

POLÍTICAS DE AVALIAÇÃO EM LARGA ESCALA DA EDUCAÇÃO BÁSICA: REVISÃO DA LITERATURA

Elisângela Teixeira Gomes Dias¹

RESUMO

Estudo de revisão sistemática, do tipo exploratório, que tem por objetivo analisar as pesquisas que abordam as políticas brasileiras de avaliação em larga escala, com foco na Educação Básica, no marco temporal de 2002 a 2011. O levantamento foi feito nos bancos de dados do Instituto Brasileiro de Informação em Ciência e em Tecnologia (Ibict), na base de dados *Scientific Electronic Library Online* (Scielo) e também nos sítios da Associação Nacional de Pós-Graduação e Pesquisa em Educação (Anped) e da Associação Brasileira de Avaliação Educacional (Abave). Os trabalhos foram organizados e analisados por temáticas afins, considerando os objetivos explicitados pelos autores. Somente 3% são estudos de revisão, mas que não abordam especificamente a temática das avaliações externas. Em linhas gerais, aproximadamente 25% das pesquisas propõem ou analisam instrumentos e metodologias para avaliação de larga escala; 9% discutem a regulação dos sistemas via avaliação externa e analisam as consequências das políticas de accountability; 20% analisam o impacto ou desenvolvimento de políticas públicas voltadas para o sistema de avaliação externa; 33% investigam o desempenho dos estudantes, escolas e redes de ensino em avaliações externas e 11% analisam os usos, impactos e implicações dos resultados dos exames em larga escala na dinâmica escolar. Os estudos evidenciaram que, observando o cenário das políticas públicas educacionais do país, é possível afirmar que a máxima “mais avaliação, mais qualidade” vem sendo disseminada, validando o discurso hegemônico do Estado. Em geral, as pesquisas analisadas têm caráter descritivo e estão mais no campo da crítica, sendo pouco propositivas, mas revelam diferentes saberes que precisam ser considerados para o desenvolvimento do campo da avaliação.

Palavras-chaves: Avaliação. Avaliação externa. Avaliação em larga escala.

POLICIES OF LARGE-SCALE EVALUATION OF BASIC EDUCATION: LITERATURE REVIEW

ABSTRACT

A systematic review, exploratory, with the objective to analyze researches on Brazilian policies of large-scale assessment, focusing on Basic Education

Recebido em: 15/4/2013.

¹ Doutoranda em Educação na UNB. Professora da SEDF. Integrante do Grupo de Estudos e Pesquisa em Avaliação e Organização do Trabalho Pedagógico. E-mail: eliangelagomesdias@gmail.com

in the time frame 2002-2011. The survey was conducted in the databases of the Brazilian Institute of Information in Science and Technology (Ibict), in the database Scientific Electronic Library Online (SciELO) and also on the websites of the National Association of Graduate and Research in Education (Anped) and Brazilian Educational Assessment Association (ABAVE). The work was organized and analyzed for themes related, considering the objectives spelled out by the authors. Only 3% are review studies, but did not specifically address the issue of external evaluations. In general, approximately 25% of research and propose tools and methodologies to analyze large-scale evaluation; 9% discuss the regulation of systems via external evaluation and analyze the consequences of political accountability; 20% analyze the impact or development of public policies for the system of external evaluation, 33% investigated the performance of students, schools and networks teaching and 11% examine the uses, impacts and implications of the results of tests on a large scale in the school dynamics. Studies have shown that, observing the scene of the educational policies of the country, we can say that the maxim “more evaluation, more quality” is being spread; validating the hegemonic discourse of the state. In general, studies have analyzed descriptive character and are more in the field of criticism, being little propositional, but reveal different skills that need to be considered for the development of the field of evaluation.

Keywords: Avaluation. External evaluation. Large-scale assessment.

1 Considerações Iniciais

A reforma dos Estados nacionais, consequência do ideário neoliberal instaurado após as crises do capitalismo, engendraram alterações substantivas nos padrões de intervenção estatal, redirecionando as formas de organização e gestão e, conseqüentemente, os modos de regulação das políticas públicas. Estas medidas, segundo Barroso (2005), podem ser: (i) de ordem mais técnica, amparadas pelo discurso da modernização, desburocratização e combate à ineficiência do Estado; (ii) de natureza política, de acordo com projetos neoliberais e neoconservadores com a justificativa de “libertar a sociedade civil” do controle do Estado por meio da privatização; (iii) de natureza filosófica e cultural, com a defesa de uma participação comunitária; (iv) de natureza pedagógica, com vistas ao ensino centrado no estudante e suas características.

No Brasil, estas medidas, no que se refere às políticas de avaliação da Educação Básica desde o texto da Constituição Federal de 1988 (BRASIL, 2007), são de ordem técnica e também política. O artigo 206 da referida Constituição indica a “garantia de padrão de qua-

lidade” como um dos princípios basilares do ensino. Esta qualidade, conforme o artigo 209, surge como condição do ensino livre à iniciativa privada e como resultante de processos de indicação de resultados (art. 214), o que indica melhorias aferidas. É a partir da prerrogativa de que a avaliação é um importante meio para a melhoria da qualidade da educação é que um sistema nacional de avaliação (externa, em larga e com foco no rendimento dos estudantes e no desempenho dos sistemas de ensino) foi introduzido no país no final dos anos de 1980.

O interesse em implementar um sistema nacional de avaliação se insere historicamente na administração do sistema educacional brasileiro a partir das reformas neoliberais; consequência da crise estrutural do modo de produção capitalista. Para tanto, concorreram às perspectivas economicista e tecnicista no tratamento da questão da avaliação educacional, de modo articulado à construção científica dos fatores de qualidade, eficiência, equidade e produtividade e que definem um quadro hegemônico de propostas educacionais (COELHO, 2008).

Com efeito, o exame é considerado avaliação, por isso preocupa-se com o aprimoramento de técnicas de mensuração, de objetividade e de neutralidade. Assim, a avaliação é tomada como mecanismo da gestão da educação para elaboração de diagnósticos que possam orientar a implementação de políticas educacionais visando à melhoria da tão almejada “qualidade total²” e em defesa da promoção da equidade.

Desde então, são muitas as contradições e as tensões em torno das avaliações em larga escala, envolvendo questões diversas, como: “[...] universal versus amostral, série versus idade; muitas matérias do currículo versus alguns conteúdos; avaliação de competências versus avaliação do efeito da escola; autonomia da escola versus fortalecimento do poder central” (WERLE, 2010, p. 33). Nesse sentido, é que para Bonamino e Sousa (2012) há três gerações de avaliação em lar-

2 Qualidade total refere-se a um programa que teve início na década de 50, em empresas do Japão, e traz uma proposta denominada ‘novo paradigma’ em administração, o qual foi implantado no mundo inteiro. “As palavras de ordem do Programa são: eficiência, controle, competitividade. Procurando ir além da administração de caráter taylorista, que privilegiava a produtividade centrada na quantidade do trabalho, instala-se a palavra de ordem da qualidade. [...] É principalmente na segunda metade da década de 80 que o Programa de Qualidade Total se instala no Brasil, ganhando espaços em diversas organizações empresariais e estendendo-as instituições escolares” (RIOS, 2008, p. 72-73) como “contraface do discurso da democratização” presente nas escolas (GENTILLI, 1997).

ga escala, considerando os objetivos e desenhos usuais em iniciativas implementadas no Brasil.

