

A QUALIDADE DA EDUCAÇÃO E A DISSEMINAÇÃO DE SISTEMAS DE AVALIAÇÃO

FERNANDO TAVARES JÚNIOR
LUIZ FLÁVIO NEUBERT

RESUMO

Nos últimos quinze anos, disseminaram-se no Brasil políticas de avaliação educacional. Os estados exerceram especial protagonismo nesse processo. Desde a Conferência de Jomtien, em 1990, observa-se o mesmo em escala global, concomitante às primeiras experiências em larga escala na América Latina. No Brasil, ocorreram dois movimentos: inicialmente, do governo federal, com avaliações bianuais, amostrais e diagnósticas; posteriormente, avaliações descentralizadas, estaduais, anuais, censitárias e relacionadas à gestão de suas redes, processo que se intensificou na última década. Este trabalho analisa a associação entre a implantação de sistemas estaduais de avaliação educacional no Brasil e o desenvolvimento da proficiência média entre os alunos das respectivas redes estaduais de ensino. Utilizam-se os resultados agregados das redes estaduais nas avaliações nacionais como parâmetro externo, e observa-se associação entre a implantação de avaliações e a elevação da proficiência média nos anos seguintes.

PALAVRAS-CHAVE AVALIAÇÃO DOS SISTEMAS DE ENSINO • SAEB • DESEMPENHO • POLÍTICAS EDUCACIONAIS.

RESUMEN

Los últimos quince años se diseminaron en Brasil políticas de evaluación educativa. Los estados ejercieron un especial protagonismo en tal proceso. Desde la Conferencia de Jomtien en 1990, se observa lo mismo en escala global, de modo concomitante a las primeras experiencias en larga escala en América Latina. En Brasil ocurrieron dos movimientos: inicialmente, del gobierno federal, con evaluaciones bianuales, realizadas por medio de muestras y diagnósticas; posteriormente, evaluaciones descentralizadas, estaduais, anuales, censitarias y relacionadas a la gestión de sus redes, proceso que se intensificó en la última década. Este trabajo analiza la asociación entre la implantación de sistemas estaduais de evaluación educativa en Brasil y el desarrollo de la proficiencia media entre los alumnos de las respectivas redes estaduais de enseñanza. Se utilizan los resultados agregados de las redes estaduais en las evaluaciones nacionales como parámetro externo, y se observa una asociación entre la implantación de evaluaciones y la elevación de la proficiencia media en los años posteriores.

PALABRAS CLAVE EVALUACIÓN DE LOS SISTEMAS DE ENSEÑANZA • SAEB • DESEMPEÑO • POLÍTICAS EDUCATIVAS.

ABSTRACT

In the last fifteen years, educational evaluation policies have been propagated in Brazil. The states have a special role in this process. Since the Jomtien Conference in 1990, the same has been observed on a global scale, concomitant with the first large-scale experiments in Latin America. There were two movements in Brazil: the former, from the federal Government, with biannual, diagnostic and sampling evaluations; the latter, with evaluations which were decentralized, state, annual, census and related to the management of their networks, a process which was intensified in the last decade. This article analyzes the association between the establishment of state systems of educational evaluation in Brazil and the development of the average proficiency among students of the respective state school networks. The aggregate results of the state networks in the national evaluations are used as an external parameter, and the association between the implementation of evaluations and the increase of the average proficiency in the following years is observed.

KEYWORDS EVALUATION OF SCHOOL SYSTEMS • SAEB • PERFORMANCE EDUCATIONAL POLICIES.

INTRODUÇÃO

Entre 1997 e 2011, o Sistema de Avaliação da Educação Básica (Saeb) permitiu o diagnóstico da proficiência média dos estudantes brasileiros da educação básica. Esse período permitiu a aferição de sete ciclos bianuais, cada qual apontando tendências de desenvolvimento da proficiência média em cada um dos estratos investigados: unidades da federação, redes de ensino, etapas dos níveis fundamental e médio etc. Este estudo analisa a relação entre essas distintas tendências de desenvolvimento e a adoção (ou não) de sistemas de avaliação no Brasil. Não é possível investigar com rigor as relações causais entre um fenômeno e outro. No entanto, em caráter exploratório e como indicativo de possíveis efeitos secundários, procura-se analisar a ocorrência mútua da elevação da proficiência média dos alunos das redes estaduais de ensino e da implantação, manutenção ou mesmo interrupção de processos avaliativos.

Assim, o estudo organizou-se em três frentes distintas. Em primeiro lugar, analisaram-se as tendências do desenvolvimento da proficiência média dos alunos avaliados pelo

Saeb. Paralelamente, investigaram-se as múltiplas ocorrências de implantação/adoção de políticas de avaliação educacional, e seu fluxo histórico (manutenção, alterações, interrupção) em cada unidade da federação. Soma-se um terceiro vetor dedicado a pesquisar as possíveis relações entre um fenômeno e outro. O principal objetivo é tentar identificar fatores que possam estar ligados à melhoria da qualidade da educação. Acredita-se também que se aprende com a análise do sucesso escolar, uma vez que sua identificação pode levar à descoberta de redes de nexos causais e processos sócio-históricos até então pouco conhecidos ou mesmo inovadores. Esse vértice está relacionado ao projeto mais amplo de investigação dos “Determinantes do Sucesso Educacional no Brasil”, pesquisa em curso que conta com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), por meio do Programa Observatório da Educação (OBEDUC).

Diversos estudos sobre educação no Brasil dedicam-se a analisar o fracasso escolar. Há ampla literatura que relaciona, a partir de diferentes metodologias e fontes, características sociodemográficas a trajetórias de insucesso, destacando-se as diferenças de classe social, raça e gênero (VALLE SILVA; SOUZA, 1986; HASENBALG; VALLE SILVA, 1990; HASENBALG; VALLE SILVA, 1999; VALLE SILVA, 2003; BARBOSA; RANDALL, 2004; FERNANDES, 2005; MONT'ALVÃO, 2011). Contudo, ainda não são muitos os estudos que se dedicam a identificar os fatores relacionados ao sucesso escolar (BONAMINO; FRANCO, 2005; BARBOSA, 2009; BROOKE; SOARES, 2008; SOARES, 2007). Destaque crescente tem sido dado às políticas de responsabilização (BROOKE, 2006) e à análise da eficácia escolar e seus fatores associados (ESPÓSITO; DAVIS; NUNES, 2000; FRANCO; MANDARINO; ORTIGÃO, 2001; ALBERNAZ; FERREIRA; FRANCO, 2002). Ainda que não seja possível apontar ou negar relações causais diretas, objetiva-se identificar correlações que permitam inferir possíveis redes de nexos já diagnosticadas e analisadas em outros estudos acerca da aprendizagem, do desempenho de alunos e do desenvolvimento educacional.

