ESTUDOS EM AVALIAÇÃO EDUCACIONAL

ISSN **0103-6831**

AVALIAÇÃO EM LARGA ESCALA E GESTÃO EDUCACIONAL

FUNDAÇÃO CARLOS CHAGAS

V.25 N.59 SETEMBRO | DEZEMBRO 2014

ESTUDOS EM AVALIAÇÃO EDUCACIONAL • N. 2 JUL./DEZ. 1990 • FUNDAÇÃO CARLOS CHAGAS • SÃO PAULO
Semestral
A partir de 2006 passa a ser QUADRIMESTRAL

ISSN 0103-6831
e-ISSN 1984-932X

1. Avaliação 2. Políticas Educacionais 3. Qualidade do Ensino
I. Fundação Carlos Chagas II. Departamento de Pesquisas Educacionais/FCC

INDEXADO EM
BAMP – Biblioteca Ana Maria Poppovic (*Brasil, FCC*)
www.fcc.org.br

BBE – Bibliografia Brasileira de Educação *(Brasil, Cibec/Inep/MEC)*
http://pergamum.inep.gov.br/pergamum/biblioteca/index.php?resolution2=1024_1

Clase – Citas Latinoamericanas en Ciencias Sociales y Humanidades *(México, Unam)*
http://biblat.unam.mx/

Edubase – *Faculdade de Educação (Brasil, Unicamp)*
http://143.106.58.49/fae/default.htm

Educ@ – Publicações on-line de Educação (*Brasil, FCC*)
http://educa.fcc.org.br/scielo.php

e-Revistas – Plataforma Open Access de Revistas Científicas Electrónicas Españolas y Latinoamericana (*Espanha*)
http://www.erevistas.csic.es/

Iresie – Índice de Revistas de Educación Superior e Investigación Educativa (*México, Cesu-Unam*)
http://www.iisue.unam.mx/iresie/

Latindex – Sistema Regional de Información en Línea para Revistas Científicas de América Latina, en Caribe, España y Portugal (*México, Unam*)
http://www.latindex.unam.mx

VERSÃO ELETRÔNICA
www.fcc.org.br

VERSÃO IMPRESSA
Dezembro de 2014
Tiragem: 500 exemplares

E-MAILS
eae@fcc.org.br *(contato)*
publicacoesfcc@fcc.org.br *(aquisição e assinaturas)*

ESTUDOS EM AVALIAÇÃO EDUCACIONAL

Periódico da Fundação Carlos Chagas criado em 1990 sucedendo *Educação e Seleção* (1980-1989). Publica trabalhos originais relacionados à educação, com perspectiva avaliativa, apresentados sob forma de relatos de pesquisa, ensaios teóricos, revisões críticas, artigos e resenhas.

As normas para a publicação de artigos e resenhas estão no final do volume.
A revista não se responsabiliza pelos conceitos emitidos em matérias assinadas.

COORDENAÇÃO DE EDIÇÕES
Adélia Maria Mariano da Silva Ferreira

ASSISTENTES DE EDIÇÕES
Camila Maria Camargo de Oliveira
Vera Eliana Rodrigues

SECRETÁRIA DE EDIÇÕES
Camila de Castro Costa

PADRONIZAÇÃO BIBLIOGRÁFICA
Biblioteca Ana Maria Poppovic

REVISÃO ESTATÍSTICA
Miriam Bizzocchi
Raquel da Cunha Valle

PROJETO GRÁFICO
Casa Rex

DIAGRAMAÇÃO
Líquido Tecnologia - Claudio Brites

IMPRESSÃO
Nova Letra Gráfica e Editora

SUMÁRIO

OUTROS TEMAS

EDITORIAL

O número 59 de *Estudos em Avaliação Educacional* apresenta, primeiramente, um conjunto de textos que discute os "Usos e efeitos das avaliações externas como objeto de pesquisa". A relevância e atualidade do tema justificam-se pelo número expressivo de edições de avaliações do rendimento escolar em nível nacional e regional que tem gerado considerável quantidade de dados e informações que, por inúmeras razões, nem sempre são analisadas em profundidade pelos órgãos públicos que as gerenciam. Dessa forma, abre-se a possibilidade de os pesquisadores brasileiros utilizarem tais dados para investigar diferentes aspectos e incrementar o conhecimento sobre a educação brasileira.

O dossiê, além da apresentação elaborada por Silva e Carvalho, é formado por quatro textos. O primeiro é de autoria de pesquisadores da Universidade Federal de Juiz de Fora (FJF) que integram o projeto "Análise da evolução da educação básica no Brasil a partir dos indicadores de fluxo e proficiência". Os demais são oriundos de pesquisas realizadas por investigadores de diferentes instituições e áreas de conhecimento que compõem o Observatório Educação e Cidade, no âmbito do projeto "Desigualdades de oportunidades educacionais e dimensões da alfabetização da educação básica no estado do Rio de Janeiro", capitaneado pela Universidade Federal do Rio de Janeiro (UFRJ) e com a participação de quadros da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) e Universidade Estadual do Rio de Janeiro (Uerj).

O primeiro artigo "A qualidade da educação e a disseminação de sistemas de avaliação", de Tavares Júnior e Neubert, trata das relações entre a implantação de sistemas estaduais de avaliação educacional no Brasil e os progressos constatados nos resultados cognitivos dos alunos.

Carvalho, Oliveira e Lima são os autores de "Avaliações externas: tensões e desafios para a gestão escolar". A partir de dados empíricos da rede municipal de ensino do Rio de Janeiro, o texto examina situações e condições vivenciadas pelos gestores escolares diante dos resultados de avaliações externas e suas consequências políticas, pedagógicas e administrativas.

"Cobranças, estratégias e 'jeitinhos': avaliações em larga escala no Rio de Janeiro", de Rosistolato, Prado e Fernández, discorre sobre como os gestores de escolas da rede municipal do Rio de Janeiro compreendem os dados de avaliações educacionais externas e os incorporam nas ações desenvolvidas nas unidades escolares.

O último texto desse conjunto é "Accountability escolar: um estudo exploratório do perfil das escolas premiadas", de Koslinski, Cunha e Andrade. O artigo versa sobre a política de responsabilização escolar implantada na Secretaria Municipal de Educação do Rio de Janeiro.

Na seção "Outros Temas", seis textos analisam outros aspectos da educação brasileira. O primeiro "Avaliações finais sobre o PNE 2001-2010 e preliminares do PNE 2014-2024", de Souza, oferece um panorama de avaliações acadêmicas sobre os processos de elaboração, aprovação, acompanhamento e efeitos dos últimos Planos Nacionais de Educação brasileiros.

Em seguida, Horta Neto, em "Avaliações educacionais e seus reflexos em ações federais e na mídia eletrônica", trata das relações entre os resultados auferidos nessas avaliações e ações e as políticas governamentais, especificamente o Plano de Desenvolvimento da Escola. O texto aborda, também, como a mídia eletrônica avalia a qualidade da educação retratada nessas testagens.

A matrícula em creches no país é tema candente na atualidade. O artigo "Normatizações federais e a oferta de matrículas em creches no Brasil", de Cruz, Farah e Sugiyama, avalia os efeitos das regulamentações da União sobre a ampliação de vagas na educação infantil, bem como o alcance de metas estabelecidas no Plano Nacional de Educação em diferentes unidades federativas para esse nível da educação básica.

Correa, Bonamino e Soares, em "Evidências do efeito da repetência nos primeiros anos escolares", discorrem a respeito de resultados de pesquisa sobre ganhos efetivos nas proficiências em Língua Portuguesa e Matemática entre alunos promovidos e repetentes, para debater os prós e contras da repetência escolar como medida pedagógica.

"Os questionários do Saeb: o que eles realmente medem?" é o artigo elaborado por Karino, Vinha e Laros, que versa sobre os construtos detectados nos instrumentos de delineamento de perfis e contextos da edição de 2009 do Sistema de Avaliação da Educação Básica e suas contribuições para explicar o desempenho escolar dos estudantes brasileiros.

Finalmente, o último texto deste número é "Confiabilidade e concordância entre juízes: aplicações na área educacional", de Matos. Esses quesitos foram analisados entre juízes que corrigiram redações de uma universidade pública de Minas Gerais no período de 2005 a 2010.