Inicialmente, o foco esteve na avaliação diagnóstica da qualidade da educação, sem atribuição de consequências diretas para as escolas e para o currículo escolar. A primeira geração, portanto, era de baixo impacto (*low stakes*) ou responsabilização branda, como definem as autoras. As outras duas gerações, diferentemente, trazem consequências fortes (*high stakes*), pautadas na lógica da responsabilização que traz consequências simbólicas ou materiais para os agentes escolares (BONAMINO; SOUSA, 2012).

Observa-se, nesse contexto, uma transformação progressiva do discurso dominante sobre a escola. O fracasso escolar, antes atribuído exclusivamente ao aluno, agora é atribuído à insuficiência dos recursos das escolas, e principalmente ao despreparo dos professores ou mesmo à lógica confusa de um sistema globalmente deficiente (BOURDIEU; CHAMPAGNE, 1998). A exclusão se dá de maneira contínua, tornando a escola uma fonte de decepção coletiva.

O fato é que, com uma história de mais de vinte anos no Brasil, a avaliação se firma cada vez mais como elemento da regulação e da administração competitiva do “Estado-Avaliador” (COELHO, 2008), de forma que a compreensão das políticas públicas de avaliação direcionadas à Educação Básica, por meio da análise das ações empreendidas pelo Estado, implica a apreensão, no feixe dessas políticas, dos limites e possibilidades da proposição e materialização das mesmas. Por isso a relevância de uma reflexão acerca da produção científica que tem sido desenvolvida e disseminada.

O presente trabalho tem por objetivo analisar as pesquisas que abordam as políticas de avaliação em larga escala, com foco na Educação Básica, no período de uma década (2002 a 2011). Acredita-se que traçar um panorama sobre as produções científicas relacionadas ao tema da avaliação externa torna-se relevante para empreendermos uma análise da conjuntura das pesquisas educacionais produzidas e discutidas no país e evidenciarmos as contribuições e as lacunas ainda latentes no campo da avaliação educacional.

2 Metodologia

Estudo de revisão sistemática, do tipo exploratório, no marco temporal de 2006 a 2011. Foram pesquisados artigos científicos, teses e dissertações que se referiam as políticas de avaliação em larga escala na Educação Básica. Os termos utilizados para a identificação das pesquisas foram: “avaliação externa”; “avaliação larga escala”; “avaliação sistemas”; “avaliação redes”.

O levantamento do material foi realizado em três etapas. A primeira etapa da revisão de literatura concentrou-se sobre as teses e dissertações nacionais. Para tanto, a pesquisa foi realizada no Banco de dados do Instituto Brasileiro de Informação em Ciência e em Tecnologia (IBICT) e também foram feitas buscas nas bibliotecas digitais da Universidade de Brasília, da Universidade Estadual de Campinas (Unicamp), da Universidade Federal de Minas Gerais e da Universidade de São Paulo (USP).

A segunda etapa da revisão de literatura compreendeu a seleção de artigos científicos publicados em revistas indexadas na base de dados *Scientific Electronic Library Online* (SciELO).

A terceira etapa foi a análise da produção científica disseminada em dois importantes congressos anuais: o da Associação Nacional de Pós-Graduação e Pesquisa em Educação (ANPED), reconhecido como o mais expressivo fórum de debates na área educacional, e a Associação Brasileira de Avaliação Educacional (ABAVE), fundação criada em 2003 com o objetivo de favorecer o intercâmbio de experiências entre os acadêmicos e os implementadores da avaliação educacional.

Destaca-se que a maioria dos trabalhos selecionados foi lida na íntegra, pois os resumos nem sempre continham todos os elementos necessários para compreensão da pesquisa empreendida pelos autores e, algumas vezes, deixavam dúvidas quanto ao objetivo do estudo ou resultados alcançados.

Partindo do quadro demonstrativo do exame das produções científicas a respeito das avaliações em larga escala na Educação Básica, os trabalhos foram organizados por temáticas afins, considerando os objetivos de estudo explicitados pelos autores.

3 Resultados e Discussão

A tabela a seguir sintetiza o total de trabalhos selecionados, os quais foram distribuídos em seis categorias que serão posteriormente analisadas.

Tabela 1 – Distribuição dos trabalhos por temáticas afins

Categoria	Nº	%
1. Estudos de revisão de literatura.	05	03
2. Proposição, análise de instrumentos ou metodologias de exames aplicados em larga escala.	37	25
3. Análise de políticas de <i>accountability</i> e regulação via exames externos.	13	09
4. Análises de políticas de exames externos municipais, estaduais ou federal.	30	20
5. Estudos com foco na análise do desempenho dos estudantes, escolas ou redes de ensino em exames externos.	49	33
6. Usos, impactos e implicações dos resultados dos exames externos.	17	11
Total Geral	151	100%

Fonte: elaboração da autora.

Com este levantamento, constatou-se que o número de teses elaboradas no país abordando a temática da avaliação externa na Educação Básica ainda é pequeno. Do total de trabalhos encontrados, apenas 6% são teses, sendo que mais da metade estão voltadas para área da Estatística, propondo ou analisando metodologias para construção de instrumentos de avaliações externas. As dissertações representam 24% dos trabalhos selecionados. Os demais estudos (70%) são artigos científicos indexados em revistas científicas ou apresentados em anais de congressos.

É possível proceder diferentes análises a partir do mapeamento dos trabalhos que abordam esta temática. Entretanto, será realizada uma análise integrativa por categoria. Ressalta-se que alguns estudos abordam mais de uma das temáticas categorizadas, mas o agrupamento está em conformidade com o objetivo anunciado na pesquisa e resultados encontrados.

a) Estudos de revisão de literatura

Há mais de duas décadas o Brasil implementou a avaliação educacional em larga escala, também conhecida como avaliação externa. Mas somente dois estudos fizeram um mapeamento abordando especificamente esta temática.

Recentemente, Fachinetti e Lima (2011) buscaram compreender quantas teses e dissertações tratavam da temática da “avaliação

institucional em larga escala” no ano de 2006 em 17 cursos de pós-graduação em Educação no país. Os autores concluíram que a temática é pouco abordada, embora seja de “extrema” importância. Este mapeamento, no entanto, demonstra duas limitações: a) a pesquisa se limitou em apenas um ano, desconsiderando inclusive os quatro anos posteriores que antecederam a publicação do artigo (2007 a 2010); b) os autores não tratam da avaliação institucional, mas sim da avaliação em larga escala, o que revela o equívoco conceitual em relação ao termo, usado em todo texto como sinônimos.

