lembra que na sociedade capitalista a qualidade está associada à excelência e esta a privilégio, nunca a um direito. Por outro lado, Mendonça (2005) destaca que qualidade da educação deve ser entendida como um direito, um bem público, e aponta alguns indicadores que poderiam nortear o que seria a qualidade social.

Entendo que para a definição de políticas públicas é importante que a qualidade possa ter características tais que permitam que ela possa ser mensurada. Caso contrário, haveria um total descompasso entre os formuladores, os executores e os beneficiários dessa política, já que não haveria uma referência comum que permitisse tanto a sua execução como a sua avaliação.

Nessa perspectiva, entendo qualidade como sendo um conceito relativo que utiliza padrões de referência para medir algo, e a partir dos resultados obtidos, compará-los a esses padrões. Uma vez definidos os padrões, levando-se em conta a sua complexidade, o tempo gasto e os recursos necessários para executar a medição, eles não devem sofrer mudanças significativas que impossibilitem a construção de uma série histórica que avalie as variações acontecidas ao longo do tempo.

Dessa forma, qualidade em educação depende da existência de um padrão, uma referência a partir da qual seja possível comparar os processos educacionais que se quer medir a esses padrões. Por meio dos dados obtidos pelas medidas, é possível adquirir indicações objetivas que permitam orientar as políticas educacionais.

Como não é possível medir diretamente a qualidade, utilizam-se alguns substitutos, como os indicadores. Segundo Carley (1985, p. 2), indicadores são "medidas de uma característica observável de um fenômeno social e que estabelecem o valor de uma característica diferente,

mas não observável”. Isto é, o autor indica que quando não é possível medir diretamente um fenômeno social, deve-se utilizar, para realizar a medida, algumas de suas características que sejam observáveis e, portanto, mensuráveis. Apesar de não dar conta de explicitar completamente o fenômeno que pretende mensurar, o indicador fornece pistas, características observáveis e que são importantes para compreender o fenômeno que se quer estudar.

Medir a qualidade da educação a partir de indicadores é um tema que levanta polêmica. Paro (1998) considera que a educação é um processo que se prolonga por toda a vida da pessoa e que, portanto, não pode ter sua qualidade avaliada em um dado momento a partir de índices de reprovação e aprovação ou por avaliações externas. Para o autor, o produto da escola é o aluno educado, e esse produto deve ter especificações bastante rigorosas quanto à qualidade que dele se deve exigir.

Concordando com essa afirmação de Paro, argumento sobre a necessidade de que todos aqueles que estejam envolvidos no processo educacional, como dirigentes, professores, pais e alunos, tenham a sua disposição dados a partir dos quais seja possível estabelecer juízos de valor sobre a qualidade da educação. Sem a existência desses dados, ficaria prejudicado um acordo mínimo sobre como atuar para aprimorar esse processo, dificultando assim a discussão que venha a contribuir para a formulação de políticas públicas adequadas.

Os indicadores usados mundialmente estão relacionados com a qualidade do processo educacional – que medem características como acesso e fluxo escolar, entre outras – e com a qualidade do que foi aprendido pelos alunos, obtidos a partir de uma avaliação externa que meça o desempenho dos alunos em um teste padrão, conseguindo assim maiores informações sobre os processos que ocorrem no interior da escola. No caso da educação básica brasileira, esses indicadores são adquiridos por meio do Saeb.

No Brasil, para obter outros indicadores relacionados à qualidade do processo educacional, utiliza-se também do Censo da Educação Básica, que levanta os dados de todas as escolas. A partir dos dados do Censo são construídos 39 indicadores educacionais divididos em: oferta, acesso e participação e eficiência e rendimento (Brasil, 2004). Para alcançar esses dados, primeiramente, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) encaminha um questionário a todas as escolas para que sejam preenchidos. Para se ter uma ideia da complexidade desse questionário, o formulário do Censo 2005 foi composto por 158 itens distribuídos em 60 páginas. As escolas fornecem os dados na sua forma bruta, sem nenhum tratamento, fornecendo informações obtidas a partir de seus registros administrativos. Em cada Estado e no Distrito Federal (DF), a entrega dos formulários preenchidos pelas escolas é centralizada na respectiva secretaria de educação do Estado ou do DF. De posse dos formulários, as secretarias processam as informações e encaminham os dados para o Inep, que faz a consolidação, o tratamento das informações, a checagem e a publicação anual da

Sinopse Estatística com os resultados obtidos. Até a publicação final dos resultados, todo o processo dura cerca de um ano. Está em fase inicial de testes o Censo em tempo real, que pretende coletar as informações da escola em tempo real.