O presente número apresenta, portanto, diferentes facetas de estudos avaliativos a partir de dados oferecidos pelas avaliações educacionais em pauta no cenário brasileiro. Esperamos que os textos apresentados possam contribuir e enriquecer o conhecimento entre aqueles que se interessam por essa área de atuação.

Comitê Editorial

TEMA EM DESTAQUE

APRESENTAÇÃO

USOS E EFEITOS DAS AVALIAÇÕES EXTERNAS COMO OBJETO DE PESQUISA

Desde o início dos anos 1990, tem-se verificado o aumento e a consolidação de sistemas de avaliação de desempenho discente no Brasil, promovidos tanto pelo governo federal como por iniciativa de secretarias estaduais e municipais de educação. A consolidação das avaliações externas se desdobra em diversos mecanismos de política educacional, desde a estratégia da ampla divulgação de resultados aferidos ao conjunto da sociedade – notadamente através das coberturas jornalísticas em tom alarmante – até a premiação de agentes escolares em função dos resultados obtidos pelos alunos, dentre outros critérios. Em meio a diversas ações feitas em nome das avaliações externas e dado o peso que a divulgação de seus resultados acarreta, cabe investigar, de forma mais detida, o impacto dessas avaliações na orientação e no aperfeiçoamento do ensino e em que medida elas têm se tornado também instrumentos relevantes e pertinentes para a atuação de professores, gestores escolares e setores técnicos das secretarias de educação.

O presente número de *Estudos em Avaliação Educacional* reúne um conjunto de artigos que analisam diferentes indícios de impacto de avaliações externas na gestão educacional, tanto no cotidiano e na gestão de unidades escolares como no desempenho discente de redes ou sistemas educacionais. Destaca-se assim a relevância dessa agenda de pesquisa que focaliza a influência e a apropriação dos resultados desse tipo de avaliação em escolas e sistemas de ensino público brasileiros como um objeto de pesquisa que merece atenção de pesquisadores e profissionais da área com diferentes abordagens de investigação e possibilidades de aportar conhecimento sobre o tema.

De fato, a adoção de avaliações externas que tenham como ponto forte o emprego de testes padronizados em larga escala, *baseada em censo ou população* de alunos (KELLAGHAN; GREANEY; MURRAY, 2011, p. 19) vai muito além da aplicação de provas, divulgação ou mesmo da análise dos resultados aferidos. O considerável volume de recursos públicos empregados e a crescente quantidade e qualidade das informações geradas impõem a necessidade política, particularmente no Brasil, de aprofundar e diversificar as pesquisas em torno das ações que utilizam as avaliações externas, como resposta responsável da comunidade acadêmica. Por outro lado, a frequente vinculação das avaliações externas à noção de *accountability* e a políticas de responsabilização pode provocar consequências indesejáveis à qualidade do trabalho escolar e ao caráter profissional de seus agentes institucionais (HAERTEL; HERMAN, 2005; LINN, 2000). Desse ponto de vista, reforça-se o interesse na investigação de eventuais efeitos dessas avaliações enquanto subsídio eminentemente pedagógico e formativo para agentes escolares e gestores educacionais. Ressalta-se, assim, que as eventuais implicações dos instrumentos de avaliação externa utilizados, seus desdobramentos e as expectativas nele depositadas demandam a verificação de sua *validade consequencial* (VIANNA, 2003) a partir dos efeitos que os resultados podem despertar no pensamento, nas atitudes e na ação dos diversos interessados.[1]

Atualmente, o crescente incentivo ao uso das avaliações por escolas e professores (GIMENES et al., 2013) parece transcender a compreensão das avaliações externas como *instrumentos de gestão* restritos ao âmbito dos diferentes órgãos decisórios do sistema educacional, como pode fazer supor a primeira geração de avaliação no Brasil, segundo classificação estabelecida por Bonamino e Sousa (2012). Para as autoras, a característica eminente das avaliações daquela primeira geração – como o Sistema de Avaliação da Educação Básica (Saeb) –[2] seria a de propiciar o diagnóstico da qualidade da educação oferecida no Brasil objetivada em uma amostra representativa de escolas e redes de ensino com os desempenhos de seus respectivos alunos, o que inviabilizava consequências diretas para as unidades escolares ou seus profissionais.

1 Segundo Heraldo Vianna (2003, p. 54) a validade consequencial refere-se "ao impacto da avaliação sobre o sistema, determinando mudanças de pensamento, gerando novos comportamentos, formando novas atitudes e promovendo novas ações. A validade consequencial reflete em que medida a avaliação faz realmente alguma diferença para a comunidade".

2 Sem dúvida, o Saeb tem servido de parâmetro para a formulação de diversos sistemas de avaliação em larga escala, em especial quanto à apropriação de suas matrizes de avaliação e a utilização da abordagem metodológica referente à Teoria da Resposta ao Item – TRI.

Além do Saeb, diversos sistemas de avaliação que compõem o atual cenário educacional apresentam como objetivo precípuo fomentar algum tipo de consequência direta para escolas e/ou para seus profissionais ao aferir e produzir resultados por unidade escolar. Tendo em vista esse desenho de avaliação externa, de caráter censitário, podemos encontrar dois grupos distintos. O primeiro deles seria composto pelas avaliações de *segunda geração*, vinculadas a consequências de caráter mais simbólico explicitadas pela ampla divulgação dos resultados à sociedade, pais de alunos e, evidentemente, pelos próprios agentes escolares (BONAMINO; SOUSA, 2012). Trata-se de um tipo de política de avaliação comumente chamada, na literatura acadêmica, de *low stakes* ou de *responsabilização branda* (BROOKE, 2006). A Prova Brasil, de iniciativa do Ministério da Educação, seria um exemplo de avaliação externa de segunda geração.

Já as avaliações de *terceira geração* seriam características das chamadas políticas de *responsabilização forte* ou *high stakes* e envolveriam não apenas a divulgação dos resultados aferidos, mas também sanções ou recompensas em função dos desempenhos obtidos por alunos e escolas. Uma ação de gestão nessa perspectiva de avaliação implicaria algum tipo de consequência material para as unidades escolares e docentes (BROOKE, 2008). No contexto específico da gestão pública brasileira, essa perspectiva de avaliação pode ser exemplificada nos processos de bonificação ou premiação de escolas e professores, verificados em diversos estados e municípios, e que apresentam como um de seus critérios os resultados obtidos pelo conjunto de alunos avaliados em uma mesma escola.

Em síntese, as avaliações de segunda e terceira geração estariam atreladas, de alguma forma, à noção de *responsabilização* de escolas e de seus agentes institucionais em relação aos resultados mensurados por avaliações externas. Sem descartar os aspectos políticos e ideológicos naturalmente vinculados à temática, é de extrema relevância investigar empiricamente algumas hipóteses, como: as políticas de incentivo material a professores e escolas, ou mesmo determinados tipos de exposição pública de resultados,

podem estimular de alguma forma os profissionais de ensino a se engajarem na busca de melhores resultados no contexto brasileiro? Tais medidas seriam capazes de desestimular os agentes públicos da educação ao invés de motivá-los?

As políticas de responsabilização partem da hipótese de que os gestores dos sistemas e as redes de ensino, bem como as equipes gestoras de escolas e os professores, podem utilizar os dados gerados pelas avaliações externas e outros indicadores educacionais para tomar decisões estratégicas e modificar práticas administrativas e pedagógicas para alcançar melhores resultados educacionais, incrementando o desempenho dos estudantes. Obviamente a perspectiva não é culpar os agentes escolares pelos resultados inferiores ao adequado, mas estabelecer sua responsabilidade perante a sociedade pelo trabalho que desenvolvem nos estabelecimentos de ensino e suscitar a reflexão sobre suas práticas para analisar se "podem fazer algo diferente do que vêm fazendo" (FERNANDES; GREMAUD, 2009, p. 222). Na mesma direção, parece também importante discutir a responsabilidade dos gestores dos sistemas e das redes de ensino, considerando que integra suas competências prover as condições de funcionamento das escolas, seja no que diz respeito à infraestrutura, seja no que se refere às condições para o trabalho de docentes e gestores escolares, viabilizando as reflexões e mudanças almejadas para melhorar a qualidade dos serviços educacionais ofertados à população.