Borges e Calderón (2011) mapearam a produção científica da “Revista Ensaio: Avaliação e Políticas Públicas em Educação”, no período de 1999 a 2008. De um total de 284 artigos publicados pela revista neste período, 131 (46, 12%) tratam do tema avaliação educacional, sendo que 89 (67, 93%) destes referem-se à Educação Básica, enquanto os demais (32,06%) abordam a Educação Superior. Os 89 artigos abordam temáticas diversas relacionadas à área de avaliação na Educação Básica, como práticas pedagógicas, formação docente e avaliação, instrumentos para avaliação, avaliação de programas e projetos, avaliação de instituições escolares, SAEB, ENEM, qualidade e educação, entre outros. Entretanto, como o estudo não focalizou o descritor “avaliação externa” ou “avaliação em larga escala”, não foi possível apresentar considerações conclusivas acerca deste nível de avaliação educacional ou mesmo precisar quantos artigos que tratam deste tema.

No entanto, nesta categoria incluir-se-á três trabalhos que dissertam sobre a temática, como também historicizam o contexto de implantação deste nível da avaliação educacional no Brasil.

Figueiredo (2009) e Gonçalves (2006) discutem que, no contexto da crise da dívida externa, a educação é concebida como parte do conjunto das reformas econômicas. Os projetos financiados pelo Banco Mundial (BM) para o ensino fundamental fazem parte do conjunto dos empréstimos de ajustes estruturais e setoriais que foram intensificados a partir da década de 1990, notadamente iniciados com as reformas engendradas com Fernando Collor de Mello (1999-1992). É neste contexto que os projetos financiados para o Ensino Fundamental apoiaram e orientaram as políticas para enfrentar o fracasso escolar (evasão e repetência), as quais trataram do processo de avaliação externa, de racionalidade econômica, de produtividade e de concorrência bem como os critérios de qualidade.

Vieira e Tenório (2009) discorrem sobre o tema da avaliação em educação incluindo o nível de avaliação em larga escala. Os autores fazem uma breve reflexão quanto ao uso dos resultados das avaliações externas, destacando os resultados do SAEB publicizados em fevereiro de 2007, mas não apontam argumentos conclusivos. A discussão feita é sintetizada em um quadro que apresenta os principais argumentos contrários à avaliação em larga escala a partir dos estudos de Carlini e Vieira (2005), Crepaldi, Santos e Galindo (2005) e Convic, Tavares e Machado (2005). Em seguida, Vieira e Tenório (2009) discutem o tema a partir do panorama apresentado por Guba e Lincoln, autores que abordam o campo da avaliação sob uma perspectiva de quatro gerações (medida, descrição de objetivos, julgamento, avaliação construtivista responsiva³), com a finalidade de questionar seus modelos correntes no final da década de 1980, e sua articulação com a questão da sustentabilidade. Vieira e Tenório (2009) identificam as lacunas presente nas proposições de Guba e Lincoln e apresentam um novo conceito que, para os autores, inclui todas as dimensões da avaliação: “[...] avaliar é o diagnóstico para a tomada de decisão com vista na melhoria do processo” (p. 11). Concluem discutindo o papel do avaliador sob a crença na sustentabilidade, a qual para eles inverte o resultado da avaliação para quem está sendo avaliado, ampliando sua ação quando todos os envolvidos se comprometem não apenas com a eficácia dos resultados, mas também com a efetividade destes resultados.

Neste contexto, observa-se a relevância de empreendermos uma revisão bibliográfica sobre a temática da avaliação externa na Educação Básica.

b) Proposição, análise de instrumentos ou metodologias de exames aplicados em larga escala

Um grupo de autores considera que a avaliação externa pode contribuir para ajudar os gestores das redes de ensino e os professores a compreenderem a realidade da sua comunidade escolar, fornecendo informações que serviriam de diretrizes em busca da melhoria do ensino (FREITAS, 2007; MOREIRA; SORDI, 2004; PERRY, 2009). Para tanto, é necessário que a avaliação não seja reduzida a apenas

3 O termo “avaliação construtivista responsiva” foi utilizado por Guba e Lincoln a partir da perspectiva do paradigma interpretativo ou hermenêutico, tendo como principal característica a defesa de práticas avaliativas participativas e negociadas.

um exame e que seus dados possam ser compreendidos no interior de cada escola, de forma que gestores e professores entendam seus objetivos e seus processos, interpretando os resultados e analisando como os mesmos poderão influenciar na tomada de decisão em relação ao trabalho pedagógico desenvolvido na sala de aula, na escola e nas redes de ensino.

O problema é que os sistemas de avaliação em larga escala utilizam metodologias e procedimentos sofisticados, apoiados em um referencial diferente dos que utilizam os professores, mesmo porque não coincidem com os apresentados em sua formação, o que pode tornar estes dados incompreensíveis a eles e também aos gestores das escolas e das redes de ensino, imprimindo resistência quanto à utilização de seus resultados (BARRETO, 2001).

Um total de 37 trabalhos, os quais representam a aproximadamente 25% dos trabalhos sobre a temática investigada, propuseram ou analisaram algum instrumento ou metodologia de avaliação em larga escala. A maior parte dos estudos trata de pesquisas que focalizam a Teoria de Resposta ao Item (TRI).

A TRI é uma metodologia adotada no Brasil a partir da edição do SAEB de 1995 e que atribuiu maior validade e utilidade às informações coletadas na medida em que permite comparabilidade entre os anos e as séries avaliadas e considera o desempenho dos estudantes para definir o grau de dificuldade do item (SOUZA, 2009). Vale ressaltar que, embora a TRI seja utilizada no Brasil para produzir índices de proficiência dos estudantes nas avaliações externas, seus diferentes modelos permitem a construção de indicadores variados, como o indicador socioeconômico utilizado na pesquisa de Pereira (2004).

Outras pesquisas analisadas reforçaram, nesse contexto, a importância da avaliação externa considerar variáveis de valor agregado, como a influência de fatores demográficos, sociológicos e relativos à trajetória escolar e fatores socioeconômicos. O estudo de Andrade e Laros (2006), por exemplo, teve como objetivo construir um modelo de desempenho escolar, utilizando-se de modelos lineares hierárquicos. Este modelo, na conclusão dos autores, pode agregar às avaliações externas variáveis diversas para compreender os desempenhos dos estudantes, além da possibilidade de ser utilizado para elaboração de políticas públicas mais efetivas para o melhoramento do sistema educacional brasileiro. Costa (2011) também discute a pertinência do uso do modelo de valor agregado no sistema de avaliação em larga esca-

la. Este modelo, aplicado em pesquisa longitudinal em Belo Horizonte, permitiu um diagnóstico mais eficiente para subsidiar a construção de políticas públicas mais efetivas, conforme analisa o autor.

Miranda (2008) e Rodrigues, S. (2005) realizaram pesquisas em torno da técnica DEA (Análise por Envoltória de Dados em português), defendendo que este é um processo alternativo de condução de avaliação em nível de sistema de ensino fundamental que pode ser um instrumento útil à gestão dos objetivos da instituição, com vistas ao seu aperfeiçoamento. O foco é na articulação entre a avaliação institucional e a avaliação em larga escala. Os autores confirmaram que a DEA, por relacionar recursos e resultados de maneira simples, é uma importante ferramenta para o gestor e para a equipe da escola conhecer e discutir seus problemas, tendo como foco a superação das dificuldades e a melhoria da escola. Todavia, como os governos têm utilizado as avaliações externas segundo a teoria da responsabilização (FREITAS, 2011). Dos 151 trabalhos identificados, apenas estas duas teses abordam o uso desta técnica, o que reforça a necessidade de mais pesquisas que investiguem novas metodologias ou mesmo o uso da DEA.