No caso específico das avaliações de sistemas educacionais, há ampla literatura dedicada a suas características, seus

modelos, seus efeitos e à análise de suas repercussões no cotidiano escolar. Em especial, a literatura internacional voltada para “Reformas Educacionais Baseadas em Padrões”, ou *Standards-based Reforms of Education*, tem evidenciado, desde os anos 1990, a partir da Conferência de Jomtien (ORGANIZAÇÕES DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, CIÊNCIA E CULTURA, 1998), que a adoção de políticas de avaliação de sistemas, com especial atenção às unidades escolares, produz efeitos sobre o desenvolvimento da proficiência média (LOCKHEED; VERSPOOR, 1991; MCLAUGHLIN; SHEPARD, 1995; BYBEE, 1997; ELMORE, 2007). Alguns estudos ao longo do tempo apontaram tanto as potencialidades quanto os limites das reformas empreendidas a partir dos anos 1990 (CHATTERJI, 2002; FALK, 2002; GAMORAN, 2007). Outros estudos utilizaram dados de avaliações nacionais como referência para analisar as reformas de unidades federativas (SWANSON; STEVENSON, 2002), bem como outras evidências empíricas acerca das reformas e sua repercussão na mudança de práticas escolares (SLAVIN, 2002).

Uma hipótese é que a implantação de sistemas de avaliação, ao dar visibilidade aos resultados de aprendizagem dos alunos, afeta a escola de múltiplas formas. Em especial três: (1) produz consequências ao trabalho docente, gerando maior ou menor prestígio à escola; (2) fornece informações relevantes que podem auxiliar no planejamento do trabalho pedagógico e motivar alterações positivas na gestão escolar; (3) cria uma nova linguagem para diálogo com a sociedade, as famílias e o sistema, fomentando a priorização da aprendizagem como principal marca da identidade escolar. Vale ressaltar que há também limites, como o risco de “redução curricular”, mas elogios, cobranças e benefícios diversos (ou sua ausência) tendem a “mexer” com a escola. Essa mudança produziria efeitos positivos sobre a gestão educacional, que acarretariam boas práticas, e essas contribuiriam para o desempenho dos alunos.

Essas reformas ocorreram tardiamente no Brasil. As primeiras iniciativas decorreram do governo federal que, por sua vez, não gere as redes públicas de ensino voltadas para a educação básica. São os estados e municípios os responsáveis

pelas redes de ensino fundamental no Brasil. Todavia, embora suas reformas e avaliações tenham sido tardias, boa parte já neste século, ocorreu rápida disseminação, já merecendo meta-avaliação (FRANCO, 2009; CASTRO, 2009). Concomitante a esse processo, observou-se também a reversão das tendências de queda da proficiência média e uma positiva e crescente tendência de melhoria do desempenho. Quais fatores podem estar associados ao desenvolvimento educacional brasileiro? Enquanto fatores, podem ser elencados projetos, iniciativas, variações conjunturais e, principalmente, políticas públicas. Pesquisas recentes têm se dedicado aos determinantes do desempenho (MENEZES-FILHO, 2007) e também aos efeitos das reformas educacionais na melhoria da qualidade da educação (VELOSO, 2009, 2011).

Seria possível aprender com as experiências de sucesso? Provavelmente, boa parte da melhoria da qualidade da educação é decorrência de boas práticas: experiências de sucesso, iniciativas inovadoras e que alcançam melhores resultados, empreendedorismo educacional e outras realizações que, por vezes, oferecem propostas eficientes, factíveis e testadas para os desafios da educação nacional (LÜCK, 2009). As pesquisas sobre eficácia escolar apontam relação entre as políticas sistêmicas, a gestão das escolas e a melhoria do desempenho (SAMMONS; HILLMAN; MORTIMORE, 1995; SOARES, 2002; UNIVERSIDADE FEDERAL DE MINAS GERAIS, 2002; SAMMONS, 2008). Mais recentemente, houve incremento de pesquisas voltadas para mensuração e análise do “Efeito Escola”, em especial a partir da implantação da Prova Brasil, em 2005, e do Ideb, em 2007 (ALVES; SOARES, 2007; BIONDI; FELÍCIO, 2007; FRANCO et al., 2007; GREMAUD; FELICIO; BIONDI, 2007). Todavia, não se deixou de considerar a relevância das características socioeconômicas da clientela sobre os resultados educacionais (BUCHMANN, 2002; ALVES; SOARES, 2013). De fato, o desempenho escolar é produto de uma série de fatores. Quanto mais próximo ao aluno, mais viável é a identificação desses fatores e a mensuração de seu efeito sobre o desempenho.

Observa-se que tem sido dada pouca atenção relativa às mudanças nas políticas educacionais. Daí o esforço de também

considerá-las na análise geral das tendências de desenvolvimento da proficiência média. Como dito, as variações no desempenho médio podem estar associadas a várias ocorrências, como reformas educacionais, sejam elas parciais ou amplas; alterações contextuais, como melhoria nas condições de vida da clientela; dentre outras. É empiricamente impossível isolar um único fator. Portanto, uma possibilidade é testar associações entre as ocorrências.

O SAEB E A EVOLUÇÃO DA PROFICIÊNCIA MÉDIA

O ponto de partida para tal consideração é o panorama da proficiência média dos alunos brasileiros ao longo das últimas décadas. Optou-se por priorizar o ensino fundamental, em seus dois ciclos, dos alunos atendidos pelas redes estaduais, a partir das mensurações realizadas pelo Saeb. A opção pelas redes estaduais deve-se à possibilidade de investigar as políticas educacionais de cada estado, em especial a adoção de sistemas de avaliação de suas escolas. Isso seria impossível para redes municipais e não faria sentido para as escolas particulares. Observou-se que os avanços não foram lineares nem distribuídos homoganeamente. Notam-se duas tendências distintas: o desempenho médio apresentava-se em acentuada queda no final dos anos 1990; essa tendência suaviza-se no início do século e inverte-se logo depois, apresentando crescimento consistente desde 2001. Esse diagnóstico só foi possível em função da série histórica produzida pelo Saeb.

O Saeb teve início em 1990, por meio de uma avaliação amostral de escolas públicas urbanas, em quatro séries (1^a, 3^a, 5^a e 7^a) do ensino fundamental e testes de Língua Portuguesa, Matemática e Ciências (avaliou-se redação na 5^a e 7^a série). Repetiu-se o formato em 1993. Todavia, seus resultados não eram comparáveis. Em 1995, adotou-se a Teoria da Resposta ao Item (TRI) para elaboração de testes e escalas, mensuração dos resultados e padronização de procedimentos. Passou-se a avaliar os anos finais dos ciclos fundamentais e médio. A partir de 2001, apenas avaliou-se Língua Portuguesa e Matemática. Para efeitos desta investigação, optou-se por considerar os

resultados médios dos alunos das escolas estaduais no ensino fundamental entre 1997 e 2011. Os dados foram obtidos por meio de consultas ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), responsável pelo Saeb.

As médias de proficiência em Língua Portuguesa ao final do primeiro ciclo fundamental (4ª série/5º ano) dos alunos das escolas estaduais brasileiras declinaram entre 1995 (187,9) e 2001 (163,3). Após esse período, apresentou tendência ascendente, atingindo 190,6 em 2011.¹ Algo similar aconteceu em Matemática: queda de 189,3 (1995) para 175,2 (2001) e posterior crescimento (209,8, em 2011). No segundo ciclo fundamental, o declínio foi mais longo e o progresso, mais curto e rápido. Em Língua Portuguesa, ao final do ciclo fundamental (8ª série/9º ano), as médias de proficiência dos alunos das escolas estaduais brasileiras declinaram entre 1995 (252,4) e 2005 (226,6), melhorando posteriormente (238,7, em 2011), ainda sem atingir os níveis anteriores. Ocorreu algo semelhante em Matemática: queda de 246,6 (1995) para 232,9 (2005), com melhoria posterior (244,7, em 2011).