Também no plano internacional, a preocupação com a questão da efetiva utilidade das avaliações externas e políticas de responsabilização para melhorar os resultados educacionais tem gerado um crescente número de pesquisas (MARSH; McCOMBS; MARTORELL, 2010; STILLMAN, 2011; SLAVIN et al., 2012; MANDINACH; GRUMMER, 2013), muitas das quais têm focalizado as percepções dos agentes sobre as políticas e as apropriações das informações geradas e disponibilizadas. De fato, a compreensão e o uso dos dados das avaliações externas pelos agentes escolares possibilitam sua intervenção como protagonistas no processo, usando efetivamente os dados como insumo estratégico para a tomada de decisões.

O estudo de Mandinach e Grummer (2013) nos EUA, por exemplo, ressalta que, apesar de a ênfase nas políticas que utilizam dados ser crescente naquele país, ainda se confere pouca atenção para a formação dos profissionais em torno do uso dos dados (*data use*). As autoras enfatizam a necessidade de desenvolver a capacidade de tomar decisões baseadas nos dados disponíveis (*Data-Driven Decision Making – DDDM)* e propõem processos de formação de profissionais de educação com esse objetivo, realçando a relevância de sua "alfabetização técnica" para o uso dos dados (*data literacy)*. Marsh, McCombs e Martorell (2010) examinaram estratégias de "treinamento" com monitores para o DDDM em escolas na Flórida (*instructional coaching*), concluindo que a melhor compreensão dos professores sobre os dados contribui para a melhoria do ensino e o desempenho dos alunos. Stillman (2011), por sua vez, chega a conclusões semelhantes, mas enfatiza a influência dos fatores contextuais, especialmente o papel do diretor ou sua liderança, formação docente em serviço comprometida com a mudança e melhoria dos resultados escolares. Segundo a autora, quando os diretores proporcionam aos professores oportunidades de formação alinhadas com a política e valorizam o papel destes como protagonistas de mudanças na escola a partir das práticas de sala de aula, gera-se uma "tensão produtiva" que favorece seu desenvolvimento profissional e seu engajamento na perspectiva da mudança de práticas docentes com vistas à melhoria dos resultados escolares.

No cenário brasileiro, determinados usos de avaliações externas, protagonizados por Secretarias de Educação (BROOKE; CUNHA, 2011), podem gerar efeitos diversos dependendo das intenções e objetivos políticos da gestão central dos sistemas de ensino, das características e do histórico da rede escolar, do perfil socioeconômico do alunado e do próprio sentido que adquirem tais ações ao longo do tempo. Em especial, quando se constata serem as escolas – e seus agentes – instâncias capazes de reação e reinterpretação das políticas e ações externas a elas direcionadas (MARTINS, 2008) e não apenas unidades executoras da orientação de uma gestão central.

De todo modo, as formas como as escolas e seus profissionais reagem à divulgação dos resultados e às diferentes políticas de responsabilização praticadas no país ainda é assunto pouco explorado empiricamente. Isto é, no âmbito das consequências das políticas de avaliação externa no Brasil, ainda pouco se conhece em relação aos seus efeitos, sobretudo quanto ao trabalho realizado por escolas e professores, mesmo sendo essa uma alegada preocupação da gestão educacional, em âmbito federal, estadual ou municipal.

O Departamento de Pesquisa Educacionais da Fundação Carlos Chagas, por meio de seu Núcleo de Estudos de Avaliação, tem realizado estudos e pesquisas sobre o eventual uso das avaliações externas, bem como analisado a literatura produzida no campo nas últimas duas décadas (SILVA et al., 2013; GIMENES et al., 2013; BAUER; REIS, 2014) no intuito de mapear e aprofundar essa temática. Visando a um maior intercâmbio e uma ampliação do debate em relação aos possíveis efeitos das avaliações externas no cotidiano escolar e na gestão educacional, o dossiê deste número apresenta alguns resultados de pesquisas realizados pelo grupo de pesquisa interdisciplinar e interinstitucional Observatório Educação e Cidade,[3] que tem estudado as desigualdades de oportunidades educacionais, o mapeamento de indicadores educacionais e a apropriação dos sistemas de informação e avaliação no estado do Rio de Janeiro. Trata-se de uma articulação que reúne pesquisadores de diversas instituições (PUC-Rio, UFRJ e Uerj) e de diferentes áreas do conhecimento. Desde 2007 o grupo vem desenvolvendo vários projetos de pesquisa com fomento da Faperj, Inep, CNPq e Capes. O dossiê traz três textos com resultados de pesquisas desse grupo, desenvolvidas com diferentes abordagens metodológicas.

Em novembro de 2013, o Observatório Educação e Cidade realizou um seminário para discutir os resultados acumulados, sobretudo no que se refere às políticas de avaliação e responsabilização e as alterações que vêm provocando no cenário educacional e na gestão das redes de ensino e das escolas. Nessa oportunidade, foi possível estreitar laços de colaboração com pesquisadores da Universidade Federal de Juiz de Fora (UFJF) que desenvolvem investigações sobre

3 Os trabalhos apresentados foram realizados no período de 2011 a 2014, com apoio ao Projeto "Desigualdades de Oportunidades Educacionais e Dimensões da Alfabetização da Educação Básica no Estado do Rio de Janeiro" coordenado pela Universidade Federal do Rio de Janeiro (UFRJ), por meio do Edital n. 038/2010/Capes/Inep – Programa Observatório da Educação.

a evolução da educação básica a partir dos indicadores de fluxo e proficiência.[4] O primeiro artigo que integra o dossiê traz resultados das pesquisas realizadas por esse grupo no âmbito dos sistemas estaduais de ensino. Fernando Tavares Júnior e Luiz Flávio Neubert discutem, nesse artigo, a evolução recente das políticas de avaliação no Brasil, realçando o protagonismo dos estados nesse processo. Os autores chamam a atenção para a disseminação crescente de avaliações estaduais anuais, censitárias e relacionadas à gestão de suas redes na última década. O 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. Para isso, os autores utilizam resultados agregados das redes estaduais nas avaliações nacionais, como parâmetro externo. Os autores concluem que há uma associação estatisticamente significativa entre a implantação de avaliações nos estados e a elevação da proficiência média nos anos seguintes.

Três dos artigos apresentados trazem resultados de pesquisas realizadas na rede municipal do Rio de Janeiro. O artigo de Cynthia Paes de Carvalho, Ana Cristina Prado de Oliveira e Maria de Fátima Magalhães de Lima apresenta três estudos, articulando análises de cunho qualitativo e quantitativo sobre dados da rede municipal do Rio de Janeiro: observações e entrevistas realizadas em uma escola pública da rede e análises de dados dos questionários contextuais dos diretores da Prova Brasil 2007, 2009 e 2011. As autoras discutem os desafios e tensões no âmbito da gestão escolar no contexto das políticas de avaliação e responsabilização. Destacam em particular a sobrecarga de trabalho burocrático numa situação frequente de precariedade de condições de administração, a mediação tensa que as equipes de gestão precisam realizar junto aos agentes escolares para efetivar a implementação das políticas, e, ainda, as tendências de mudança na gestão pedagógica para promover melhores resultados escolares nas avaliações externas. O estudo propõe a necessidade de maior reflexão sobre a pesquisa e a formação em gestão escolar frente aos desafios políticos e pedagógicos no contexto atual.

4 Os pesquisadores integram o Projeto "Análise da evolução da Educação Básica no Brasil a partir dos indicadores de fluxo e proficiência", apoiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes, por meio do Edital n. 038/2010/Capes/Inep – Programa Observatório da Educação.

Com base em dados coletados em grupos focais realizados com gestores escolares, membros da gestão central e técnicos dos órgãos intermediários de gestão da rede municipal de educação do Rio de Janeiro, o artigo de Rodrigo Rosistolato, Ana Pires do Prado e Silvina Julia Fernández descreve e analisa as percepções de profissionais da gestão sobre as avaliações externas de aprendizagem e seus reflexos no cotidiano escolar. Os autores identificam uma tendência à utilização dos dados das avaliações como orientadores das práticas de gestão, mas também observam lacunas de conhecimento sobre os aspectos mais técnicos das avaliações. Essa dinâmica produz o que os autores classificaram como "efeito dominó": um conjunto de cobranças sucessivas que induzem, simultaneamente, estratégias de gestão que visam à melhoria do desempenho institucional e "jeitinhos" para ampliar os índices das escolas.