Aguiar (2009) e Rodrigues, E. (2008) realizaram um estudo abordando a DIF (Funcionamento diferencial em português) e revelam que, considerando diferenças de regiões, gênero e raça ou outras especificidades, o procedimento mais adequado para validação dos itens que compõem os bancos de dados das avaliações externas nem sempre é a sua retirada dos testes.

Algumas pesquisas apresentaram ou discutiram ainda análises e interpretações pedagógicas das escalas de proficiência. De forma geral, essas apontam a necessidade urgente de um investimento em melhorias da divulgação e apresentação dos resultados dos testes, inclusive com possibilidades de usos dos resultados em favor da prática pedagógica. Concorda-se com a consideração de Perry (2009) de que o professor, independentemente do método utilizado para composição da escala, possa compreender seus objetivos e possibilidades, de forma a servir como construção de um diagnóstico pedagógico na medida em que este conseguir agregar suas observações diárias às informações da escala de proficiência.

Por fim, encontramos estudos que tratam da construção de metodologias para avaliar a produção escrita de estudantes do Ensino Fundamental e do Ensino Médio em exames externos de Língua

Portuguesa. Há ainda as pesquisas de Oliveira e Guimarães (2011) e Marques *et al.*, (2011), as quais analisaram a Provinha Brasil de Matemática, e apontaram conclusões que podem contribuir para aperfeiçoar o instrumento e desafiar os responsáveis pela elaboração das provas a buscarem metodologias cada vez mais adequadas.

c) Políticas de *accountability* e regulação via exames externos

A normatização da avaliação como política de governo foi acentuada nos anos de 1980 e 1990 com a expansão e a universalização da Educação Básica, o que foi se constituindo como uma nova ordem de direitos que demandou a atuação do Estado. Com isso, ao mesmo tempo em que a educação torna-se um espaço para construção da cidadania, novas competências são exigências ao trabalhador moderno, como responsabilidade, iniciativa, dedicação, capacidade de aprender a aprender e de trabalhar em equipe, etc. (SOUZA, 2009).

A retórica oficial é a de qualidade e excelência, mas estes termos são usados para abrandar a ideia de produtividade, segundo a lógica do mercado: aumentar a quantidade de rendimentos, com menores custos (DIAS SOBRINHO, 1998). É neste cenário, ligado ao contexto de crise do Estado desenvolvimentista e ao movimento reformista da década de 1990 que os exames externos são introduzidos como regulação da Educação Básica no Brasil, produzindo significativas mudanças na área educacional e na própria orientação das políticas públicas para a educação. Portanto, a montagem do sistema nacional de avaliação brasileiro, efetivado por meio da criação do SAEB, serviu para atender a lógica do mercado capitalista, embebida no discurso da transparência das instituições públicas (*accountability*) e colocando temas como equidade e qualidade em evidência.

Segundo Cassetari (2011), os programas de remuneração por desempenho para professores da Educação Básica, uma das estratégias da política de *accountability*, desdobram-se em três diferentes modelos. O primeiro é o pagamento por mérito individual, o qual depende de avaliações subjetivas dos professores, provocando disputas e conflitos que interferem de forma negativa no espaço escolar. Por isso este é o modelo mais criticado e menos adotado hoje. O segundo é a bonificação a partir dos resultados das escolas. Este modelo também é criticado por incentivar a disputa entre os professores e por ampliar esta competição entre as escolas. Além disso, esta política pode reduzir o currículo escolar às matrizes de referências dos exames,

reforçando a lógica da avaliação classificatória do espaço escolar. O terceiro modelo, que é o mais recente, baseia-se nos conhecimentos e habilidades dos professores aferidos por um único instrumento, em geral um exame de proficiência, ou envolvendo múltiplos instrumentos, dependendo do consenso sobre quais competências são necessárias para avaliar um bom desempenho profissional.

Como ressalta a autora, apesar do grande debate sobre a remuneração por desempenho na literatura e da falta de estudos conclusivos sobre os seus efeitos, diversos sistemas educacionais brasileiros têm adotado estas propostas, o que indica a necessidade de maior reflexão sobre o tema. Além disso, como revelam suas pesquisas (CASSETTARI, 2011, 2008) não existem consensos sobre qual seria a melhor maneira de executar estas propostas, pois são inúmeras as dificuldades técnicas tanto em relação ao processo de formulação quanto ao processo de implementação dos programas de pagamento por performance. Outra constatação é que não há evidências significativas sobre os resultados da implementação destes programas. Acrescenta-se que no contexto brasileiro não é possível atribuir aos professores a responsabilidade por um ensino de qualidade sem que as condições de trabalho, como formação docente e infraestrutura das escolas, sejam adequadas.

Várias pesquisas lançaram olhar diferenciado sobre as perspectivas e desafios na implantação de sistemas de responsabilização educacional no Brasil, focalizando a introdução dos programas de remuneração variável por desempenho educacional; a percepção de professores acerca do desenvolvimento de processos de *accountability* e o processo de regulação da rede de ensino. Dentre os resultados evidenciados, destaca-se que: a) a maioria dos professores desconhecem tanto a fase de elaboração quanto a fase de divulgação dos exames externos, o que torna os resultados gerados incompreensíveis a eles; b) os resultados das avaliações externas são pouco utilizados como subsídio para planejamento e implementação de políticas; c) a utilização das informações produzidas por exames em larga escala, principalmente diante dos *rankings* promovidos, incita a competitividade entre as escolas e evidencia o predomínio da regulação para o controle, na visão da teoria da responsabilização. Assim, cabe interrogar: a quem tem servido esses exames?

Silva (2011), analisando em sua tese o uso que os diretores escolares fazem dos resultados de avaliações externas implementadas

pelo Programa de Avaliação da Rede Pública de Educação Básica em Minas Gerais (PROEB), com foco no processo de regulação, acrescenta outras consequências da pressão por produtividade e alcance de metas. Entre elas, cita: estrangimentos, exposições públicas, ameaças e críticas, de forma que o processo de ensino-aprendizagem pode ser desqualificado pelo imediatismo, pela superficialidade e reducionismo dos conteúdos. Com efeito, “[...] incorre-se também na possibilidade de desvalorizar ainda mais o trabalhador docente ao julgar todo o seu trabalho pelos resultados de uma avaliação” (p. 187).

Registra-se ainda que muitas pesquisas verificaram a interferência de setores da economia nacional e de organismos internacionais, como o Banco Mundial e o FMI, instituindo a lógica mercadológica que iguala a escola como produto. No entanto, observaram que tal acontecimento não se deu sem confronto.

d) Análises de políticas de exames externos municipais, estaduais ou federal

Ao longo das duas últimas décadas, em diferentes contextos sociopolíticos, primeiro nos países centrais, depois nos países semi-periféricos e periféricos, a agenda avaliativa, nas suas diversas configurações e domínios de incidência, tem assumido uma enorme centralidade (SÁ, 2009).