Como os dados do Censo são adquiridos a partir de informações preenchidas no formulário enviado às escolas, dando a ele um caráter declaratório, foram observadas que existem algumas discrepâncias. Elas se manifestam principalmente quanto ao número de alunos matriculados, quando comparados com outros levantamentos feitos pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Isso foi notado quando, a partir dos anos 1980, foram feitas comparações entre os dados do Censo e aqueles obtidos pela Pesquisa Nacional por Amostra de Domicílios (Pnad), desenvolvida pelo IBGE. A Pnad é uma pesquisa amostral realizada a cada dois anos em todo o Brasil, excluindo-se a área rural da região Norte, e é realizada por meio de pesquisadores que, munidos de um questionário, entrevistam diretamente as famílias. O questionário engloba uma série de temas, entre eles a educação, sem que nesse caso tenha a profundidade que tem o Censo da Educação Básica. Por serem coletados diretamente junto às famílias, os dados da Pnad são considerados mais fidedignos que os do Censo da Educação Básica, principalmente quando se busca acompanhar, por exemplo, taxas de evasão escolar, ingresso por idade e aprovação (Golgher, Rios-Neto, 2005).

Os números finais do Censo da Educação Básica podem ser usados, dentro dos limites que possuem, para definir algumas políticas educacionais. Por exemplo, dados referentes a elevadas taxas de reprovação podem fornecer indicações de que alguma coisa grave está acontecendo no interior de uma escola, talvez ligado a processos educacionais inadequados. Portanto, indicam uma determinada situação, mas não são capazes de fornecer detalhes, pois no caso do exemplo citado, os dados são claramente insuficientes para se compreender o que está acontecendo naquela escola. Dessa forma, começamos a perceber que os indicadores, apesar de serem importantes instrumentos para a definição e avaliação de políticas públicas, apresentam limitações: eles indicam um fato, mas são incapazes de explicar os porquês desse fato estar ocorrendo. Para entender os porquês, são necessários estudos mais aprofundados que venham a determinar quais seriam os fatores que estariam sendo responsáveis pelo fato indicado.

Uma discussão importante, e também não consensual, envolvendo a avaliação é a identificação das diversas instâncias que são responsabilizadas pelos resultados obtidos nela. Segundo alguns, responsabilizar é indicar os responsáveis pelo sucesso ou fracasso com base nesses resultados; sem dúvida, essa é uma forma muito limitada de compreender a responsabilização. Essa questão é refletida de maneira diferente por alguns autores.

Ristoff (1995, p. 63) acentua o caráter ético da avaliação como um valor inerente ao próprio processo, colocando-a a serviço de valores

públicos e dos justos interesses da sociedade: “avaliar é uma forma de restabelecer compromissos com a sociedade; de repensar objetivos, modos de atuação e resultados; de estudar, propor e implementar mudanças nas instituições e em seus programas; deve-se avaliar para poder planejar, para evoluir”.

Uma outra forma de compreender a questão é considerar a educação como um sistema que funciona com responsáveis que prestam contas pela forma como utilizam seus recursos e pelos resultados que obtêm. Para esta forma de encarar o resultado da avaliação a pergunta central é: quem deve fazer-se responsável e como?

Para Ravela (2005), existe uma lógica que implica numa guerra de todos contra todos, na qual todos se desfazem da sua responsabilidade e a atribuem a outros. O professor responsabiliza o aluno porque ele não estuda, a família do aluno porque não o apóia e o sistema ou as autoridades porque não dão aos professores meios necessários para desenvolver seu trabalho. A família responsabiliza o docente e a escola porque não ensinam ou porque não ensinam como anteriormente. As autoridades, por meio dos sistemas de prestação de contas/responsabilização pelos resultados, buscam responsabilizar as escolas e os professores e exercer pressão sobre eles. As autoridades centrais responsabilizam as autoridades locais e vice-versa. Os políticos, que dirigem o sistema durante um período limitado, responsabilizam os funcionários de carreira, que formam parte do aparato técnico e burocrático permanente, e vice-versa. Por esse caminho, os sistemas não avançam nem melhoram. A alternativa seria, segundo o autor, estabelecer uma lógica de colaboração entre os atores.