Mariane Campelo Koslinski, Carolina Portela da Cunha e Felipe Macedo de Andrade investigam o perfil das escolas que foram premiadas por um dos programas que compõem a política de *accountability* escolar da Secretaria Municipal de Educação do Rio de Janeiro (SME/RJ). Partindo da revisão bibliográfica nacional e internacional que indica a relevância dos desenhos de políticas de responsabilização e as atitudes ou cursos de ação que podem induzir nos agentes escolares, os autores descrevem a política de responsabilização escolar da SME/RJ e discutem alguns de seus efeitos com base em dados da própria Secretaria e dos questionários contextuais da Prova Brasil. Dessa forma, identificam algumas limitações do programa de responsabilização e tecem considerações sobre algumas alterações ocorridas na regulamentação no período que buscam corrigir alguns dos sinais/incentivos emitidos pelo programa estudado.

Espera-se, assim, contribuir para a discussão sobre as políticas de avaliação e responsabilização levadas a cabo em redes públicas brasileiras e propiciar a discussão e a reflexão sobre a pluralidade de percepções e reações que geram nos agentes educacionais no âmbito das unidades escolares. Acredita-se que tal reflexão, lastreada por dados de pesquisa empírica, possa também favorecer iniciativas de formação

dos profissionais da educação para que se responsabilizem como protagonistas tecnicamente competentes e comprometidos com a democratização efetiva do direito de aprender das crianças e jovens brasileiros.

VANDRÉ GOMES DA SILVA
vgomes@fcc.org.br

CYNTHIA PAES DE CARVALHO
cynthiapaesdecarvalho@puc-rio.br

REFERÊNCIAS

BAUER, Adriana; REIS, Adriana Teixeira. *Base de Dados Avaliação*: balanço da produção teórica sobre avaliação de sistemas educacionais no Brasil: 1988 a 2011. São Paulo: Fundação Carlos Chagas, 2014. Disponível em: <http://acervo.fcc.org.br>. Acesso em: nov. 2014.

BONAMINO, Alicia; SOUSA, Sandra Zákia. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo , v. 38, n. 2, jun. 2012 .

BROOKE, Nigel. O futuro das políticas de responsabilização educacional no Brasil. *Cadernos de Pesquisa*, São Paulo, v. 36, n. 128, p. 377-401, maio/ago. 2006. Disponível em: <http://www.scielo.br/pdf/cp/v36n128/v36n128a06.pdf>. Acesso em: nov. 2014.

_____. Responsabilização Educacional no Brasil. *Revista Iberoamericana de Evaluación Educativa*, Madrid, v. 1, n. 1, p. 93-109, 2008.

BROOKE, Nigel; CUNHA, Maria Amália de A. A avaliação externa como instrumento da gestão educacional nos estados. *Estudos & Pesquisas Educacionais*, São Paulo, n. 2, p. 17-79, nov. 2011.

FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. Qualidade da educação: avaliação, indicadores e metas. In: VELOSO, Fernando et al. (Org.). *Educação Básica no Brasil*: construindo um país do futuro. Rio de Janeiro: Elsevier, 2009. p. 213-238.

GIMENES, Nelson et al. Além da Prova Brasil: investimento em sistemas próprios de avaliação externa. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 55, p. 12-32, ago. 2013.

HAERTEL, Edward G.; HERMAN, Joan. A Historical Perspective on Validity Arguments for Accountability Testing. In: HAERTEL, E. G.; HERMAN, J. (Ed.). Uses and Misuses of data for educational Accountability and Improvement: *The 104th Yearbook of the National Society for Study of Education*, Part. 2, J. L. Herman and E. H. Haertel (Org.). Malden MA: Blackwell, 2005. p. 1-34.

KELLAGHAN, Thomas; GREANEY, Vicent; MURRAY, T. Scott. *O uso dos resultados da avaliação de desempenho educacional*. Rio de Janeiro: Elsevier, 2011.

LINN, Robert L. Assessments and Accountability. *Educational Researcher*, v. 29, n. 2, p. 4-16, 2000.

MANDINACH, Ellen; GRUMMER, Edith. A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, v. 42, p. 30-37, jan./fev. 2013. Disponível em: < http://edr.sagepub.com/content/42/1/30.short>. Acesso em: nov. 2014.

MARSH, Julie A.; McCOMBS, Jennifer Sloan; MARTORELL, Francisco. How Instructional Coaches Support Data-Driven Decision Making. *Educational Policy*, v. 24, n. 6, p. 872-907, nov. 2010. Disponível em: <http://epx.sagepub.com>. Acesso em: nov. 2014.

MARTINS, Angela. Aspectos organizacionais e dinâmicos na gestão da escola pública: dilemas e conflitos. *Revista Brasileira de Política e Administração da Educação*, v. 24, p. 135-152, 2008.

SILVA, Vandré Gomes et al. *Uso da avaliação externa por equipes gestoras e profissionais docentes:* um estudo em quatro redes de ensino público. São Paulo: FCC/SEP, 2013. Disponível em: <http://www.fcc.org.br/biblioteca/publicacoes/textos_fcc/arquivos/38/arquivoAnexado.pdf>. Acesso em: nov. 2014.

SLAVIN, Robert E. et al. Effects of a Data-Driven District Reform Model on State Assessment Outcomes. *American Educational Research Journal*, v. XX, n. 10, p. 1–26, 2012. Disponível em: <http://aer.sagepub.com/content/50/2/371.full.pdf+html>. Acesso em: nov. 2014.

STILLMAN, Jamy. Teacher Learning in an Era of High-Stakes Accountability: Productive Tension and Critical Professional Practice. *Teachers College Record*, v. 113, n. 1, p. 133–180, jan. 2011.

VIANNA, Heraldo Marelim. Avaliações nacionais em larga escala: análises e propostas. *Estudos em Avaliação Educacional*, São Paulo, n. 27, p. 41-76, jan./jul. 2003.

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, estaduales, 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 estaduales de evaluación educativa en Brasil y el desarrollo de la proficiencia media entre los alumnos de las respectivas redes estaduales de enseñanza. Se utilizan los resultados agregados de las redes estaduales 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 homogeneamente. 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 (1ª, 3ª, 5ª e 7ª) do ensino fundamental e testes de Língua Portuguesa, Matemática e Ciências (avaliou-se redação na 5ª e 7ª 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 esforço[2] 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.

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.

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

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 (4ª série/5º ano e 8ª série/9º 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:

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

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 explanações 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
			ΔLP 4ª/5º	ΔLP 8ª/9º	ΔM 4ª/5º	ΔM 8ª/9º	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.
- *4ª/5ª*: quarta série ou quinto ano; *8ª/9ª*: 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 28ª 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

VARIAÇÃO DO DESEMPENHO MÉDIO NOS BIÊNIOS	SISTEMA ESTADUAL DE AVALIAÇÃO DA EDUCAÇÃO		
	IMPLANTOU	MANTEVE	INTERROMPEU
Δ LP 4ª/5º	0,18	0,10	-0,22
Δ LP 8ª/9º	0,23	0,06	-0,28
Δ M 4ª/5º	0,18	0,11	-0,17
Δ M 8ª/9º	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: Argvmentvm, 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: Ibmec-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.ufjf.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@ufjf.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 ( ) Bianual
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

AVALIAÇÕES EXTERNAS: TENSÕES E DESAFIOS PARA A GESTÃO ESCOLAR

CYNTHIA PAES DE CARVALHO
ANA CRISTINA PRADO DE OLIVEIRA
MARIA DE FÁTIMA MAGALHÃES DE LIMA

RESUMO

No contexto das políticas de avaliação e responsabilização na rede pública municipal de ensino do Rio de Janeiro, o artigo traz três estudos que partem de observações e entrevistas realizadas em uma escola pública da rede e de incursões exploratórias sobre os dados dos questionários contextuais dos diretores da Prova Brasil 2007, 2009 e 2011. São discutidos os desafios e as tensões cada vez mais presentes na pauta da gestão escolar, com destaque para a sobrecarga de trabalho burocrático num contexto de precariedade de condições de administração, de uma mediação tensa das políticas junto aos agentes escolares e ainda das tendências de mudanças na gestão pedagógica. Conclui-se ressaltando a importância do adensamento da reflexão sobre a pesquisa e a formação em gestão escolar diante dos novos desafios políticos e pedagógicos, sem perder de vista suas limitações administrativas e político-legais.