Um conjunto de pesquisas analisou, sob diferentes perspectivas, as concepções e implicações dos “sistemas” de avaliação externa na área educacional. Não é objetivo deste estudo dissertar acerca de cada uma delas, mas sim de fazer uma leitura crítica dos seus resultados. Para facilitar a análise, as pesquisas foram organizadas em dois grupos que serão detalhados a seguir.

As políticas de avaliação externa em âmbito federal foram objeto de estudo e análise do primeiro grupo de pesquisas. Em linhas gerais, os resultados das investigações indicam que existe um hiato entre a construção, a disseminação de indicadores educacionais e o processo de formulação e implementação de políticas públicas educacionais. Há, neste sentido, a necessidade de revisão da concepção subjacente em nível de avaliação externa e da forma como esta vem sendo conduzida. Entretanto, a maior parte das pesquisas, apesar de indicarem a necessidade de mudanças em diferentes aspectos das avaliações em larga escala, mantém uma postura hegemônica ao validarem o discurso do poder público na medida em que os argumentos usados

não são significativos a ponto de apresentar alternativas para romper com a visão de responsabilização.

Do grupo de trabalhos analisados, somente Freitas (2007) sustenta crítica às formas de implementar políticas de avaliação baseadas na “teoria da responsabilização”, em curso nos estados e municípios e propõe uma forma alternativa de lidar com o problema, baseada em “qualidade negociada” via avaliação institucional, com envolvimento bilateral do Estado e da escola. Argumenta que as políticas de responsabilização unilaterais conduzem à configuração de escolas para pobres e escolas para ricos, e alerta para o risco de que, mesmo com o discurso da transparência e responsabilidade, os “sistemas” de avaliação externa centralizados na Federação ocultem, em indicadores estatísticos como o IDEB, as dificuldades que as classes populares estão tendo para aprender no interior da escola, legitimando estratégias que somente conduzem ao adiamento da sua exclusão.

O segundo grupo de pesquisas analisa como os exames externos são desdobrados no âmbito da instância estadual ou municipal. O foco é a discussão de como esses testes são tomados como uma das medidas usadas pelo governo para implementar medidas empresariais como solução dos problemas educacionais, na visão da responsabilização e regulam o ensino pela lógica do controle. Os achados de Silva (2006), por exemplo, revelam a inquietação que há entre os educadores no sentido de aguardar medidas mais incisivas e contínuas do governo após a divulgação dos resultados dos exames aplicados em larga escala.

Novamente observa-se que as pesquisas não negam a constituição dos sistemas de avaliação externa em âmbito estadual e municipal, embora alguns tenham críticas quanto à forma, às concepções subjacentes à constituição destes sistemas ou aos efeitos negativos na prática pedagógica, principalmente em decorrência dos diversos ranqueamentos feitos com os resultados dos exames. As conclusões convergem para a assertiva de Lima:

Por sua vez, a avaliação [externa] passa a ser concebida como uma ferramenta estratégica da gestão do sistema para melhoria dos indicadores educacionais, materializada no modelo administrativo de gestão por resultados. Observa-se o uso da avaliação como instrumento indutor da melhoria da educação, explicitado, sobretudo, com a utilização dos resultados das avaliações atrelados à política de incentivo e premiação, com o propósito de motivar os agentes envolvidos no processo educacional. (LIMA, 2011, p. 17-18).

Além disso, é possível perceber que a política de responsabilização (*accountability*), composta por três elementos (medição do desempenho dos estudantes, publicização dos resultados e recompensa ou sanções baseadas em alguma medida de desempenho – pagamento por mérito) tem avançado nos Estados e municípios mais que no âmbito federal; que foi freada na gestão do governo de Luiz Inácio Lula da Silva (FREITAS, 2011) e agora será implementada por meio do Pacto Nacional pela Alfabetização na Idade Certa.

e) Estudos com foco na análise do desempenho dos estudantes, escolas ou redes de ensino em exames externos.

A maioria das pesquisas identificadas (33%) tinha como objetivo analisar os resultados das avaliações externas com foco no desempenho dos estudantes, de escolas ou redes de ensino. Em geral, buscavam compreender as relações entre fatores associados e a proficiência dos estudantes ou os índices (médias) das escolas ou redes de ensino pesquisadas nos exames externos.

A limitação desses estudos está na tentativa de relacionar o desempenho a variáveis isoladas, como gestão, práticas pedagógicas, formação do professor, investimento financeiro dos governos na educação, a municipalização, participação dos pais, política de organização das turmas dentro da escola, nível socioeconômico, desigualdades regionais, infraestrutura (biblioteca, quadra de esportes, laboratório de informática e laboratório de ciências), ou mesmo a elaboração da prova ou forma de condução da aplicação.

É importante destacar que os resultados das pesquisas evidenciavam que, analisando determinadas variáveis, não eram encontradas diferenças significativas para explicar o desempenho tanto dos estudantes quanto das escolas ou redes de ensino. Embora algumas evidências fossem positivas, não eram passíveis de generalizações.

Como mostraram algumas investigações (BAUER, 2011; CASTRO, 2009), um conjunto de variáveis concorre para explicar os resultados dos estudantes em exames externos, como nível socioeconômico, contexto familiar, contexto escolar (atuação dos professores, dos gestores, clima escolar, tempo de permanência na escola), entre outros. Mesmo a variável fator socioeconômico, como destacam Almeida (2009) e Klein, Fontanive e Elliot (2006), somente quando cruzada com outras variáveis apresenta resultados relevantes para explicar as

diferenças de desempenho nos testes padronizados. Da mesma forma, Andrade e Soares (2008) concluíram que indicadores de qualidade de um mesmo sistema educacional podem variar dependendo de decisões tomadas a priori, tais como: mudanças nas linhas de corte para escolha dos diferentes níveis de desempenho ou mudanças na interpretação dos níveis da escala de proficiência.

A pesquisa de Freitas, Baruffi e Real (2011) também se inscreve neste cenário. Os autores trazem constatações de estudo realizadas em redes escolares municipais com resultados positivos no IDEB com vistas a identificar no contexto e na política local fatores que podem ter concorrido para esses resultados. Os autores concluem que, mesmo tendo elencado dez indicadores que podem ter contribuído para explicar quais fatores que melhor explicam os resultados positivos obtidos por diferentes redes de ensino pesquisadas, não é possível defini-los com segurança.

Neste sentido, compactua-se com Freitas (2007) e Hippólyto, Carvalho e Sousa (2011) que as mudanças somente trarão qualidade ao sistema público de ensino quando as ações entre Estado, município e escola forem desenvolvidas em conjunto.

Outro grupo de pesquisas descreve e analisa os resultados de redes de ensino em relação aos indicadores nacionais ou de programas específicos, conduzidos por empresas do âmbito da iniciativa privada para avaliar desempenho nos estados ou municípios. Os autores levantam, de modo geral, a hipótese de que a melhoria observada nos desempenhos escolares em algumas séries avaliadas pode ser reflexo da política educacional, mediante estabelecimento do pacto de compromisso para implementação do Plano de Desenvolvimento da Educação (PDE) e das metas estabelecidas com a criação de indicadores de desempenho – como o Índice de Desenvolvimento da Educação Básica (IDEB), e também da implementação das políticas de incentivo e práticas de *accountability*.