¹ Há resultados preliminares para a edição de 2013. Todavia, ainda cabem recursos e consolidações. Optou-se por utilizar os resultados até 2011.

OS SISTEMAS ESTADUAIS DE AVALIAÇÃO

A crescente implantação de sistemas estaduais de avaliação da educação neste século atraiu a atenção de pesquisas recentes. A mais importante foi conduzida por Brooke, Cunha e Faleiros (FUNDAÇÃO VICTOR CIVITA, 2011) e produziu um diagnóstico desse processo no Brasil, instruindo e motivando vários outros trabalhos, inclusive este. Portanto, para efeitos deste estudo, a principal referência utilizada foi o citado relatório. Somou-se a ele um largo esforço² de busca de dados sobre as avaliações, seu histórico e suas características. Para cada avaliação e fonte de informação, completava-se uma ficha (semelhante à apresentada no Anexo I), que depois se somava a outras fichas para a consolidação das informações. Foram encontrados muitos dados incongruentes. Em todas as situações buscou-se triangular as fontes e aferir a informação mais fidedigna.

Embora quase sempre adotassem a mesma escala de proficiência e processo de mensuração de resultados (Escala

² Este trabalho contou com o valioso apoio das bolsistas Daniele Pires, Evelize Moreira e Tuane Gonçalves, sob a supervisão da mestranda Liliane Mendonça. Soma-se ainda a colaboração dos bolsistas Emílio Clemente e Marcos Alves nas etapas iniciais do trabalho. A equipe de pesquisa agradece aos bolsistas e à Capes, por meio do Observatório da Educação, pelo apoio para a realização do projeto, sem o qual este não seria possível.

SAEB equalizada pela TRI), diferentes das avaliações nacionais, os sistemas estaduais caracterizaram-se majoritariamente por edições anuais, censitárias e com foco na gestão de suas redes, ou seja, com capacidade de prover informações sobre as unidades escolares e instruir seu diálogo com a gestão central (Secretarias de Educação).

Ceará, Minas Gerais e São Paulo podem ser considerados os estados que melhor desenvolveram sistemas próprios de avaliação da educação, embora com características e percursos significativamente distintos. No entanto, em geral os estados caracterizam-se por experiências muito plurais: adoção, alterações de modelo, interrupções, nova implementação etc. São três as principais ocorrências que interessam mais a este estudo: (1) a primeira implementação de um sistema de avaliação da rede estadual de ensino fundamental; (2) a opção por manter a avaliação e continuar o ciclo de testes nos anos seguintes; (3) a interrupção do sistema. A análise, portanto, dirige-se às seguintes questões:

- a. *A implantação de um sistema de avaliação pode estar associada a uma alteração positiva no desempenho médio dos alunos aferido na edição seguinte do Saeb, acima da média das unidades da federação (UFs) que não implantaram?*
- b. *A manutenção da avaliação produziria efeitos duradouros de incremento no desempenho médio?*
- c. *A interrupção das avaliações estaria relacionada à queda do desempenho dos alunos, pelo menos ao nível similar às UFs que não implantaram?*

Se a resposta às três perguntas for positiva, imagina-se que há indícios de que a implantação de sistemas de avaliação pode estar relacionada à transformação de práticas no cotidiano escolar, com priorização da aprendizagem e melhoria geral do desempenho. Espera-se que o primeiro ciclo avaliativo tenha maior impacto que os demais, uma vez que representa maior inovação. De igual forma, espera-se que a continuidade das avaliações represente um reforço nas boas práticas dedicadas à melhoria da qualidade e que a interrupção das avaliações, ainda mais se precoce, esteja associada à

inibição de tais práticas e, portanto, à queda das tendências de progresso do desempenho médio.

Segue uma breve sinopse com as principais observações a respeito de cada UF e suas opções em relação à avaliação educacional. De maneira geral, mesmo que pontualmente, ao longo da última década, a maioria das UFs apresentaram iniciativas de implantação ou manutenção de sistemas de avaliação educacional em suas redes de ensino. As principais exceções são Santa Catarina, Amapá, Mato Grosso, Pará, Paraíba, Rondônia, Roraima e Rio Grande do Norte. Como dito, Minas Gerais, São Paulo e Ceará destacam-se com experiências mais longevas de avaliação.

Em 1992, o Ceará implantou o Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece). Ainda incipiente e precário, foi um sistema limitado, mas que denotou preocupação com o tema. A partir de 2001, o Spaece-NET (Sistema Permanente de Avaliação da Educação Básica do Ceará Via Internet) foi implantado e mantido desde então. Houve aprofundamento da avaliação, com iniciativas voltadas também para alfabetização. A manutenção ininterrupta da avaliação educacional em larga escala apresenta correlação com a tendência ascendente na curva de desempenho médio da rede estadual, observadas nas edições do Saeb. Minas Gerais implementou o Programa de Avaliação do Estado de Minas Gerais (AVA-MG), em 1992. Em 2000, adotou o Sistema Mineiro de Avaliação (Simave) que atua conjuntamente com o Programa de Avaliação da Rede Pública de Educação Básica (Proeb), sem apresentar interrupções. Implantou também, em 2005, o Programa de Avaliação da Alfabetização (Proalfa), que atua de forma complementar ao Simave, mas com escala própria. A continuidade do Simave apresenta certa relação com a curva positiva de desempenho médio da rede estadual. O Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp) iniciou-se em 1995 e foi mantido desde então. Tal como em Minas Gerais, a continuidade do Saresp também aparece associada a curvas de melhoria de desempenho superiores à média nacional.

As demais UFs registram experiências idiossincráticas. Em 2009, o Acre implantou o Sistema Estadual de Avaliação

da Aprendizagem Escolar (Seape) e o manteve desde então. A aferição do Saeb, em 2011, consecutiva à adoção do Seape, apresentou queda na curva de desempenho de Língua Portuguesa e crescimento na curva de desempenho de Matemática. No entanto, o Seape é incipiente e não permite análises mais aprofundadas. Algo parecido acontece com o Maranhão. Esse estado implantou em 2010 o Sistema Maranhense de Avaliação de Desempenho (Simade). Por ser muito incipiente, o contexto também não apresentou tendências claras, o que não proporcionou condições adequadas para análise das curvas de desempenho médio. Por fim, Tocantins avaliou pontualmente sua rede de ensino pela Avaliação de Escolas Públicas (AEP) e implantou posteriormente, em 2010, o Sistema de Avaliação de Desempenho (SAD). Para os objetivos deste estudo, pode ser considerado um processo também incipiente. A edição de 2011 do Saeb, logo após a adoção do SAD, registrou curvas ascendentes no desempenho médio na maioria dos conteúdos e etapas. A exceção foi Língua Portuguesa no 1º ciclo.