A aprendizagem, para Ravela (2005), não é o resultado direto e automático do que faz o docente. Seu trabalho é essencial, mas para que o aluno aprenda é necessário uma acumulação cultural prévia. Depende da sua motivação para aprender, que deve ser construída com a ajuda do professor, mas que também está determinada por valores sociais (tem sentido estudar?), pela importância que a família dá à educação e pelo apoio concreto que ela pode oferecer ao aluno durante seu processo de aprendizagem. Depende também do ambiente na sua escola, assim como do ambiente social e cultural em relação à educação.

Na linha da relação entre avaliação e sociedade, Paro (1998) salienta que todos aqueles envolvidos no processo educacional têm refletido muito pouco sobre a utilidade dos serviços que oferecem à sociedade, ao mesmo tempo que os próprios usuários e contribuintes não demonstram consciência sobre o que deve ser exigido da escola. O autor chama atenção para o fato de que a escola pública tem resistido a responsabilizar-se pelo seu produto, do qual ela deve prestar contas ao Estado e à sociedade.

Essa responsabilização pode ser entendida como um dos desafios a ser perseguido, como aponta Souza (1998), na medida em que a avaliação pode ajudar oferecendo elementos para se analisar em que se pode aperfeiçoar o ensino.

Diante dessas considerações, responsabilização da escola não significa que ela deva assumir, sozinha, o ônus do processo educacional e que o Estado deva se ausentar de seu papel constitucional de prover os recursos tanto materiais quanto técnicos para que o processo educacional possa acontecer de forma a garantir o sucesso do aluno. Apesar de não responderem unicamente pelo sucesso dos alunos, o diretor, a equipe pedagógica e os professores que atuam na escola têm um papel importante em todo o processo educacional, e exatamente devido a este papel de destaque é que devem prestar contas à sociedade, em especial aos alunos e seus familiares, sobre seus atos, assumindo, ao mesmo tempo, responsabilidades sobre o resultado desses atos. Isso não significa que se deva imputar a esses agentes uma culpabilidade sobre os resultados negativos que o processo educacional vem alcançando; mas deve-se buscar formas de colaboração em que juntos – agentes, poder público e sociedade – tenham fóruns adequados que lhes permitam discutir esses resultados e encontrar soluções que possibilitem superar os problemas encontrados.

Entendo como responsabilização, um processo em que ao mesmo tempo que se identificam as causas e os resultados dos problemas existentes nos processos educacionais, buscam-se alternativas que levem a encontrar soluções que possam modificar as situações problemáticas. Assim, prestar contas à sociedade do trabalho que está sendo desenvolvido na escola é uma das formas de aprimorar a qualidade da educação. A responsabilidade pelos resultados deve envolver o governo federal, os governos regionais, os profissionais que atuam na escola e a comunidade escolar. Os resultados educacionais dos estudantes são o produto de um complexo conjunto de fatores (que inclui o próprio esforço individual do estudante para aprender), e que cada ator tem a responsabilidade de encontrar as melhores formas de aperfeiçoá-los dentro do âmbito de suas competências. Assim, a responsabilidade é compartilhada por todos.

Outro aspecto importante no processo de responsabilização é a possibilidade de a sociedade ter acesso a documentos que identifiquem os resultados alcançados, comparando-os com os resultados desejados. Ao mesmo tempo, esses documentos devem garantir a possibilidade de construir séries históricas consistentes que permitam a comparação desses resultados ao longo do tempo.

Como a escola se encontra integrada a uma política educacional, Dalben (2003), reforçando a importância da responsabilização, afirma que é fundamental que os diferentes níveis do sistema educacional (governo, instituições, professores, comunidade escolar) sejam responsabilizados em relação aos resultados obtidos. Ao mesmo tempo, sugere que os professores se debrucem sobre os resultados das avaliações para também exigir a implementação de novas políticas de gestão do sistema.

Uma breve referência ao percurso da avaliação externa até o Saeb ganhar centralidade na política educacional

No caso brasileiro, um árduo caminho foi percorrido até que fosse instituído um instrumento de avaliação externo voltado para a educação básica, com a institucionalização do Sistema de Avaliação da Educação Básica (Saeb).