No entanto, os estudos são inconsistentes, não apresentam argumentos sustentados teórica e empiricamente. Além disso, há pesquisas nesta mesma direção, como o de Lima, Pequeno e Melo (2008), que revelam que os resultados das avaliações externas indicam um desempenho insatisfatório em relação à proficiência dos estudantes e ao índice de desempenho das escolas; o que vem sendo explorado recorrentemente na mídia e reforça o esvaziamento da função docente e o enfraquecimento da escola pública brasileira.

Há ainda investigações que utilizaram os resultados para estabelecer comparações. A tese de Almeida (2009), por exemplo, objetivou comparar a eficiência de escolas públicas cicladas e não cicladas do Ensino Fundamental com uso da técnica DEA. A dissertação de Barreto (2009) e a pesquisa de Klein *et al.* (2008) compararam o desempenho de estudantes de um grupo de escolas privadas que tem um sistema próprio de avaliação externa com o desempenho das escolas privadas no SAEB. A dissertação de Costâncio (2009) mostrou diferenças de desempenho em exames externos – SAEB, Programa de Avaliação da Rede Pública de Educação Básica (Proeb) e Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul (Saers) –, entre alunos declarados pardos ou brancos e os que se consideram negros. O estudo de Mandelert e Mello (2009) comparou o risco de atraso escolar no Brasil com outros três países latinoamericanos, o México, a Argentina e a Colômbia, considerando o nível socioeconômico e a dependência administrativa da escola, controlando os resultados pelo desempenho em leitura dos estudantes utilizando os dados do *Programme for International Student Assessment (Pisa)*. Graça (2009) realizou uma análise comparativa dos resultados obtidos pela rede municipal de Aracaju na Prova Brasil e no IDEB nos anos de 2005 e 2007.

Os estudos comparativos conseguem perceber diferenças mais significativas, mas não são passíveis de generalizações por focalizarem contextos específicos. Além disso, a maioria apresenta problemas metodológicos, principalmente quando compararam grupos com características diferenciadas, específicas, e que acabam não sendo consideradas.

f) Usos, impactos e implicações dos resultados das avaliações externas

Um pequeno número de pesquisas (11%) teve por objetivo analisar os usos, impactos e implicações dos resultados das avaliações externas. De modo geral, os estudos encontrados com esse fim buscaram investigar as repercussões da avaliação externa (práticas pedagógicas, práticas de gestão, metodologias de ensino, escolha de livros didáticos, planejamento de ações ou reformulações curriculares) e compreender o processo de apropriação da política de avaliação da Educação Básica por professores ou gestores das escolas e das redes de ensino.

Estas pesquisas revelaram que os professores e os gestores se sentem pressionados pelos resultados dos exames externos e se observa um estreitamento curricular das disciplinas avaliadas, que passam a priorizar os conteúdos que compõem as matrizes de referência dos testes. Marçal e Carrijo (2011) apontam que tais fatos são decorrentes principalmente das políticas de meritocracia implementadas por alguns Estados. Com efeito, a avaliação formativa vem perdendo espaço para práticas de avaliação classificatórias, com foco em resultados aferidos de forma objetiva.

Desataca-se que a apropriação dos resultados, como revelou a pesquisa de Silva, (2011), acontece em maior ou menor grau, dependendo da escola. Os resultados dos exames externos geram transtornos, pressões e constrangimentos, com conseqüente interferência nos aspectos administrativo, pedagógico e financeiro da escola.

Além disso, alguns gestores, como evidenciou o estudo de Neto (2006) realizado com os gestores da rede pública de ensino do Distrito Federal, conhecem pouco os resultados e não compreendem as informações produzidas por avaliações externas de forma suficiente, o que faz com que não haja uma atividade sistemática de planejamento de suas ações.

Em síntese, estas pesquisas mostram que as avaliações externas estão mudando a rotina escolar e servem de referência para o direcionamento do currículo. Contudo, os dados gerados não estão sendo devidamente aproveitados pelos professores e gestores para o planejamento de ações que possam, de fato, melhorar o ensino.

4 Considerações Finais

As pesquisas analisadas evidenciaram a tendência da adoção de exames externos de rendimento dos estudantes, nos moldes conduzidos pelo governo federal brasileiro, como parâmetro nuclear da gestão da educação nos sistemas estaduais e municipais. A emergência da regulação avaliativa na Educação Básica tem proporcionado um novo paradigma de gestão educacional e, gradativamente, difunde uma nova “cultura de avaliação” na rede de escolas. O discurso permanece atrelado à falsa ideia de que o investimento em exames externos, na perspectiva da responsabilização, gera a tão desejada qualidade do sistema educacional brasileiro.

Os trabalhos que, em sua maioria, apresentam resultados de pesquisas empíricas, ou tecem críticas à forma como os sistemas de avaliação em larga escala são constituídos e implementados, bem como aos usos e implicações dos resultados das avaliações externas no campo educacional são importantes, mas não são propositivos a ponto de romper com o desenho que atualmente é utilizado. De certa maneira, há um consenso que avaliação externa gera qualidade na medida em que, a partir de seus resultados, índices e metas são estabelecidos. Portanto, observando o cenário das políticas públicas educacionais do país, é possível afirmar que a máxima “mais avaliação, mais qualidade” vem sendo disseminada, validando o discurso hegemônico do Estado.

A ressalva está em alguns dos trabalhos que propõem ou analisam instrumentos e metodologias para avaliação em larga escala. Estas pesquisas reforçam a importância de as avaliações externas considerarem variáveis de valor agregado, como a influência de fatores demográficos, sociológicos e relativos à trajetória escolar e fatores socioeconômicos. Mas são poucos os estudos que apresentam modelos alternativos que podem ser utilizados para elaboração de políticas públicas mais efetivas para o melhoramento do sistema educacional brasileiro.

Em geral, as pesquisas têm caráter descritivo e estão mais no campo da crítica, sendo pouco propositivas, embora revelem diferentes saberes que precisam ser considerados para o desenvolvimento do campo da avaliação. Faltam pesquisas que ofereçam análises pontuais acerca das implicações das avaliações externas, com foco não somente no trabalho do professor ou do gestor, para uma análise que considere a organização do trabalho pedagógico.

Concordamos que não mais é possível que a avaliação se limite ao nível micro da sala de aula, o que não significa que devemos aceitar acriticamente os resultados dos exames externos. Ao contrário, precisamos dialogar entre os diferentes níveis da avaliação, buscando a compreensão de como as políticas públicas influenciam a organização do trabalho pedagógico e participando de forma efetiva dos processos de avaliação institucional, por meio de posturas colaborativas e crítico-reflexivas.

Referências

AGUIAR, G. da S. O funcionamento diferencial do item (DIF) como estratégia para captar ênfases curriculares. In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

ALMEIDA, I. B. P. de 2009. **Análise do desempenho de escolas públicas cicladas e não cicladas pertencentes ao ensino fundamental**. Tese de Doutorado, Faculdade de Educação. Unicamp: São Paulo, 2009. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=000470700>>. Acesso em: 19 nov. 2011.