Sergipe realizou avaliações de 2004 a 2006: o Exame de Avaliação da Educação Básica do Estado de Sergipe (Exaeb-SE). A edição de 2005 do Saeb, logo após a adoção do Exaeb-SE, apresenta curvas ascendentes. A continuidade desse exame no período é concomitante ao crescimento do desempenho médio no 1º ciclo, e sua interrupção apresenta relação com a alteração dessa tendência. A edição de 2007 do Saeb registra declive na curva de desempenho. Alagoas adotou o Sistema de Avaliação de Alagoas (Saveal) em 2001, no entanto houve descontinuidade. O Saveal retornou em 2005, mas também sofreu interrupções, sendo retomado em 2011. As edições do Saeb posteriores à aplicação do Saveal apresentam em geral curvas ascendentes, mas há casos de estagnação e declive. A descontinuidade da política de avaliação externa relaciona-se com a alteração nas tendências das curvas, apontando vetores negativos. A partir de 2008, o Amazonas implantou o Sistema de Avaliação de Desempenho Educacional do Amazonas (Sadeam). A partir da adoção do Sadeam, a curva do desempenho médio mostrou-se ascendente para ambos os conteúdos.

A Bahia também foi um dos primeiros estados a adotar políticas de avaliação externa. Em 2001, implantou o Projeto de Avaliação Externa (PAE), que foi interrompido apenas em 2004. Esse período apresenta curva ascendente de desempenho médio da rede, com uma única exceção: Matemática da 8ª série. A interrupção em 2004 apresenta certa relação com a alteração da curva de desempenho, aferida a partir do Saeb de 2005, com tendência de declive. Um caso parecido foi o Espírito Santo, que, em 2000, implantou o Programa de Avaliação da Educação Básica do Espírito Santo (Paebes), mas houve interrupções até 2007. A partir de 2008 houve continuidade da política. Há relação entre as interrupções do Paebes, entre 2000 e 2007, e oscilações frequentes nas tendências das curvas de desempenho médio da rede estadual. Com a continuação do Paebes a partir de 2008, a curva do desempenho médio dos alunos em ambos os conteúdos e ciclos apresentou crescimento, aferida na edição do Saeb de 2009.

O Sistema de Avaliação de Desempenho das Instituições Educacionais do Distrito Federal (Siade) foi implantado em 2008 e mantido desde então. Observam-se curvas ascendentes de desempenho médio em ambos os conteúdos e ciclos, desde a aplicação da Prova Brasil de 2005. A edição de 2009 do Saeb aponta tendência positiva, com declive somente para o 2º ciclo em Matemática. Nesse caso, parece que a redundância das avaliações não provocou um efeito nulo, mas, sim, uma interação positiva. Essa é uma evidência de algo também encontrado em outras avaliações. A percepção de que “quanto mais prova, melhor” (Betti, 2011) parece encontrar nas experiências estaduais indícios positivos de complementaridade, em detrimento de repetição, redundância ou efeito nulo.

Em 2001, Goiás implantou o Sistema de Avaliação da Educação de Goiás (Saego), mas interrompeu em 2003, retomando em 2004. Após a adoção do Saego, o desempenho médio da rede estadual apresentou crescimento significativo. A descontinuidade desse sistema apresenta relação com a descontinuidade dessa tendência, com resultados difusos no Saeb de 2005. Em 2003, o Sistema de Avaliação da Educação Básica do Mato Grosso do Sul (Saems) foi implantado, mas

também sofreu interrupções em 2004, 2006 e 2010. A adoção do Saems estava relacionada à curva ascendente de desempenho, e sua descontinuidade apresenta concomitância com alterações nas tendências das curvas e alguns declives no desempenho médio da rede estadual.

O Programa de Avaliação do Sistema Educacional do Paraná (AVA) foi aplicado entre 1995 e 2002, apresentando interrupções em 1999 e 2001. Suas características também diferem substancialmente de outros programas de avaliação, o que dificulta sua comparabilidade. Após a implementação do AVA, a edição seguinte do Saeb manifestou curva ascendente para ambos os ciclos, tanto para Língua Portuguesa quanto para Matemática. Em 2003, uma nova interrupção do AVA apresentou relação com a alteração na tendência da curva de desempenho médio dos estudantes do 2º ciclo em Matemática. O Piauí instituiu a Prova Piauí somente em 2008 e de forma pontual. Suas características diferem de outros sistemas. Entre 1997 e 2005, o estado apresentou curva oscilante com tendência a declive no desempenho médio educacional dos alunos do 1º e 2º ciclos. A alteração na curva aconteceu após a implantação da Prova Brasil em 2005, com tendência de melhoria do desempenho educacional a partir da edição de 2007 do Saeb, em ambas as disciplinas e ambos os ciclos.

O Sistema de Avaliação Educacional de Pernambuco (Saepe) foi implantado em 2000, mas também apresentou formatos diferentes e múltiplas interrupções. A edição do Saeb subsequente à adoção da primeira edição do Saepe registrou curva em declive para ambos os ciclos nas disciplinas de Língua Portuguesa e Matemática. Observa-se, por outro lado, que o contrário ocorreu na aferição do Saeb subsequente à edição de 2002 do Saepe, quando observam-se curvas ascendentes para ambos os ciclos e conteúdos. Esse é um indício de que não é qualquer modelo de sistema que apresenta efeitos positivos. Políticas de avaliação devem ter características mínimas comuns para que gerem os efeitos esperados. A descontinuidade também apresentou efeitos perversos. A não realização do Saepe em 2003 e 2004 apresenta certa relação com o declínio das curvas de desempenho médio. Contudo, a continuidade posterior do Saepe apresenta efeito positivo

no desempenho, em ambos os conteúdos e ciclos, aferidos nas edições Saeb a partir do ano de 2007. Exceção apenas para Língua Portuguesa, 2º ciclo, que apresentou declínio na edição do Saeb de 2011.

O Rio Grande do Sul implantou em 2005 a primeira versão do Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul (Saers). Os anos seguintes foram marcados por descontinuidade. Sua rede estadual apresentou curva oscilante com tendência à declive em relação a outras UFs. Como foi concomitante à realização da Prova Brasil, há um efeito interativo, impossível de ser isolado. Em 2007, edição subsequente à criação do Saers, observa-se uma curva de desempenho ascendente em ambas as disciplinas do 1º e 2º ciclos. A interrupção do Saers nos anos seguintes teve pouco efeito nessa tendência. É um modelo bastante questionado, principalmente pelo sindicato docente. Modificações em seu formato são esperadas, sem condições ainda de serem avaliadas.

O Rio de Janeiro também teve diferentes modelos. O Sistema de Avaliação da Educação do Estado do Rio de Janeiro (Saerj), implantado em 2008, apresenta estabilidade e continuidade. O estado ainda possui o Sistema de Avaliação Bimestral do Processo Ensino-Aprendizagem (Saerjinho). A curva de desempenho médio de sua rede estadual se apresentou oscilante até 2007. A edição de 2009 do Saeb, posterior à adoção desse modelo do Saerj, registra curvas ascendentes, tanto em Língua Portuguesa quanto em Matemática, em ambos os ciclos. A continuidade da política de avaliação educacional no Rio de Janeiro apresenta relação com a tendência ascendente de suas curvas de desempenho.

METODOLOGIA

Como princípio, pressupõe-se que com base na variação dos resultados do Saeb entre 1997 e 2011, para cada unidade federativa do Brasil, é possível determinar a importância da implantação, manutenção ou interrupção de sistemas estaduais de avaliação da educação básica sobre o nível de proficiência entre estudantes da 4ª série/5º ano e 8ª série/9º ano do ensino fundamental.