PALAVRAS-CHAVE GESTÃO ESCOLAR • AVALIAÇÃO DA EDUCAÇÃO • RESPONSABILIZAÇÃO • RIO DE JANEIRO.

RESUMEN

En el marco de las políticas de evaluación y responsabilización en la red pública municipal de enseñanza de Rio de Janeiro, el artículo presenta tres estudios que parten de observaciones y entrevistas realizadas en una escuela pública de la red y de incursiones exploratorias de los datos de los cuestionarios contextuales de los directores de la Prova Brasil *2007, 2009 y 2011. Se discuten los desafíos y tensiones crecientemente presentes en la pauta de la gestión escolar, con destaque para la sobrecarga de trabajo burocrático en un contexto de precariedad de condiciones de administración, una mediación tensa de las políticas junto a los agentes escolares y asimismo las tendencias de cambios en la gestión pedagógica. Se concluye resaltando la importancia del incremento de la reflexión sobre la investigación y la formación en gestión escolar frente a los nuevos desafíos políticos y pedagógicos, sin perder de vista sus limitaciones administrativas y político-legales.*

PALABRAS CLAVE GESTIÓN ESCOLAR • EVALUACIÓN DE LA EDUCACIÓN • RESPONSABILIZACIÓN • RIO DE JANEIRO.

ABSTRACT

In the context of evaluation and accountability policies in the public municipal school network in Rio de Janeiro, this article deals with three studies based on observations and interviews conducted in one public school in the network and exploratory inroads about the data from the contextual questionnaires of the directors of the 2007, 2009 and 2011 Prova Brasil. *The increasing challenges and tensions present in the agenda of school management are discussed, highlighting the bureaucratic work overload in a context of the precarious conditions of administration, of a tense mediation of the policies together with the school agents and, also, the trends of changes in school management. It concludes by highlighting the importance of the consolidation of thinking about research, and the preparation of school management in the face of the new political and pedagogical changes, without losing sight of their administrative and politico-legal limitations.*

KEYWORDS EDUCATIONAL MANAGEMENT • EDUCATIONAL EVALUATION • ACCOUNTABILITY • RIO DE JANEIRO.

INTRODUÇÃO

Propõe-se, neste artigo, a discussão sobre as demandas da gestão escolar frente às estratégias de avaliação externa e às políticas de responsabilização adotadas em alguns segmentos – estaduais e/ou municipais – da educação pública brasileira. O texto encontra-se organizado em seis seções, incluindo esta introdução. O contexto dessa análise – tratado na segunda seção deste artigo – é a introdução de sistemas de avaliação externa para aferição da aprendizagem nas escolas brasileiras, que se inicia na década de 1990. As mudanças na rotina escolar, trazidas pela introdução dessas políticas, estabelecem demandas específicas para o trabalho da gestão da escola.

A terceira seção trata da centralidade da gestão para a eficácia escolar e dos impactos percebidos no trabalho da direção e da coordenação de uma escola municipal do Rio de Janeiro, campo de um estudo de caso nos anos 2010 e 2011 (OLIVEIRA, 2012a). Entre os resultados desse estudo, destacam-se as preocupações da equipe de gestão com as metas de desempenho que a escola devia alcançar, já que a

equipe tinha que lidar com desafios que comprometiam suas condições de trabalho, como a ausência de funcionários ou docentes.

Na quarta seção, essa problemática é analisada com base em um levantamento de dados provenientes das edições de 2007, 2009 e 2011 da Prova Brasil. Nesse segundo movimento de pesquisa, são estabelecidas algumas comparações com o objetivo de examinar as possíveis mudanças nas estratégias adotadas por gestores escolares a partir da premiação (ou não premiação) recebida em reconhecimento pelo fato de a referida escola ter atingido a meta estabelecida pela Secretaria Municipal de Educação do Rio de Janeiro (SME/RJ). Os resultados desse estudo exploratório são apresentados na quinta seção do artigo, seguidos de uma exposição das características das políticas de responsabilização.

Finalmente, a sexta seção traz uma análise do cenário em que se desenvolveram os três estudos, propondo uma reflexão sobre o papel da gestão escolar em um contexto de avaliações externas e políticas de responsabilização.

POLÍTICAS DE AVALIAÇÃO E RESPONSABILIZAÇÃO

No Brasil, desde 1990, o Sistema de Avaliação da Educação Básica (Saeb) tornou possível conhecer – com base na aplicação de avaliações externas a amostras representativas por unidades da federação – o desempenho dos alunos em Língua Portuguesa e Matemática em cada segmento do ensino fundamental e no ensino médio. Em 2005, o Ministério da Educação (MEC) ampliou o monitoramento do ensino fundamental por meio da Prova Brasil, que avalia o desempenho em Língua Portuguesa e Matemática dos alunos dos anos finais de cada segmento dessa etapa de ensino (5º e 9º anos) em todas as escolas da rede pública do país com mais de 20 estudantes matriculados nos anos avaliados. Sua intenção diagnóstica é defendida pelo Plano de Metas Todos pela Educação:

> Com os resultados do Prova Brasil, as secretarias e o MEC têm um diagnóstico da Educação brasileira, podendo detectar desigualdades nas escolas e entre elas. A partir disso,

> esses órgãos devem definir ações e direcionar recursos para corrigir essas distorções e melhorar a qualidade do ensino. (BRASIL, 2007)

Esse movimento de focalização das políticas federais de avaliação, ocorrido a partir de 2007, tem enfatizado os resultados por escolas e redes de ensino e catalisado ou induzido à adoção de políticas locais (estaduais e municipais) que, em muitos casos, se associam a políticas de responsabilização. Pesquisas recentes de Brooke (2008) e do Grupo de Avaliação e Medidas Educacionais – Game – da Faculdade de Educação da Universidade Federal de Minas Gerais – UFMG (2011) registram o crescimento de sistemas subnacionais de avaliação, intensificado nos últimos dez anos, após a aplicação bianual de exames padronizados censitários (Prova Brasil) iniciada em 2005.

Um dos principais fatores que contribuíram para a expansão de sistemas de avaliação foi a criação do Índice de Desenvolvimento da Educação Básica (Ideb) e sua centralidade na política educacional. Por meio desse índice, inscrito no Plano de Desenvolvimento da Educação (PDE), de 2007, foi possível arbitrar e fixar metas bianuais por escola e rede de ensino. O Ideb – descrito por Fernandes e Gremaud (2009, p. 1) como "um indicador sintético da qualidade da educação básica, que considera tanto o desempenho dos estudantes em exames padronizados quanto a progressão desses alunos no sistema" – combina o fluxo escolar (taxas de reprovação e aprovação) com o resultado do desempenho padronizado na Prova Brasil, monitora as ações do PDE e é divulgado a cada dois anos, com acesso público no portal do MEC e do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Cabe sublinhar, também, a importância estratégica do "Plano de Metas Compromisso Todos Pela Educação" como instrumento que provocou a assunção de uma responsabilidade pactuada e propiciou uma relação de colaboração mais sistemática entre a União, os estados e os municípios, condicionando o apoio técnico e as transferências voluntárias do MEC à adesão ao PDE e ao consequente alcance das 28 diretrizes nele estabelecidas. O Plano obteve a adesão de

100% dos municípios e vincula o aporte técnico e financeiro à elaboração dos Planos de Ações Articuladas, que deve ocorrer com a participação dos gestores, da sociedade civil e dos educadores locais.

É importante ressaltar que, por meio da divulgação de resultados e da fixação de metas, o MEC pôde estabelecer uma forma branda de *accountability*, na qual estados e municípios passaram a monitorar a qualidade do ensino, prestar informações sobre os resultados de desempenho e intermediar apoio técnico e financeiro para as escolas que apresentam piores resultados.