ANDRADE, J. M.; LAROS, J. A. Construção de um modelo explicativo de desempenho escolar: um estudo multinível com dados do Saeb. In Abave, II Reunião Anual, 2006. **Anais eletrônicos...** Abave, 2006. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 16 out. 2011.

ANDRADE, R. J.; SOARES, J. F. O problema do efeito da escola. In Abave, IV Reunião Anual, 2008, Rio de Janeiro, RJ. **Anais eletrônicos...** Rio de Janeiro – RJ: Abave, 2008. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 18 out. 2011.

BARRETO, E. S. de S. Avaliação na educação básica: entre dois modelos. **Revista Educ. e Soc.** Campinas, ano XXII, n. 75, p. 49-66, ago. 2001.

BARRETO, H. P. D. **A avaliação em larga escala no Brasil: análise comparativa entre o SAEB e um sistema privado**. Dissertação (Mestrado em Educação). UCB, 2009. Disponível em: <http://www.bdtd.ucb.br/tede/tde_busca/arquivo.php?codArquivo=1076>. Acesso em: 13 nov. 2011.

BARROSO, J. A regulação da educação. In: _____. **Políticas educativas e organização escolar**. Lisboa: Universidade Aberta, 2005. p. 63-92.

BAUER, A. Percalços da elaboração de propostas de avaliação de impacto no Brasil: é possível avaliar impactos de programas de formação docente? In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

_____. Avaliação de impacto de programas de formação docente em serviço: uma proposta metodológica para o Estado. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

BONAMINO, Alicia; SOUSA, Sandra Zákia. Três gerações de avaliação da educação básica: interfaces com o currículo da/na escola. **Educação e pesquisa**, v. 38, n. 2, abr./jun. 2012. Disponível em: <<http://dx.doi.org/10.1590/S1517-97022012005000006>>. Acesso em: 26/03/2013.

BORGES, R. M.; CALDERON, Adolfo Ignacio. Avaliação na Educação Básica. Mapeamento da produção científica disseminada na Revista Ensaio: Avaliação e Políticas Públicas em Educação (1999-2008). In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

BOURDIEU, P.; CHAMPAGNE, P. Os excluídos do interior. In: NOGUEIRA, Maria Alice; CATANI, Afrânio (Org.). **Escritos de Educação**. Petrópolis, RJ: Vozes, 1998.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**: texto constitucional promulgado em 5 de outubro de 1988, com as alterações adotadas pelas Emendas Constitucionais nos 1/92 a 3/2006 e pelas Emendas Constitucionais de revisão nos 1 a 6/94. Brasília: Senado Federal, 2007.

CARLINI, A. L.; VIEIRA, M. C. T. SAEB, ENEM, ENC (Provão). Por que foram criados? Que pensar de seus resultados? In: CAPPELLETTI, I. F. **Análise crítica das políticas públicas de avaliação**. São Paulo: Ed. Articulação Universidade/Escola, 2005.

CASSETTARI, N. Pagamento por Performance na Educação Básica. In: ANPED, 31ª Reunião Anual, 2008, Caxambu. **Anais eletrônicos...** Caxambu: ANPED, 2008. Disponível em: <<http://www.anped.org.br/reunioes/31ra/1trabalho/GT05-4496-Int.pdf>>. Acesso em: 12 set. 2011.

_____. O uso das avaliações para diferenciação salarial de professores. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

CASTRO, R. Eficácia e Equidade em Escolas Públicas da Bahia. In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

COELHO, M. I de M. Vinte anos de avaliação da educação básica no Brasil: aprendizagens e desafios. **Revista Ensaio: Avaliação e Políticas Públicas em Educação**, 2008. Disponível em: <<http://www.scielo.br/pdf/ensaio/v16n59/v16n59a05.pdf>>. Acesso em: 12 jul. 2010.

CONVIC, A. N.; TAVARES, C. Z.; MACHADO, M. A. J. As avaliações externas: uma percepção dos educadores In: CAPPELLETTI, I. F. **Análise crítica das políticas públicas de avaliação**. São Paulo: Ed. Articulação Universidade/Escola, 2005.

COSTÂNCIO, J. **O negro e as políticas de ação afirmativa no contexto da desigualdade educacional**. Dissertação (Mestrado em Educação). Universidade Federal de Juiz de Fora: Juiz de Fora, 2009. Disponível em: <<http://www.ufjf.br/ppge/files/2010/04/disserta%C3%A7%C3%A3o-Julimar.pdf>>. Acesso em: 13 nov. 2011.

COSTA, S. de A. O valor agregado em leitura como indicador de qualidade das escolas municipais de Belo Horizonte. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

CREPALDI, M. de L., dos SANTOS, R. L. L.; GALINDO, R. C. Políticas de avaliação educacional sob a ótica de professores. In: CAPPELLETTI, I. F. **Análise crítica das políticas públicas de avaliação**. São Paulo: Ed. Articulação Universidade/Escola, 2005.

DIAS SOBRINHO, J. A Avaliação Institucional da Educação Superior: fontes externas e fontes internas. **Revista Avaliação**, Campinas, v. 3, n. 4, p. 29-35, 1998.

FACHINETTO, J. L.; LIMA, M. A. Avaliação Institucional como instrumento de verificação de qualidade: o que dizem as teses e dissertação de pós-graduandos em Educação no Brasil. In: ANPED, 31ª Reunião Anual, 2008, Caxambu. **Anais eletrônicos...** Caxambu: ANPED, 2008. Disponível em: <<http://www.anped.org.br/reunioes/31ra/1trabalho/GT05-4496-Int.pdf>>. Acesso em: 12 set. 2011.

FIGUEIREDO, I. Z. Os projetos financiados pelo banco mundial para o ensino fundamental no Brasil. **Educ. Soc.** [online]. 2009, v.30, n.109, p. 1123-1138. Disponível em: <<http://dx.doi.org/10.1590/S0101-73302009000400010>>. Acesso em: 21 out. 2011.

FREITAS, D.; BARUFFI, A.; REAL. Resultados positivos do IDEB em redes escolares municipais: evidências de estudo exploratório. In: ANPED, 34ª Reunião Anual, 2011, Natal. **Anais eletrônicos...** Natal – RN: ANPED, 2011.

FREITAS, L. C. de. Eliminação adiada: o caso das classes populares no interior da escola e a ocultação da (má) qualidade do ensino. **Educação & Sociedade**, CEDES, v. 28, n. 100, p. 965-987, 2007.

_____. Responsabilização, meritocracia e privatização: conseguiremos escapar ao neotecnicismo? Texto apresentado no III Seminário de Educação Brasileira promovido pelo CEDES no Simpósio PNE: Diretrizes para avaliação e regulação da Educação Nacional, 2011. Versão sujeita a modificações. Disponível em: <http://www.cedes.unicamp.br/seminario3/luiz_freitas.pdf>. Acesso em: 4 mai. 2011.

GENTILI, Pablo. O discurso da qualidade como nova retórica conservadora no campo educacional. In: SILVA, T. T. **Neoliberalismo, Qualidade Total e Educação**. Petrópolis, RJ: Vozes, 1997.