Para elaborar a tabulação de dados que será apresentada em seguida, foram utilizados dados do Saeb e informações sobre políticas de avaliação em cada unidade federativa. Essa sinopse contribui para identificar sistemas estaduais de avaliação da educação presentes em cada unidade federativa brasileira. A sinopse foi organizada em duas partes: a primeira inclui a reunião e a organização das informações sobre sistemas estaduais de avaliação da educação, o que resultou na produção de variáveis dicotômicas para cada fenômeno. A segunda parte é dedicada às mudanças nas políticas de avaliação da educação, em especial a criação e aplicação da Prova Brasil em 2005 e a proposição do Ideb em 2007 e sua divulgação no ano seguinte. Já do Saeb foram utilizados os dados relativos ao desempenho médio dos alunos das redes estaduais de ensino nos anos finais dos ciclos fundamentais (4^a série/5^o ano e 8^a série/9^o ano).

O Plano Nacional da Educação (PNE), mesmo ao destacar o crescimento das avaliações de larga escala no sistema educacional brasileiro, apresenta uma escassez de informações a respeito do desenvolvimento e da consolidação de sistemas de avaliação em âmbito estadual e municipal. A urgência de se focar em sistemas estaduais é, nesse sentido, imediata: as unidades federativas e municípios, de forma desigual, têm investido em sistemas próprios de avaliação da educação e, com isso, demonstram uma tendência que perdurará no futuro próximo.

FONTES DE DADOS

A tabela primária de dados foi elaborada com base nos resultados do Saeb entre os anos de 1997 e 2011, os quais, por sua vez, foram relacionados com informações sobre os sistemas estaduais. Foram incluídas as 27 unidades federativas brasileiras, levando-se em conta sete biênios: 1997/1999; 1999/2001; 2001/2003; 2003/2005; 2005/2007; 2007/2009 e 2009/2011. A organização final das informações resultou em uma tabela com 189 casos (linhas da tabela) e 11 variáveis (colunas da tabela). Foram cumpridas, portanto, as seguintes etapas de seleção e organização de informações para a construção da Tabela 1:

- a. Realizou-se o cálculo da diferença entre os resultados do Saeb para cada biênio, os deltas da tabela. Para tanto, subtraiu-se o escore médio final do escore médio inicial, relativos a cada biênio. Isso foi realizado para cada unidade federativa, cada biênio de referência, cada ciclo (4ª série/5º ano ou 8ª série/9º ano) e conteúdo (se Matemática ou Língua Portuguesa). Verificou-se a ocorrência de deltas positivos e negativos, sendo que o primeiro grupo indica crescimento da proficiência e o segundo, declínio da proficiência em determinada unidade federativa, disciplina, ano/série e biênio. O cálculo do delta é dado por:

$$\Delta = P_{\text{final}} - P_{\text{inicial}}$$

Onde:

Δ = Delta;

P_{final} = Proficiência média no ano final do biênio;

P_{inicial} = Proficiência média no ano inicial do biênio.

- b. Identificação do *status* de adoção de sistemas estaduais de avaliação da educação nas unidades federativas nos biênios.
- c. Registro da realização da Prova Brasil e divulgação do Ideb para cada biênio.

Dado o excessivo espaço que a Tabela 1 ocuparia no corpo do texto, apenas uma parte dela foi reproduzida a seguir, juntamente de explicações sobre seus elementos constitutivos, em seguida.

TABELA 1 – Dados sobre variação de desempenho dos alunos no Saeb por série e disciplina nas unidades federativas brasileiras; sistemas estaduais de avaliação; Prova Brasil e Ideb, no período de 1997 a 2011

N.	UF	Biênio	Variação do desempenho médio no Saeb				Sistema estadual de avaliação da educação			Prova Brasil	Ideb
			Δ LP 4 ^o /5 ^o	Δ LP 8 ^o /9 ^o	Δ M 4 ^o /5 ^o	Δ M 8 ^o /9 ^o	Implantou	Manteve	Interrompeu		
1	Rondônia	97-99	-10,2	-22,7	-7,2	-4,5	0	0	0	0	0
2	Acre	97-99	-7,7	-14,2	-2,9	-1,2	0	0	0	0	0
3	Amazonas	97-99	-5,8	-9,7	-0,3	3,3	0	0	0	0	0
4	Roraima	97-99	1,3	-6,8	1,6	8,1	0	0	0	0	0
5	Pará	97-99	-10,2	-17,2	1,7	-2,8	0	0	0	0	0
6	Amapá	97-99	-4,8	-14,8	-3,6	0,2	0	0	0	0	0
7	Tocantins	97-99	-22,4	-10,2	-10,1	-3,0	0	0	0	0	0
8	Maranhão	97-99	-16,7	-14,6	-10,0	8,7	0	0	0	0	0
9	Piauí	97-99	-29,9	-5,2	-10,5	2,8	0	0	0	0	0
10	Ceará	97-99	-25,9	0,3	-13,2	9,9	1	0	0	0	0
11	Rio Grande do Norte	97-99	-17,5	-20,3	-15,5	-13,5	0	0	0	0	0
12	Paraíba	97-99	-11,2	-13,7	-7,1	2,9	0	0	0	0	0
13	Pernambuco	97-99	-11,1	-30,7	-5,0	-8,4	0	0	0	0	0
14	Alagoas	97-99	-14,6	-3,8	-6,3	0,9	0	0	0	0	0
15	Sergipe	97-99	-20,0	-2,5	-8,3	8,6	0	0	0	0	0
16	Bahia	97-99	-18,5	-18,3	-15,7	-4,7	0	0	0	0	0
17	Minas Gerais	97-99	-37,5	-19,7	-32,5	-11,8	0	0	1	0	0
18	Espírito Santo	97-99	-1,4	0,8	6,1	18,0	0	0	0	0	0
19	Rio de Janeiro	97-99	1,3	-7,8	4,5	10,9	0	0	0	0	0
20	São Paulo	97-99	-9,4	-17,4	-2,2	1,1	0	1	0	0	0
21	Paraná	97-99	-30,0	-26,9	-25,1	-17,3	0	0	1	0	0
22	Santa Catarina	97-99	-16,2	-19,3	-7,7	7,2	0	0	0	0	0
23	Rio Grande do Sul	97-99	-5,1	-15,2	2,3	8,7	0	0	1	0	0
24	Mato Grosso do Sul	97-99	-14,7	-25,4	-10,5	-17,8	0	0	1	0	0
25	Mato Grosso	97-99	-15,9	-6,5	-4,8	9,9	0	0	0	0	0
26	Goiás	97-99	-14,6	-23,3	-6,3	-12,5	0	0	0	0	0
27	Distrito Federal	97-99	-13,7	-14,9	-4,4	-0,8	0	0	0	0	0
28	Rondônia	99-01	-1,4	15,9	-3,3	4,2	0	0	0	0	0

Fonte: Elaboração própria a partir das informações disponibilizadas pelo Inep/MEC e sinopse de avaliações estaduais (2014).

Notas:

- LP: Língua Portuguesa; M: Matemática.

- 4^o/5^o: quarta série ou quinto ano; 8^o/9^o: oitava série ou nono ano.