Em Horta Neto (2007) são analisados documentos oficiais e legislação federal, traçando o caminho que parte das primeiras medições educacionais no início do século 20 até a avaliação educacional ganhar destaque na Lei de Diretrizes e Bases de Educação Nacional (LDB), de 1996, passando a ser parte integrante da Lei 10.172/01 que instituiu o Plano Nacional de Educação (PNE). Esse trabalho indica que apesar de o Saeb ter ganhado centralidade em uma determinada época, ele não foi um instrumento desenvolvido por um único governo.

O primeiro esboço do que seria o Saeb foi um instrumento desenvolvido nos anos 1980, que tinha por objetivo avaliar os impactos do Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste (Edurural). Esse Programa – desenvolvido pelo Ministério da Educação (MEC) – era voltado para os anos iniciais do ensino fundamental (educação primária, nos termos da época) da região rural nordestina, e 35% de seus recursos foram financiados pelo Banco Mundial. O instrumento utilizado foi uma prova para avaliar o desempenho dos alunos que residiam em municípios atendidos pelo Programa, tendo sido contratada pelo MEC a Fundação Carlos Chagas para desenvolver os estudos, elaborar os projetos e implementar essa avaliação. Os dados foram coletados nos anos de 1981, 1983 e 1985, em 60 municípios de três Estados: Pernambuco, Ceará e Piauí, em uma amostra de 600 escolas e 6.000 alunos.

Com base nessa experiência, o MEC decidiu expandir a avaliação a todo o território nacional. Em novembro de 1988, foi feito o teste piloto nos Estados do Paraná e do Rio Grande do Norte, com o objetivo de testar os instrumentos e os procedimentos com vistas à sua aplicação em nível nacional no início de 1989. Em virtude de problemas orçamentários do MEC, a avaliação, chamada de Sistema de Avaliação da Educação Primária (Saep), foi aplicada em agosto de 1990 (Brasil, 1992). Em 1991, ela assumiu seu nome atual, Sistema de Avaliação da Educação Básica (Saeb).

O ciclo seguinte do Saeb, previsto para acontecer em 1992, só foi aplicado no ano seguinte, novamente devido a problemas orçamentários do MEC. A partir de 1993, a avaliação vem ocorrendo a cada dois anos. Em termos mundiais, o desenvolvimento de sistemas nacionais de avaliação, que começou nos Estados Unidos em fins dos anos 1960, expande-se para a Europa na década de 70 e nos anos 80 atinge a Ásia e a Oceania. Na década de 1990, em quase todos os países da América Latina foram criados sistemas nacionais de avaliação de aprendizagem, com destaque para Cuba, que realizou sua primeira avaliação em 1975.

A Tabela 1 apresenta a evolução da amostra utilizada pelo Saeb, das séries e disciplinas avaliadas no período compreendido entre 1990 e 2003.

Tabela 1 – Evolução da amostra utilizada pelo Sistema de Avaliação da Educação Básica (Saeb) – 1990-2003

Participantes	Ano						
	1990 ⁽¹⁾	1993 ⁽¹⁾	1995	1997	1999	2001	2003
Escolas	4.790	2.806	2.839	1.933	6.890	6.935	6.270
Diretores	3.546	2.294	2.214	2.302 ⁽³⁾	6.890	6.820	6.000
Professores	17.814	7.814	4.971	13.267	53.815	21.754	17.000
Alunos	108.982	133.114	90.499	167.196	279.764	287.719	275.081
Séries	1 ^a , 3 ^a , 5 ^a e 7 ^a	1 ^a , 3 ^a , 5 ^a e 7 ^a	4 ^a , 8 ^a , 2 ^a e 3 ^a	4 ^a , 8 ^a e 3 ^a	4 ^a , 8 ^a e 3 ^a	4 ^a , 8 ^a e 3 ^a	4 ^a , 8 ^a e 3 ^a
Disciplinas avaliadas	Matemática	Matemática	Matemática	Matemática	Matemática	Matemática	Matemática
	Língua Portuguesa	Língua Portuguesa	Língua Portuguesa	Língua Portuguesa	Língua Portuguesa	Língua Portuguesa	Língua Portuguesa
	Ciências	Ciências		Ciências ⁽⁴⁾	Geografia		
	Redação ⁽²⁾				História		
				Ciências ⁽⁴⁾			

Notas:

⁽¹⁾ As provas aplicadas às 1^a e 3^a séries foram de Língua Portuguesa e Matemática, de caráter semiobjetivo. Nas 5^a e 7^a séries, além delas, foram aplicadas também a de Ciências, todas de caráter objetivo (Brasil. Inep, 1992, 1995).