GONÇALVES, L. de F. A intensificação da avaliação educacional em larga escala, na Educação Básica, a partir da década de 90, e suas relações com as reformas educacionais. In Abave, II Reunião Anual, 2006. **Anais eletrônicos...** Abave, 2006. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 16 out. 2011.

GRAÇA, H. C. da. No espelho da avaliação externa: Análise comparativa dos resultados da Prova Brasil e do IDEB da Rede Municipal de Ensino de Aracaju. In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

HIPPÓLYTO, L de Q; CARVALHO, F. E. D de; SOUSA, R. T de. Avaliação dos resultados obtidos no Spaece, pelos alunos da 3ª série do Ensino Médio, em Matemática. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

KLEIN, R. et al. O desempenho dos alunos da Fundação Bradesco: uma comparação com os resultados do SAEB. In Abave, IV Reunião Anual, 2008, Rio de Janeiro – RJ. **Anais eletrônicos...** Rio de Janeiro – RJ: Abave, 2008. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 18 out. 2011.

KLEIN, R.; FONTANIVE, N. S.; ELLIOT, L. G. O Exame Nacional do Ensino Médio: tecnologia e principais resultados em 2005. In Abave, II Reunião Anual, 2006. **Anais eletrônicos...** Abave, 2006. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 16 out. 2011.

LIMA, A. C.; PEQUENO, M. I. C.; MELO, M. N. R. Avaliação da Alfabetização no Ceará: Principais Resultados da primeira edição do Spaece-alfa. In Abave, IV Reunião Anual, 2008, Rio de Janeiro – RJ. **Anais eletrônicos...** Rio de Janeiro – RJ: Abave, 2008. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 18 out. 2011.

LIMA, A. C. Política de avaliação Educacional do Ceará no período 2007 a 2010: Principais resultados. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

MANDELERT, D.; MELLO, J. C. R. S. A defasagem escolar em camadas altas no Pisa 2006. In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

MARÇAL, M de P. V.; CARRIJO, V. L. A. Avaliações do Simave/Proeb: uma análise de suas repercussões na formação de professores de Língua Portuguesa. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

MARQUES, E. de M. R.; [et al]. Provinha Brasil de Matemática: Primeiras considerações sobre a Aplicação Piloto. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

MIRANDA, A. C. **O desafio da construção de referências de qualidade para os sistemas de ensino: uma avaliação com o uso de análise de envoltória de dados – DEA.** Tese (Doutorado em Educação). Unicamp: São Paulo, 2008. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000436607>>. Acesso em: 19 nov. 2011.

MOREIRA, R. dos S. M.; SORDI, M. R. L. de. Avaliação externa como instrumento da gestão do sistema de ensino: a adesão e os impasses para a busca de melhoria na educação. In: ANPED, 27ª Reunião Anual, 2004, Caxambu. **Anais eletrônicos...** Caxambu:

ANPED, 2004. Disponível em: <<http://www.anped.org.br/reunioes/27/gt05/p053.pdf>>. Acesso em: 07 set. 2011.

NETO, J. L. H. Um olhar retrospectivo da Avaliação Externa no Brasil: das primeiras medições em educação até o SAEB de 2005. In Abave, II Reunião Anual, 2006. **Anais eletrônicos...** Abave, 2006. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 16 out. 2011.

OLIVEIRA, P. N.; GUIMARÃES, G. L. A abordagem estatística na Provinha Brasil de Matemática. In Abave, VI Reunião Anual, 2011, Fortaleza – CE. **Anais eletrônicos...** Fortaleza – CE: Abave, 2011. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 31 out. 2011.

PEREIRA, M. C.; CALDERANO, M. da A. In Abave, V Reunião Anual, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

PEREIRA, V. R. **Métodos alternativos no critério Brasil para construção de indicadores sócio-econômicos.** Dissertação (Mestrado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro: Rio de Janeiro, 2004. Disponível em: <http://fourier.lambda.ele.puc-rio.br/5253/5253_1.PDF>. Acesso em: 13 nov. 2011.

PERRY, F. A. **Escalas de proficiência:** diferentes abordagens de interpretação na avaliação educacional em larga escala. Dissertação (Mestrado em Educação). Universidade Federal de Juiz de Fora, 2009. Disponível em: <http://www.btdt.ufjf.br/tde_busca/arquivo.php?codArquivo=560>. Acesso em: 13 nov. 2011.

RIOS, T. A. **Compreender e ensinar:** por uma melhor docência da melhor qualidade. São Paulo: Cortez, 2008.

RODRIGUES, E. N. **O funcionamento diferencial do item de língua portuguesa:** análise das causas e conseqüências no contexto do programa Nova Escola – RJ e do PROEB – MG. Dissertação (Mestrado em Educação). Universidade Federal de Juiz de Fora: Juiz de Fora, 2008. Disponível em: <http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=135513>. Acesso em: 13 nov. 2011.

RODRIGUES, S. C. **Construção de uma metodologia alternativa para a avaliação das escolas públicas de ensino fundamental por meio do uso da análise por envoltório de dados (DEA):** uma associação do quantitativo ao qualitativo. Tese

de Doutorado, Faculdade de Educação. Unicamp: São Paulo, 2005. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=vtls000436607>>. Acesso em: 19 nov. 2011.

SA, V. A (auto)avaliação das escolas: “virtudes” e “efeitos colaterais”. **Ensaio: aval.pol.públ.Educ.** [online]. 2009, vol.17, n.62, pp. 87-108. ISSN 0104-4036. Disponível em: <<http://dx.doi.org/10.1590/S0104-40362009000100005>>. Acesso em: 21 out. 2011.

SILVA, M. J de A. O Sistema Mineiro de Avaliação da Educação Pública: impactos na escola fundamental de Uberlândia. In Abave, II Reunião Anual, 2006. **Anais eletrônicos...** Abave, 2006. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 16 out. 2011.

_____. **Regulação educativa:** o uso dos resultados de proficiência das avaliações do PROEB por diretores escolares em Minas Gerais. Tese (Doutorado em Educação). UFMG: Belo Horizonte, 2011. Disponível em: <www.bibliotecadigital.ufmg.br/.../tesemariajulianaalmeidasilva.pdf>. Acesso em: 13 nov. 2011.

SOUZA, E. R. de. **Accountability de professores:** um estudo sobre o efeito da prova Brasil em escolas de Brasília. Dissertação (Mestrado em Educação). Unicamp: São Paulo, 2009. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?code=000476835>>. Acesso em: 19 nov. 2011.

VIEIRA, M. A; TENÓRIO, R. M. Avaliação e sustentabilidade: as quatro gerações de Guba e Lincoln em debate. In Abave, V Reunião Anual da ABAVE, 2009, Salvador – BA. **Anais eletrônicos...** Salvador – BA: Abave, 2009. Disponível em: <<http://www.abave.org.br/>>. Acesso em: 23 out. 2011.

WERLE, Flávia Obino Corrêa. Sistema de Avaliação da Educação Básica no Brasil: abordagem por níveis de segmentação. In: _____. (Org.). **Avaliação em larga escala:** foco na escola. Bsb: Líber, 2010.