Dessa maneira, percebe-se que o uso dos resultados das avaliações de exames padronizados articulados à política de responsabilização, com consequências fortes para as escolas brasileiras, não foi preconizado por um sistema central de avaliação no Brasil. Na verdade, pode-se levantar a hipótese de que o desenho de política, com base nos dados das avaliações externas, está mais estreitamente relacionado à autonomia das redes públicas de ensino na gestão de seus sistemas, ao funcionamento descentralizado do federalismo brasileiro e às responsabilidades dos estados e municípios pela qualidade do ensino médio e fundamental, conforme inscrito na Constituição Federal de 1988 e na Lei de Diretrizes e Bases de 1996. Adicionalmente, o Ideb tem induzido o desenvolvimento de sistemas de avaliação, planos de metas e políticas de responsabilização com o intuito de melhorar o desempenho acadêmico dos alunos. Esses aspectos são importantes para compreender o contexto favorável ao desenvolvimento de formas de *accountability* nas redes públicas de ensino brasileiras.

As chamadas "políticas de responsabilização" adotadas em alguns sistemas educacionais estaduais ou municipais comumente oferecem um bônus ou prêmio para aquelas escolas que atingirem as metas estabelecidas para suas unidades de ensino a cada divulgação do Ideb. Algumas características dessas políticas são tributárias das políticas de *accountability* dos Estados Unidos e foram incorporadas logo no início das políticas de responsabilização no Brasil. Segundo Andrade (2008), a partir de 2006, as políticas nacionais passam a uti-

lizar, além da avaliação em larga escala, outra característica da política de *school accountability*: a divulgação dos resultados por escola. De acordo com o autor, as políticas educacionais locais não haviam incorporado, até aquele momento, dois outros pontos característicos da *school accountability*, os que se referem ao foco na melhoria do desempenho do aluno e à responsabilização. Até então, em nível nacional, "pode-se dizer que o governo federal adotou uma política branda de *school accountability*" (ANDRADE, 2008, p. 446).

Dentre as iniciativas localizadas de avaliação do ensino público, destacam-se as experimentadas em Minas Gerais, São Paulo e Rio de Janeiro, oriundas de políticas dos governos desses estados ou de alguns de seus municípios. A intenção dessas iniciativas, em geral, **é** monitorar o ensino em suas unidades (escolas) a fim de garantir que os alunos estejam sendo efetivamente preparados para a participa**ção** na Prova Brasil. Além desse monitoramento, visam também diminuir o índice de reprovação, corrigindo distorções no fluxo dos alunos e garantindo a adequação idade/série. A melhora no resultado da Prova Brasil e no fluxo dos alunos da escola resultaria na elevação do Ideb, meta perseguida pelas secretarias de educação municipais e estaduais. Algumas dessas iniciativas são acompanhadas por estratégias de responsabilização, como o pagamento de recompensa aos professores das escolas que atingem ou ultrapassam as metas estipuladas.

No município do Rio de Janeiro, no qual as pesquisas ora apresentadas foram realizadas, a Secretaria Municipal de Educação desenvolveu, a partir de 2009, uma nova política educacional em consonância com a política nacional, à qual se integrou também uma política de responsabilização, que será descrita detalhadamente na quinta seção. Nessa seção, também serão apresentados os aspectos da política que dizem respeito ao novo conjunto de modalidades avaliativas instituído na rede: a Prova Rio[1] e as Avaliações Bimestrais da Rede.

A Prova Rio é uma avaliação anual externa realizada por meio de exames padronizados que seguem, conforme divulgado pela SME/RJ, os moldes da Prova Brasil, ou seja, é composta de testes de múltipla escolha que abrangem as áreas de matemática e língua portuguesa. A SME/RJ instituiu um

1 Segundo o Centro de Seleção e de Promoção de Eventos (CESPE) – da Universidade de Brasília (UnB) –, responsável pela elaboração e aplicação da Prova Rio, ela é um sistema de avaliação da rede municipal de ensino do Rio de Janeiro e configura-se como uma avaliação longitudinal de desempenho. A Prova é aplicada aos estudantes do 3º, 4º, 7º e 8º anos do ensino fundamental desde 2010. A avaliação do 5º e do 9º anos é realizada por meio da Prova Brasil, cuja matriz de referência é compatibilizada para a elaboração das questões da Prova Rio. Seu resultado gera o Índice de Desenvolvimento da Educação do Município do Rio de Janeiro (IDE-Rio) no ano subsequente à aplicação. Esse índice e o Ideb (medido pela Prova Brasil e pela taxa de aprovação) servem como parâmetros para o estabelecimento de bonificações anuais às escolas da rede, conforme o Acordo de Resultados – anos ímpares IDE-Rio e anos pares Ideb. Disponível em: <http://provario.cespe.unb.br/login/index.php> e <http://www.cespe.unb.br/noticiashtml/LerNoticia.asp?IdNoticia=752>. Acesso em: 22 jan. 2014.

processo de treinamento para a Prova Rio, que envolve, além da preparação dos alunos (por meio da ênfase no ensino dos conteúdos a serem testados nas avaliações e de exercícios que incluem a familiarização com o tipo de teste), a realização de provas ou simulados nas séries que serão avaliadas. Os simulados são testes aplicados também anualmente pela SME/RJ, seguindo os padrões das avaliações externas como treinamento para a Prova Rio e para a Prova Brasil.

As Avaliações Bimestrais da Rede são testes de múltipla escolha de Língua Portuguesa, Matemática e Ciências, aplicados a todas as séries do ensino fundamental. Seus resultados são monitorados pela SME/RJ e compõem a nota média do aluno na escola. Essas avaliações estão articuladas a dois outros instrumentos didáticos introduzidos pela mesma política: os Descritores de Ensino (listas bimestrais de habilidades e conteúdos a serem contemplados nas Avaliações Bimestrais da Rede) e os Cadernos de Apoio Pedagógico (apostilas bimestrais de exercícios das disciplinas avaliadas que visam complementar o trabalho do professor e preparar os alunos para a Avaliação Bimestral da Rede). Esse conjunto de políticas passou a vigorar como resposta e ruptura em relação à política educacional da gestão anterior (2005/2008), que organizava os anos escolares em ciclos e que ficou popularmente conhecida como "aprovação automática", intensificada e estendida a todo o ensino fundamental no último ano daquela gestão municipal.

Quando essa nova "política avaliativa", com todos os seus instrumentos e demandas, chegou ao contexto escolar, alterou significativamente sua rotina. Como os educadores incorporaram esses instrumentos em sua prática? Como lidaram com o desafio de gerenciar tais mudanças? O que cabe à gestão da escola nesse contexto? Para pensar essas questões, busca-se refletir, na próxima seção, sobre a relação entre gestão e eficácia escolar.

GESTÃO ESCOLAR

Ball (2001, p. 108) considera que, "durante os últimos vinte anos, a gestão tem sido um mecanismo chave tanto na

reforma política, quanto na reengenharia cultural no setor público [...]. Enquanto os mercados trabalham de fora para dentro, a gestão funciona de dentro para fora".

Os estudos sobre a eficácia escolar vêm tentando determinar o efeito das experiências vivenciadas na escola (que são particulares e relativas às características da instituição) na aprendizagem dos alunos. Assim, controlando as características iniciais dos alunos (origem socioeconômica, cor/raça, organização familiar, por exemplo), essas pesquisas buscam aferir quais fatores estruturais e organizacionais das escolas estão associados a diferentes resultados acadêmicos entre elas, ou seja, entre escolas que atendem alunos com perfis semelhantes, e o que aquelas que apresentam melhores resultados acadêmicos têm e fazem para alcançá-los.

Sammons (2008, p. 336) levantou onze "características-chave das escolas eficazes", dentre as quais se destacam estrutura da escola, aspectos da gestão, valorização de um ambiente organizado para a aprendizagem e estratégias do professor. Alves e Franco (2008, p. 494) resumem, no caso brasileiro, tais fatores em cinco categorias: recursos escolares, organização e gestão da escola, clima acadêmico, formação e salário docente e ênfase pedagógica. Ambos os trabalhos apontam a gestão escolar como um dos fatores intraescolares que podem influenciar positivamente o desempenho do aluno e a eficácia escolar. Com base na proposição de um modelo conceitual para abordar o processo de aprendizagem em uma perspectiva multidisciplinar, Soares (2007, p. 153) considera que "dentro da escola há dois importantes processos que interagem para a produção do desempenho dos alunos: a gestão escolar e o ensino". De acordo com o autor, a gestão da escola, liderada pelo(a) diretor(a), compreende as tarefas relativas à garantia do funcionamento da instituição "de forma que os recursos nela existentes possam ser usados para atender às necessidades de aprendizagem dos alunos" (SOARES, 2007, p. 153). Isso envolve tarefas rotineiras, como acompanhar e efetivar a assiduidade de alunos e professores, identificar os problemas de aprendizagem e buscar recursos para saná-los, estabelecer contatos com as famílias dos alunos em prol de seu acompanhamento escolar, garantir o

acesso de alunos e professores aos materiais didáticos necessários para o trabalho escolar, entre outras. Sendo a gestão escolar um dos fatores intraescolares que podem influenciar o desempenho dos estudantes, como seu trabalho é afetado pelas políticas de avaliação e responsabilização em vigor?