⁽²⁾ A prova de redação foi aplicada somente aos alunos das 5^a e 7^a séries. Apenas fizeram essa prova aqueles que realizaram também a de Língua Portuguesa (Brasil. Inep, 1992).

⁽³⁾ Há dupla contagem de diretores, quando a escola possui duas ou mais séries avaliadas.

⁽⁴⁾ Na 3^a série do ensino médio, as disciplinas de Ciências são Biologia, Física e Química (Brasil. Inep, 2000).

Para o ciclo de 2005, o Saeb foi modificado, transformando-se em um sistema composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica (Aneb), que mantém a mesma característica e objetivos do Saeb aplicado até 2003, e a Avaliação Nacional do Rendimento Escolar (Anresc), também conhecida como Prova Brasil, destinada a avaliar apenas as escolas públicas do ensino básico.

Paralelamente aos esforços do governo federal, outras instituições produziram diversas pesquisas na área. Destacam-se os estudos produzidos pela Universidade Federal do Ceará, entre os anos de 1970 e 1980; os estudos produzidos pela Fundação Carlos Chagas, que também realizou as pesquisas Avaliação da Jornada Única, em 1992, e a Avaliação do Processo de Inovações no Ciclo Básico e seu Impacto sobre a Situação de Ensino/Aprendizagem na Região Metropolitana de São Paulo, no período entre 1992 e 1994 (Vianna, 2005).

Instituíram-se também sistemas de avaliação em diversos Estados, destacando-se: o Programa de Avaliação do Desempenho da Rede Pública Escolar do Estado de Pernambuco, de 1991; o Programa de Avaliação das Escolas da Rede Estadual de Ensino de Minas Gerais, de 1992, que deu origem ao Sistema Mineiro de Avaliação da Educação Pública (Simave), existente desde 1999; o Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp), iniciado em 1996 e em pleno funcionamento; na Bahia, foi criada uma Agência de Avaliação reunindo diversos órgãos, entre eles a Secretaria Estadual de Educação e a Universidade Federal da Bahia, formando uma agência externa com o objetivo de conduzir o processo de avaliação; no Ceará, em 1992, foi criado um Programa de Avaliação do Rendimento Escolar dos Alunos de 4^a e 8^a séries, que em 2000 institucionalizou-se como Sistema Permanente de Avaliação Educacional do Ceará (Spaee); o Estado do Paraná

institui, em 1995, o Programa de Avaliação do Rendimento Escolar do Paraná; em Pernambuco, os estudos sobre avaliação que se iniciaram ainda nos anos 1980 levaram a criação, em 2000, do Sistema de Avaliação Educacional de Pernambuco – Saepe (Vianna, 2005; Bonamino, Bessa e Franco, 2004). Além desses, diversos municípios têm procurado desenvolver sistemas de avaliação que seguem o modelo adotado pelo Saeb.

Considerações finais

Foram discutidos neste trabalho três temas que estão afetos à avaliação educacional em larga escala, quais sejam: a qualidade educacional, os indicadores utilizados para medir a educação e finalmente a prestação de contas.

Melhorar a qualidade do ensino vem sendo um dos maiores objetivos dos sistemas de ensino e dos esforços dos pesquisadores em educação. Contudo, como o termo qualidade tem diferentes significados é muito difícil chegar a um consenso sobre qual o seu significado em determinado contexto educativo. Conceituar qualidade não é uma tarefa simples nem existem consensos sobre o melhor caminho a seguir. Como é um conceito polissêmico, primeiramente é preciso identificar qual é o conceito de qualidade a que estamos nos referindo. No caso da definição e avaliação de políticas públicas, defendo que a qualidade tenha características tais que permitam que ela possa ser mensurada, sem que se perca a noção de que essa medida não é absoluta e que deve ser relativizada, pois ela depende do contexto em que foi analisada e está intimamente ligada ao objetivo para a qual foi desenvolvida.