- implantou, manteve, interrompeu, Prova Brasil, Ideb: (1) = "sim"; (0) = "não".

As primeiras colunas da Tabela 1 permitem identificar cada caso de análise. Por exemplo, a primeira linha abaixo do cabeçalho da tabela corresponde ao caso de Rondônia no biênio 1997/1999. Já a 28^a linha trata-se também de Rondônia,

contudo, no biênio 1999/2001, correspondendo, portanto, a outro caso para análise e comparação.

Da quarta até a sétima coluna são apresentados os “deltas” (Δ), os quais correspondem à operação de subtração da proficiência média dos alunos no fim do biênio em relação ao início do mesmo biênio, em uma determinada unidade federativa, série/ano e disciplina (Língua Portuguesa ou Matemática).

Desde a oitava até a 12ª coluna é demonstrada a situação na qual o sistema nacional e os sistemas estaduais se encontram em cada caso. Assim, para cada unidade federativa em determinado biênio é possível saber a situação do sistema estadual (se fora implantado, mantido ou interrompido, situações excludentes entre si) e, ainda, se participou da Prova Brasil e se havia o Ideb disponível para o período. Cada caso foi enquadrado como:

- a. *Implantou*: se aplica à ocorrência ou implantação de um sistema estadual de avaliação em um determinado biênio, não importando se fora de forma esporádica ou continuada no biênio.
- b. *Manteve*: foi aplicada aos casos em que se verificou a existência de um sistema de avaliação estadual no biênio anterior, o qual fora mantido no seguinte.
- c. *Interrompeu*: foi aplicado aos casos em que se constatou a existência de um sistema estadual de avaliação no biênio anterior que fora descontinuado no seguinte.
- d. *Prova Brasil*: aplicação da Prova Brasil no biênio.
- e. *Ideb*: disponibilidade de informações sobre o Ideb no biênio.

Para melhor explorar os dados na análise descritiva com base na Tabela 1, foi calculado o coeficiente de correlação de Pearson (r) entre as variáveis que indicam a ocorrência dos fenômenos (o que inclui desde a quarta até a décima coluna da tabela). Desse modo, é possível mensurar a intensidade com que se dá a associação linear entre duas variáveis. O coeficiente é calculado da seguinte forma:

$$r = \frac{C_{X,Y}}{S_X S_Y}, \quad r \in [-1, 1]$$

Onde:

$C_{x,y}$: covariância ou variância conjunta das variáveis X e Y;

S_x : desvio padrão da variável X;

S_y : desvio padrão da variável Y.

O coeficiente em questão varia entre -1 e 1. Quanto mais próximo de 1 for o valor de r , maior é o grau de associação entre as variáveis X e Y, seja ele positivo (quando X é diretamente proporcional a Y) ou negativo (quando X é inversamente proporcional a Y). Quanto mais próximo de 0 for o valor do coeficiente de Pearson, menor será a intensidade da associação entre as variáveis.

ANÁLISE DOS COEFICIENTES DE CORRELAÇÃO

A Tabela 2 apresenta os coeficientes que indicam o grau de associação entre a variação do desempenho médio nos biênios e a presença/ausência de um sistema estadual de avaliação da educação nas unidades federativas. Embora os valores estejam próximos de zero, são todos positivos, indicando tendência de melhora no desempenho médio quando ocorre a implantação de um sistema avaliativo próprio, independentemente da série/ano, disciplina e biênio que se leve em conta. O r variou entre 0,08, relativo à Matemática na 8ª série/9º ano, e 0,23, relativo à Língua Portuguesa também na 8ª série/9º ano. O que sugere indícios da importância de se elaborar e gerir um sistema estadual de avaliação com vistas à melhora na qualidade da educação.

Quando se dá continuidade no biênio seguinte, mantendo-se o sistema de avaliação próprio, os coeficientes de correlação também são positivos, porém mais próximos de zero, e indicam menor intensidade na interação entre a existência da avaliação e a melhora do desempenho médio nos estados. O maior valor do coeficiente foi verificado em Matemática na 4ª série/5º ano ($r = 0,11$) e a menor em Língua Portuguesa e

Matemática, ambas disciplinas da 8ª série/9º ano ($r = 0,06$). Isso indica que a associação entre “implantação” e desempenho médio, independentemente do ciclo escolar e da disciplina, é mais intensa do que a associação entre “manutenção” e desempenho médio. Contudo, ambos os casos exibem coeficientes de correlação positivos, sugerindo que a implantação do sistema avaliativo, assim como a garantia de sua continuidade, são situações com consequências virtuosas.

Por fim, a interrupção do sistema estadual de avaliação tem relação inversa com o desempenho médio dos alunos nos estados brasileiros; e, também, de forma mais intensa, pois os coeficientes de correlação relativos à interrupção são bastante próximos, em valor absoluto, aos coeficientes relativos à implantação, porém, como ressaltado, com efeito inverso sobre o desempenho. O coeficiente de maior valor absoluto foi verificado entre Língua Portuguesa da 8ª série/9º ano e a interrupção do sistema de avaliação ($r = -0,28$), enquanto o de menor valor absoluto, entre Matemática da 4ª série/5º ano e a interrupção ($r = -0,17$).

TABELA 2 – Coeficiente de correlação de Pearson (r) entre a variação do desempenho médio dos alunos no Saeb em cada biênio (Δ) e a presença/ausência de um sistema estadual de avaliação da educação nas unidades federativas, período 1997 a 2011

VARIACÃO DO DESEMPENHO MÉDIO NOS BIÊNIOS	SISTEMA ESTADUAL DE AVALIAÇÃO DA EDUCAÇÃO		
	IMPLANTOU	MANTEVE	INTERROMPEU
Δ LP 4ª/5ª	0,18	0,10	-0,22
Δ LP 8ª/9ª	0,23	0,06	-0,28
Δ M 4ª/5ª	0,18	0,11	-0,17
Δ M 8ª/9ª	0,08	0,06	-0,19

Fonte: Informações disponibilizadas pelo Inep/MEC e sinopse de avaliações estaduais, 2014 (Tabela 1).

CONSIDERAÇÕES FINAIS

Observou-se crescimento nas curvas de desempenho médio das redes estaduais de ensino no Brasil, em especial após 2001. As tendências gerais são de maiores avanços no primeiro ciclo fundamental. Embora não tenha sido testado neste estudo, o período em que a proficiência do primeiro

ciclo fundamental mais se eleva é concomitante à implantação de avaliações de alfabetização, como é o caso de Minas Gerais e Ceará. Esse indício pode ser objeto de estudos posteriores. O segundo ciclo fundamental também melhorou, contudo, menos e mais tardiamente. Todavia, tal melhoria ocorreu após longo período de quedas, como apresentado na introdução deste trabalho. Portanto, a educação brasileira ainda está se recuperando de uma “crise” de qualidade vivida nos anos 1990. Assim como argumentado em trabalhos anteriores:

A democratização da educação brasileira parece ser caracterizada por uma dualidade [...] Se por um lado avança na direção das principais demandas sociais, por outro, num mesmo movimento, também tende a anular esses avanços, gerando efeitos perversos em outra ponta. (TAVARES Jr. et al., 2012, p. 64)

Embora a tendência geral seja de melhoria, vários biênios registraram estagnação ou declínio em alguns estados. Em especial, tais estados encontram-se no norte e no nordeste do Brasil, regiões mais pobres e com menos sistemas estaduais de avaliação. Não foi objeto deste estudo, mas há alguns indícios de elevação da desigualdade educacional no Brasil, o que é típico de nosso “cobertor curto”: um movimento geral de avançar em uma direção e regredir em outras. Ainda assim, o efeito geral é positivo.