A pesquisa de Oliveira (2012a), desenvolvida durante os anos de 2010 e 2011,[2] ressaltou que as determinações das políticas públicas, como as que instituem avaliações externas para acompanhar o trabalho desenvolvido nas escolas, dependem, em sua maioria, da articulação do trabalho da gestão para chegar ao professor e à sala de aula. De acordo com os dados desse estudo, a mediação realizada pela equipe gestora da escola, entre as determinações da SME/RJ e o trabalho dos professores, foi fundamental para o bom andamento da atividade docente. Muitas dúvidas surgiam na organização do calendário de avaliações e no cumprimento das exigências burocráticas e cabia à gestão da escola esclarecer e direcionar o trabalho.

2 Trata-se de um estudo baseado num trabalho intenso de mais de um ano de observação do cotidiano escolar, que incluiu, ainda, entrevistas com a equipe de gestão da escola (direção e coordenação pedagógica) e professores.

Entre as mudanças provocadas no trabalho dos gestores pela aplicação das avaliações aos alunos da rede pública municipal do Rio de Janeiro, mencionadas anteriormente, o estudo em tela destacou as seguintes: reorganização do calendário e da rotina da escola, alterando horários e definindo salas para a aplicação de provas, horários para esclarecer dúvidas e instruir os professores sobre os instrumentos, prazos e demandas de cada avaliação; monitoramento, publicação e comparação dos resultados dos alunos, propondo estratégias de recuperação para aqueles com dificuldades e atividades de revisão para as questões com grande índice de erro. O estudo mostrou um expressivo impacto no trabalho desenvolvido pela gestão da escola, ressaltando que as entrevistas e as observações realizadas no ambiente investigado enfatizaram a queixa quanto ao excesso de avaliações e a prevalência do trabalho burocrático do gestor, dada a necessidade constante de organização e registro dos dados. Contudo, a equipe gestora incorporou as novas atribuições em sua rotina, buscando desempenhá-las de acordo com o que lhes era determinado e com o compromisso de manter os bons resultados que a escola tradicionalmente atingia antes da política.

Talvez estas sejam mais algumas características de uma gestão comprometida com a eficácia escolar: a capacidade de adaptação às mudanças e a proatividade frente aos desafios. Entretanto, o mesmo estudo mostrou que, muitas vezes, as estratégias da equipe de gestão eram afetadas pela sobrecarga de trabalho e pelas dificuldades para manter a estabilidade e a assiduidade no quadro de funcionários da escola. Na seção que se segue, propõe-se uma reflexão sobre o impacto das condições de trabalho do diretor na gestão escolar.

CONDIÇÕES DE TRABALHO DA GESTÃO ESCOLAR

As condições de gestão das redes públicas de ensino e, particularmente, das escolas dependem de insumos que interferem em seu funcionamento, produzindo efeitos nas condições de trabalho e na qualidade do ensino. Contudo, os diretores geralmente têm pouca margem de decisão ou poder para alterar algumas das dimensões da gestão escolar e o quadro frequente de escassez de recursos que afetam seu cotidiano.

Torna-se, assim, imperativo que formuladores de políticas considerem, nos contextos escolares, a disponibilidade e as condições de uso dos recursos relacionados à infraestrutura, ao quadro de pessoal, aos materiais pedagógicos e **às dependências** físicas que, conforme evidências de pesquisas (GOMES; REGIS, 2012; SOARES; SATYRO, 2008; SOARES et al., 2010; ALVES, 2009; SOARES; TEIXEIRA, 2006), **não são** distribuídos homogeneamente nas redes de ensino ou entre as escolas.

Nesse sentido, os questionários contextuais da Prova Brasil podem oferecer informações sobre alunos e escolas, em particular sobre as percepções dos agentes escolares – professores e diretores – no âmbito das condições de gestão. Essas informações podem ainda ser complementadas pelas percepções dos aplicadores dos exames sobre as condições de infraestrutura do ambiente escolar, identificando problemas estruturais e de manutenção ou uso efetivo de recursos que podem influenciar o desempenho do aluno. Dentre os estudos citados, cabe ressaltar o de Gomes e Regis (2012),

que se vale dessas informações para discutir sua influência nos resultados escolares na região metropolitana do Rio de Janeiro. Esses aspectos parecem particularmente relevantes quando se trata de implementar políticas de avaliação e de responsabilização pelos resultados, que estabelecem a meritocracia e a bonificação como mecanismos indutores de qualidade.

Dentre os elementos da gestão que podem ser apreendidos nas respostas dos questionários contextuais da Prova Brasil, destacam-se a gestão de pessoal como uma das dimensões que potencialmente afetam o trabalho dos profissionais de educação, o funcionamento das escolas e o desempenho dos alunos (LIMA; ALCÂNTARA; ALMEIDA, 2013). Cabe assinalar que essa dimensão sofre ação limitada dos diretores, em vista da legislação que orienta a organização das redes públicas de ensino, a lotação de funcionários e docentes, bem como o eventual contingenciamento de recursos financeiros e a dotação orçamentária fixada para essa finalidade. Em que pese a centralidade dos diretores e seu papel de liderança na organização do trabalho escolar, suas possibilidades de intervenção são dependentes da política e da ação estatal para atender, por exemplo, a demanda de professores e funcionários das escolas.

No segundo movimento de pesquisa aqui considerado, analisam-se as respostas dos diretores da rede pública municipal de ensino do Rio de Janeiro aos questionários contextuais da Prova Brasil em 2007, 2009 e 2011, como mostram as Tabelas 1 e 2 apresentadas a seguir.

TABELA 1 – Pessoal Administrativo

OCORREU NA ESCOLA CARÊNCIA DE PESSOAL ADMINISTRATIVO	2007	2009	2011
Não	41%	32%	38%
Sim, mas não foi um problema grave	36%	34%	38%
Sim, e foi um problema grave	21%	31%	19%
Não respondeu	1%	4%	4%
Total	100%	100%	100%

Nota: Em 2007 a pergunta do questionário contextual teve a seguinte redação: "*Problema que ocorreu na escola dificultando o funcionamento: Carência de pessoal administrativo?*". Já em 2009 e em 2011, a redação foi: "*Ocorreu na escola carência de pessoal administrativo?*".

Fonte: Prova Brasil 2007, 2009 e 2011.

Na Tabela 1, as respostas dos diretores em 2009 indicam uma percepção mais aguda da carência de pessoal administrativo naquele ano, que parece corroborar as observações do estudo de caso de Oliveira (2012b) a respeito da sobrecarga de tarefas burocráticas da gestão, sem uma contrapartida de apoio administrativo de acordo com as novas demandas surgidas com as políticas de avaliação da SME/RJ. Ainda que os dados mostrem uma alteração em 2011 quanto à percepção da gravidade do problema, 57% dos diretores continuam a identificar "carência de pessoal técnico administrativo" como um problema da escola.

TABELA 2 – Pessoal de Apoio Pedagógico

OCORREU NA ESCOLA CARÊNCIA DE PESSOAL DE APOIO PEDAGÓGICO	2007	2009	2011
Não	69%	64%	71%
Sim, mas não foi um problema grave	20%	17%	17%
Sim, e foi um problema grave	10%	16%	8%
Não respondeu	1%	4%	4%
Total	100%	100%	100%

Nota: Em 2007 a pergunta do questionário contextual teve a seguinte redação: "*Problema que ocorreu na escola dificultando o funcionamento: Carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?*". Já em 2009 e em 2011, a redação foi: "*Ocorreu na escola carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional)?*".

Fonte: Prova Brasil 2007, 2009 e 2011.