Nesse sentido, defendo que a avaliação educacional pode ser um instrumento importante para avaliar as políticas educacionais desenvolvidas pelos diversos níveis de governo e os seus resultados devem ser discutidos e utilizados para modificar os aspectos negativos identificados.

No caso brasileiro, o Saeb deve ser cada vez mais aprimorado. Uma de suas fraquezas ainda é a divulgação de seus resultados. É importante desenvolver instrumentos e estratégias que possibilitem que tanto os profissionais da educação como a sociedade em geral possam se apropriar de seus resultados. Para tanto, é importante que a linguagem utilizada na sua divulgação seja nítida e que aponte com clareza, inclusive com a apresentação de exemplos, qual o significado das escalas apresentadas e de como seus resultados podem ser utilizados para melhorar a qualidade da educação brasileira.

Referências bibliográficas

BARRETO, Elba S. S.; A avaliação na educação básica entre dois modelos. *Educação e sociedade*, Campinas, v. 22, n. 75, p. 48-66, ago. 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_issuetoc&pid=0101-733020010002&lng=pt&nrm=iso>.

BARRETO, Elba S. S. et al. Avaliação na educação básica nos anos 90, segundo os periódicos acadêmicos. *Cadernos de Pesquisa*, São Paulo, n. 114, p. 49-88, nov. 2001a.

BELLONI, Isaura; BELLONI, José Ângelo. Questões e propostas para uma avaliação institucional formativa. In: FREITAS, Luiz Carlos (Org.), BELLONI, Isaura; SOARES, José Francisco. Avaliação de escolas e universidades. Campinas: Komedi, 2003.

BONAMINO, Alicia; BESSA, Nícia; FRANCO, Creso (Org.). *Avaliação da educação básica*. Rio de Janeiro: Ed. PUC-Rio; São Paulo: Loyola, 2004.

BRASIL Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Relatório do Sistema Nacional de Avaliação da Educação Básica: ciclo 1990*. Brasília, 1992.

_____. *Relatório do 2º ciclo do Sistema de Avaliação da Educação Básica: Saeb 1993*. Brasília, 1995.

_____. *Sistema Nacional de Avaliação da Educação Básica: resultados 99*. Brasília, 2000.

_____. *Dicionário de indicadores educacionais: fórmulas de cálculo*. Brasília, 2004.

CARLEY, Michael. *Indicadores sociais: teoria e prática*. Rio de Janeiro: Zahar, 1985.

DALBEN, Ângela. Das avaliações exigidas às avaliações necessárias. In: VILLAS BOAS, Benigna M. F. *Avaliação: políticas e práticas*. Campinas: Papirus, 2003.

DIAS SOBRINHO, José. Campo e caminhos da avaliação: a avaliação da educação superior no Brasil. In: FREITAS, L. C. (Org.). *Avaliação: construindo o campo e a crítica*. Florianópolis: Insular, 2001.

_____. *Avaliação da educação superior*. Petrópolis, Vozes, 2000.

_____. *Avaliação institucional: instrumento da qualidade educativa*. In: BALZAN, Newton; DIAS SOBRINHO, José (Orgs.). *Avaliação institucional: teoria e experiências*. São Paulo: Cortez, 1995.

FRANCO, M. L. B. Pressupostos epistemológicos de avaliação educacional. *Cadernos de Pesquisa*, São Paulo, n. 74, p. 63-67, 1990.

GENTILI, Pablo; SILVA, Tomaz Tadeu da. *Neoliberalismo, qualidade total e educação: visões críticas*. Petrópolis: Vozes, 1995.

GOLGHER, André B.; RIOS-NETO, Eduardo L. G. *Uma comparação entre os modelos Profluxo e IPC quando aplicados aos dados do sistema educacional brasileiro*. Brasília: Inep, 2005. (Série Documental, v. 19).

HORTA NETO, João Luiz. *Os desafios do SINAES*. Trabalho apresentado no XXII Simpósio Brasileiro de Política e Administração da Educação. Rio de Janeiro, 2005.

_____. Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o SAEB de 2005. *Revista Iberoamericana de Educación*, Madrid, v. 42, n. 5, p. 1-14, abr. 2007. Disponível em: <<http://www.rieoei.org/deloslectores/1533Horta.pdf>>.