Na análise exploratória dos coeficientes de correlação, observou-se que a implantação de avaliações está relacionada a efeitos positivos de melhoria de desempenho médio. A primeira edição tem efeitos mais fortes que as demais. Além disso, a continuidade das políticas, por meio da manutenção das avaliações, também está relacionada à preservação de tendências ascendentes nas curvas de desempenho. Unidades da federação que adotaram políticas de avaliação externa de suas redes e conferiram estabilidade ao processo, com continuidade dos ciclos avaliativos, apresentaram melhores resultados nos desempenhos médios aferidos pelo Saeb. Minas Gerais e Ceará são exemplos disso.

Por outro lado, a descontinuidade das políticas tem efeito nocivo sobre o desempenho. Os ganhos antes auferidos

tendem a ser anulados ou regredidos posteriormente; Alagoas e Paraná podem ser lembrados. Algo similar acontece quando mudanças significativas nas políticas são feitas durante o percurso; Rio de Janeiro é um dos exemplos. O governo fluminense avalia sua rede há muitos anos, mas adotou diferentes modelos e políticas. Seus resultados são positivos, mas inferiores a outros estados com período similar de implantação de avaliações. Seus melhores resultados foram obtidos exatamente quando manteve por mais tempo o mesmo modelo de política. Mais uma vez, a continuidade das políticas públicas mostra-se mais eficaz e positiva do que seguidas experimentações.

Importa também considerar as avaliações estaduais em paralelo à implantação da Prova Brasil. Poderia ocorrer um efeito de redundância, ou seja, as avaliações estaduais deixariam de ter efeito sobre o desempenho a partir da implantação da Prova Brasil, uma vez que já estaria em curso a avaliação externa nacional; ou poderia ocorrer um efeito de complementaridade, com as avaliações estaduais preservando seu efeito e se somando ao efeito da Prova Brasil. Observou-se que as UFs que mantiveram sistemas próprios de avaliação estadual apresentaram resultados melhores nas edições do Saeb posteriores a 2005. Parece que quando as avaliações têm características distintas e complementares, ser avaliado por mais de um sistema impulsiona melhores resultados.

A recuperação da qualidade da educação na primeira década deste século, decorrente do avanço do desempenho médio das redes públicas, esteve associada à adoção de políticas educacionais e reformas que sempre tiveram nas avaliações em larga escala um elemento central. As “Reformas Educacionais Baseadas em Padrões”, ou *Standards-based Reforms of Education*, representaram um avanço na educação brasileira, ainda que atualmente se verifique crescimento mais lento e desigual do que a queda verificada nos períodos anteriores. É preciso avançar mais, mais rapidamente e com mais equidade.

REFERÊNCIAS

- ALBERNAZ, A.; FERREIRA, F. H.; FRANCO, C. A escola importa? Determinantes da eficácia e da equidade no Ensino Fundamental brasileiro. *Pesquisa e Planejamento Econômico- PPE*, Rio de Janeiro, v. 32, n. 3, dez. 2002.
- ALVES, M. T. G.; SOARES, J. F. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação. *Sociedade e Estado*, v. 22, n. 2, p. 435-473, 2007.
- _____. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. *Educação e Pesquisa*, v. 39, n. 1, p. 177-194, 2013.
- BARBOSA, M. L. *Desigualdade e desempenho: uma introdução à sociologia da escola brasileira*. Belo Horizonte: Argymentvm, 2009.
- BARBOSA, M. L. O.; RANDALL, L. Desigualdades sociais e a formação de expectativas familiares e de professores. *Caderno CRH*, Salvador, v. 17, n. 41, p. 299-308, maio/ago. 2004.
- BETTI, R. Quanto mais prova, melhor. *Revista Veja*, São Paulo, p. 128-130, 26 out. 2011. Seção Educação.
- BIONDI, R. L.; FELÍCIO, F. *Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados Saeb*. Brasília: Inep, 2007.
- BONAMINO, A.; FRANCO, C. A pesquisa em eficácia escolar no Brasil. *Educação on line*. Rio de Janeiro: PUC-RIO, 2005.
- BROOKE, N. O futuro das políticas de responsabilização educacional no Brasil. *Cadernos de Pesquisa*, v. 36, n. 128, p. 377-401, maio/ago. 2006.
- BROOKE, N.; SOARES, J. F. *Pesquisa em eficácia escolar, origens e trajetórias*. Belo Horizonte: Editora UFMG, 2008.
- BUCHMANN, C. Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges. In: PORTER, A.; GAMORAN, A. (Ed.). *Methodological Advances in Cross-National Surveys of Educational Achievement*. Washington, D.C.: National Academy Press, 2002. p. 150-197.
- BYBEE, R. W. *Achieving scientific literacy: from purposes to practices*. Heinemann, 88 Post Road West, PO Box 5007, Westport, CT 06881, 1997.
- CASTRO, M. H. G. A consolidação da política de avaliação da educação básica no Brasil. *Meta-Avaliação*, Rio de Janeiro, v.1, n. 3, p. 271-296, dez. 2009.
- CHATTERJI, M. Models and methods for examining standards-based reforms and accountability initiatives: have the tools of inquiry answered pressing questions on improving schools?. *Review of Educational Research*, v. 72, n. 3, p. 345-386, 2002.
- ELMORE, R. F. *School reform from the inside out: policy, practice, and performance*. Cambridge, MA: Harvard Education Press, 2007.