LUCKESI, C. Avaliação da aprendizagem escolar: apontamento sobre a pedagogia do exame. *Tecnologia Educacional*, Rio de Janeiro, v. 20, n. 101, p. 82-86, jul./ago. 1991.

_____. Avaliação educacional: pressupostos conceituais. *Tecnologia Educacional*, Rio de Janeiro, v. 25, n. 130/131, p. 26-29, maio/ago. 1996.

_____. Planejamento e avaliação na escola: articulação e necessária determinação ideológica. *Série Ideias*, FDE, São Paulo, n. 15, p. 115-125, 1992.

MENDONÇA, Erasto Fortes. *Inclusão e sucesso escolar: elementos para formulação de indicadores de qualidade social da educação*. Brasília, 2005. [não publicado].

GENRO, Tarso [ministro da Educação]. 2005 será o ano da educação. [entrevistadora: Paola Gentile]. *Nova Escola*, São Paulo, n. 180, Seção "Fala, mestre!", mar. 2005.

PARO, Vitor. A gestão da educação ante as exigências de qualidade e produtividade da escola pública. In: SILVA, Luiz (Org.). *A escola cidadã no contexto da globalização*. Petrópolis: Vozes, 1998.

PIRSIG, Robert M. *Zen e a arte da manutenção de motocicletas: uma investigação sobre valores*. Rio de Janeiro: Paz e Terra, 1984.

RAVELA, Pedro. *Para comprender las evaluaciones educativas: fichas didacticas*. Santiago do Chile: Preal, 2005.

RISTOFF, Dilvo. Avaliação institucional: pensando princípios. In: BALZAN, Newton; DIAS SOBRINHO, José (Orgs.). *Avaliação institucional: teoria e experiências*. São Paulo: Cortez, 1995. p. 37-51.

SAUL, A. M. Avaliação emancipatória: uma abordagem crítico-transformadora. *Tecnologia Educacional*, Rio de Janeiro, v.21, n. 104, p. 24-31, jan./fev. 1992.

SCHWARTZMAN, Simon. La calidad de la educación superior en América Latina. In: SEMINARIO PERMANENTE SOBRE LA CALIDAD, EFICIENCIA Y EQUIDAD DE LA EDUCACIÓN SUPERIOR COLOMBIANA. 1990, Bogotá. *Anais do...*Bogotá: ICFES, 1990. Tomo I, p. 27-60.

SOUZA, Clarilza Prado. Descrição de uma trajetória na/da avaliação educacional. *Série Idéias*, FDE, São Paulo, n. 30, p. 161-174, 1998.

TOMASSIA, Cláudia V. *Avaliação educacional: experiências e metodologias*. Texto apresentado no evento Escola de Avaliação, promovido pela Unesco em Brasília, em 2004. [Não publicado].

TYLER, Ralph. *An interview with Ralph Tyler conducted by Jeri Ridings Nowakowski*. Ann Arbor, MI: University of Michigan, 1981. (Occasional Paper Series, n. 13). Disponível em <<http://www.wmich.edu/evalctr/pubs/ops/ops13.html>>. Acesso em 19 maio 2009.

TURCHI, Maria Lenita. *Qualidade total: afinal de que estamos falando?* Brasília: IPEA, 1997. (Texto para discussão n. 459).

VIANNA, Heraldo Marelim. *Fundamentos de um programa de avaliação educacional*. Brasília: Líber Livro, 2005.

_____. Fundamentos de um programa de avaliação educacional. *Estudos em Avaliação Educacional*, São Paulo, n. 28, p. 23-37, jul./dez. 2003.

_____. Programas de avaliação em larga escala: algumas observações. *Estudos em Avaliação Educacional*, São Paulo, n. 23, p. 93-104, jan./jun. 2001.

João Luiz Horta Neto, doutorando em Políticas Públicas pela Universidade de Brasília (UnB), é pesquisador-tecnologista em avaliação e informações educacionais do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), atuando na Coordenação de Instrumentos e Medidas da Diretoria de Avaliação da Educação Básica.
jlhorta@gmail.com

Recebido em 2 de setembro de 2009.

Aprovado em 15 de dezembro de 2009.