- ESPÓSITO, Y. L.; DAVIS, C.; NUNES, M. M. Sistema de avaliação do rendimento escolar: o modelo adotado pelo estado de SP. *Revista Brasileira de Educação*, n. 13, p. 25-53, jan./abr. 2000.
- FALK, B. Standards-Based Reforms: Problems and Possibilities. *Phi Delta Kappan*, v. 83, n. 8, p. 612-20, 2002.
- FERNANDES, D. C. *Estratificação educacional, origem socioeconômica e raça no Brasil: as Barreiras da Cor*. In: INSTITUTO DE PESQUISAS ECONÔMICAS APLICADAS (Org.). Prêmio IPEA 40 Anos. Brasília: Ipea, 2005.
- FRANCO, C. Avaliação de sistemas educacionais no Brasil. *Revista de Ciências da Educação*, n. 9, 2009.
- FRANCO, C. et al. Qualidade e equidade em educação: reconsiderando o significado de “fatores intraescolares”. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 15, n. 55, p. 277-298, 2007.
- FRANCO, C.; MANDARINO, M.; ORTIGÃO, M. I. Projeto pedagógico da escola promove eficácia e equidade em educação? *Revista Undime*, Rio de Janeiro, v. 2, n. 7, p. 30-46, 2001.
- FUNDAÇÃO VICTOR CIVITA. *A Avaliação externa como instrumento da gestão educacional nos estados: Relatório Final 2011*. São Paulo: Fundação Victor Civita. Disponível em: <www.fvc.org.br/estudos-e-pesquisas>. Acesso em: 6 mar. 2013
- GAMORAN, A. (Ed.). *Standards-based reform and the poverty gap: Lessons for no child left behind*. Washington, D.C.: Brookings Institution, 2007.
- GREMAUD, A. P.; FELICIO, F.; BIONDI, R. L. *Indicador de efeito escola: uma metodologia para a identificação dos sucessos escolares a partir dos dados da Prova Brasil*. Brasília: Inep, 2007.
- HASENBALG, C.; VALLE SILVA, N. Família, cor e acesso à escola no Brasil. In: HASENBALG, C.; VALLE SILVA, N.; LIMA, M. (Org.). *Cor e estratificação social*. Rio de Janeiro: Contracapa, 1999. p.126-147.
- _____. Raça e oportunidades educacionais no Brasil. *Caderno de Pesquisa*, São Paulo, v. 73, p. 5-12, maio 1990.
- LOCKHEED, M. E.; VERSPOOR, A. M. *Improving primary education in developing countries*. Oxford: Oxford University Press for the World Bank, 1991.
- LÜCK, H. *Dimensões de gestão escolar e suas competências*. Curitiba: Positivo, 2009.
- MCLAUGHLIN, M. W.; SHEPARD, L. A. *Improving Education through Standards-Based Reform*. A Report by the National Academy of Education Panel on Standards-Based Education Reform. Washington, DC: National Academy of Education, 1995.
- MENEZES-FILHO, N. A. *Os determinantes do desempenho escolar no Brasil*. São Paulo: Ibmecc-SP, FEA-USP, 2007.
- MONT´ALVÃO, A. Estratificação educacional no Brasil do século XXI. *DADOS – Revista de Ciências Sociais*, Rio de Janeiro, v. 54, n. 2, p. 389-430, 2011.

ORGANIZAÇÕES DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, CIÊNCIA E CULTURA. *Declaração Mundial sobre Educação Para Todos* (Conferência de Jomtien). Tailândia: Unesco, 1998. Disponível em: <http://unesdoc.unesco.org/images/0008/000862/086291por.pdf>. Acesso em: 29 maio 2014.

SAMMONS, P. As características-chave das escolas eficazes. In: BROOKE, N.; SOARES, J.F. (Org.). *Pesquisa em eficácia escolar: origens e trajetórias*. Belo Horizonte: UFMG, 2008. p. 335-392.

SAMMONS, P.; HILLMAN, J.; MORTIMORE, P. *Key characteristics of effective schools: A review of school effectiveness research*. London: Office for Standards in Education, 1995.

SLAVIN, R. E. Evidence-based education policies: transforming educational practice and research. *Educational Researcher*, v. 31, n. 7, p. 15-21, 2002.

SOARES, J. F. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de pesquisa*, v. 37, n. 130, p. 135-160, jan./abr. 2007.

SOARES, J. F. et al. *Escola eficaz: um estudo de caso em três escolas*. Belo Horizonte: Fundação Ford, 2002.

SWANSON, C. B.; STEVENSON, D. L. Standards-based reform in practice: evidence on state policy and classroom instruction from the NAEP state assessments. *Educational Evaluation and Policy Analysis*, v. 24, n. 1, p. 1-27, 2002.

TAVARES Jr., F.; FARIA, V. B.; LIMA, M. A. Indicadores de fluxo escolar e políticas educacionais: avaliação das últimas décadas. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 52, p. 48-67, maio/ago. 2012.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. Grupo de Avaliação e Medidas Educacionais. *Escola eficaz: um estudo de caso em três escolas da rede pública de ensino do estado de Minas Gerais*. Belo Horizonte: Game, 2002. 114p.

VALLE SILVA, N. Expansão escolar e estratificação educacional no Brasil. In: HASENBALG, C.; VALLE SILVA, N. (Org.). *Origens e destinos: desigualdades sociais ao longo da vida*. Rio de Janeiro: Topbooks, 2003.

VALLE SILVA, N.; SOUZA, A. Um modelo para análise da estratificação educacional no Brasil. *Cadernos de Pesquisa*, São Paulo, v. 58, p. 49-57, ago. 1986.

VELOSO, F. Experiências de reforma educacional nas últimas duas décadas: o que podemos aprender? In: VELOSO, F. et al. (Org.). *Educação básica no Brasil: construindo o país do futuro*. Rio de Janeiro: Elsevier, 2009. p. 191-211.

_____. A evolução recente e proposta para a melhoria da educação no Brasil. In: BACHA, E.; SCHWARTZMAN, S. (Org.). *Brasil: a nova agenda social*. Rio de Janeiro: LTC, 2011. p. 215-253.

FERNANDO TAVARES JÚNIOR

Doutor em Sociologia pelo Instituto Universitário de Pesquisa do Rio de Janeiro (Iuperj), Rio de Janeiro, Rio de Janeiro, Brasil. Professor do Departamento de Ciências Sociais do Instituto de Ciências Humanas da Universidade Federal de Juiz de Fora (ICH-UFJF), Juiz de Fora, Minas Gerais, Brasil
ftavares@caed.ufff.br

LUIZ FLÁVIO NEUBERT

Doutor em Sociologia pela Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brasil. Professor do Departamento de Ciências Sociais do Instituto de Ciências Humanas da Universidade Federal de Juiz de Fora (ICH-UFJF), Juiz de Fora, Minas Gerais, Brasil
luiz.neubert@ufff.edu.br

ANEXO I

Ficha de caracterização de sistemas de avaliação (exemplo)

Características dos sistemas estaduais de avaliação educacional	
ESTADO:	
SISTEMA:	
ANO DE IMPLEMENTAÇÃO :	PERIODICIDADE: () Anual () Bianaual
SÉRIES AVALIADAS: () Alfabetização () 5º ano () 9º ano () Ensino médio	
TIPO DE SISTEMA: () Amostral () Censitário	
DOMÍNIOS AVALIADOS: () Língua Portuguesa () Matemática () Outro:	
RESULTADOS DIVULGADOS POR: () Rede () Escola () Aluno () Outro	
PERÍODO:	
ABRANGÊNCIA: () Só a escola sabe seus resultados () Todas as escolas sabem, mutuamente () Os resultados das escolas são tornados públicos	
VEÍCULOS DE DIVULGAÇÃO: () Boletim de resultados () Boletim pedagógico () Caderno do professor () Página web () Lista pública () Outros:	
OS RESULTADOS IMPLICAM ALGUM TIPO DE AÇÃO? () Sim () Não	
O ESTADO PRATICA ALGUMA AÇÃO DE RESPONSABILIZAÇÃO: () Sim () Não	
Se sim, descreva a ação:	
INSTITUIÇÕES ORGANIZADORAS (por ano):	
CUSTO ESTIMADO DA AVALIAÇÃO:	
O ESTADO IMPLEMENTOU CONCOMITANTEMENTE AO SISTEMA DE AVALIAÇÃO OUTRO TIPO DE AVALIAÇÃO: () Não () Sim. Se sim, descreva o ano e a ação. Em caso de programas complementares, faça uma ficha para o programa implementado	
OBSERVAÇÕES IMPORTANTES E AÇÕES NÃO ELENCADAS:	

Recebido em: AGOSTO 2014

Aprovado para publicação em: OUTUBRO 2014