No que se refere ao pessoal de apoio pedagógico, a Tabela 2 registra um crescimento da percepção da carência de coordenadores e outros profissionais de apoio pedagógico em 2009, que se dilui em 2011. Essa alteração positiva da percepção de falta de profissionais também é observada nos dados relativos ao corpo docente e se explica pelo ingresso na rede de cerca de 9.000 profissionais de educação[3] entre 2009 e 2011 (FUNDAÇÃO JOÃO GOULART, s.d.).

Ainda assim, os dados sinalizam que a sobrecarga de trabalho dos diretores já era acentuada, notadamente diante do volume das atribuições da gestão, ampliado em decorrência da proliferação de programas e projetos educacionais e da interface social da escola na última década. A indução à melhoria da qualidade requer, além de novos conhecimentos técnicos, tempo para planejar, executar despesas, prestar contas, manter reuniões com a comunidade escolar e acompanhar o processo de ensino-aprendizagem desenvolvido pela escola. Desse ponto de vista, a expectativa seria que os diretores, para além da capacidade técnica necessária para conduzir a gestão, estivessem desimpedidos das atribuições burocráticas que podem ser cumpridas por funcionários administrativos, para dedicarem-se aos aspectos políticos da gestão e às questões pedagógicas que afetam a aprendizagem e o desempenho dos alunos.

Cabe ainda ressaltar que, na rede pública do Rio de Janeiro, no caso dos funcionários (merendeiras e serventes), não há concurso público há mais de dez anos e a maioria dos profissionais que atuam na limpeza, na alimentação e na portaria das escolas são terceirizados[4] ou mesmo, em alguns casos, voluntários. A precariedade dessas situações e soluções, no entanto, tampouco supre a demanda de serviços administrativos das escolas, como mostram as respostas dos diretores.

3 Ingressaram na rede municipal do Rio de Janeiro, por meio de concurso público, no período de 2009 a 2011, 8.612 profissionais, dos quais 2.891 eram professores de educação infantil; 1.801, professores de anos iniciais e 3.920, professores de anos finais do ensino fundamental, segundo dados da Fundação João Goulart, responsável pelos concursos.

4 Parte da demanda é atendida por meio de contratos com empresas que terceirizam serviços de limpeza e preparo da alimentação escolar.

A POLÍTICA DE BONIFICAÇÃO NA REDE MUNICIPAL DO RIO DE JANEIRO

Como já assinalado, a SME/RJ desenvolveu, a partir de 2009, uma política de avaliação articulada a uma política de

responsabilização e bonificação dos agentes escolares de acordo com os resultados anuais alcançados pelas escolas da rede, considerando as metas estabelecidas pela política nacional e pela própria Secretaria com base nos resultados do Ideb e do IDE-Rio. Nesta seção, apresenta-se essa política, a qual integra o escopo mais amplo da denominada "gestão de resultados" que se constituiu como uma marca de todo o governo municipal no período analisado, formalizada por meio da assinatura anual de um Acordo de Resultados e Termo de Compromisso, que define metas para os gestores (centrais, intermediários e locais) e funcionários dos diferentes órgãos da administração direta e entidades da administração pública indireta da Prefeitura (RIO DE JANEIRO, 2009).[5]

No caso das escolas, as metas são previamente pactuadas mediante o Termo de Compromisso de Desempenho Educacional, assinado pelo diretor de cada unidade. Diretores, professores e funcionários que atingem as metas recebem uma bonificação anual correspondente a um salário extra ou o 14º salário. Embora o detalhamento do desenho da política tenha sofrido algumas adaptações no período analisado (2009 a 2011), sua análise foge ao escopo do terceiro movimento de pesquisa, integrado neste artigo, que busca identificar eventuais tendências de mudança na gestão pedagógica das escolas a partir da premiação, ou não, dos agentes escolares.

Do ponto de vista da gestão escolar, as consequências das políticas de bonificação pelo resultado parecem integrar a pauta da relação entre os agentes escolares e desafiar a competência pedagógica e política da equipe de gestão na condução das ações com vistas ao alcance das metas definidas pela SME/RJ. Nesse terceiro movimento de investigação, busca-se identificar possíveis tendências de alteração nas práticas dos diretores referentes à gestão pedagógica (atividades de formação continuada, critérios de matrícula e de organização de turmas) que, potencialmente, estariam relacionadas às políticas de avaliação e responsabilização/bonificação. Considera-se a hipótese de que a política de bonificação tensiona as relações na escola, num contexto de gestão já desafiado pela frequente escassez de recursos humanos

5 "Para o melhor acompanhamento do Plano Estratégico de Gestão da Prefeitura foram pactuados Acordos de Resultados como uma ferramenta usada na administração pública para: (i) garantir um maior comprometimento de toda a máquina pública municipal com os resultados da Prefeitura; (ii) institucionalizar uma nova cultura que privilegie o planejamento com metas claras; (iii) motivar a participação dos servidores com um modelo de meritocracia que avalie e premie aqueles que atingirem bons resultados e; (iv) antecipar problemas e apontar soluções através do acompanhamento formal dos resultados obtidos. O que está se propondo é a implementação de um modelo de gestão de desempenho que privilegie o planejamento (metas claras) e a meritocracia como elementos centrais para a melhoria dos serviços prestados à população. A implantação bem-sucedida desse modelo de gestão de desempenho garantirá um maior comprometimento de toda a máquina pública municipal com os resultados da prefeitura, aumentando significativamente nossas chances de atingir as metas propostas à sociedade carioca durante o primeiro ano da administração" (RIO DE JANEIRO, 2013).

para o apoio administrativo. Desenvolve-se o estudo exploratório ora apresentado tendo em vista as prováveis necessidades de intervenção pedagógica associadas à implementação na rede municipal do Rio de Janeiro das políticas avaliativas e de bonificação a partir de 2009.

Para analisar os efeitos dessa política de bonificação sobre a gestão pedagógica, foram selecionadas as escolas da rede municipal do Rio de Janeiro que participaram das edições de 2009 e 2011 da Prova Brasil, perfazendo um total de 697 unidades. Dessas, 170 não foram premiadas no período e 113 foram premiadas nos dois anos. Esses dois grupos representariam, respectivamente, as escolas que não conseguiram (1º grupo) ou conseguiram (2º grupo) melhorar seus resultados no Ideb e no IDE-Rio até as metas estabelecidas pela SME/RJ para os anos de 2009 e 2010.

As escolas que atingiram suas metas na Prova Brasil 2009 foram informadas sobre a bonificação em 2010, após a divulgação dos resultados. Portanto, as respostas aos questionários contextuais de 2009 foram registradas antes de elas receberem ou serem informadas de que receberiam qualquer bonificação. Por outro lado, as respostas aos questionários contextuais da Prova Brasil 2011 foram dadas após as escolas receberem (ou não) a premiação referente às metas do Ideb 2009 em 2010 e a premiação referente às metas do IDE-Rio 2010 paga em 2011. Cabe assinalar que se trata, na verdade, de um intervalo ainda pequeno para avaliar alterações de monta em práticas escolares, o que explica o caráter apenas exploratório do estudo desenvolvido.

Foram analisadas as respostas dos diretores aos questionários da Prova Brasil de 2009 e 2011 para esses dois grupos com o objetivo de identificar alterações e/ou continuidades em práticas de gestão pedagógica relacionadas, em maior ou menor medida, a alguns aspectos da gestão mais diretamente associados ao desempenho discente. As Tabelas 3 e 4 apresentam as respostas sobre a oferta de formação continuada aos docentes dessas escolas e a participação destes nas atividades.

TABELA 3 – Premiação e Promoção de Atividade de Formação Continuada

PROMOVEU ALGUMA ATIVIDADE DE FORMAÇÃO CONTINUADA NA ESCOLA	SEM PREMIAÇÃO		PREMIADA EM 2009/2010 E 2010/2011	
	2009	2011	2009	2011
Sim	40%	55%	45%	67%
Não	57%	38%	48%	29%
Não respondeu	3%	7%	7%	4%
Total	100%	100%	100%	100%

Nota: Em 2011 esta pergunta do questionário contextual foi levemente modificada, ganhando a seguinte redação: "*Você promoveu alguma atividade de formação continuada (atualização, treinamento, capacitação, etc.) nesta escola?*".

Fonte: Elaboração das autoras com base nos dados da SME/RJ, 2010; SME/RJ, 2011; Prova Brasil 2009 e 2011.

TABELA 4 – Premiação em 2009 e 2011 e Participação Docente na Formação Continuada

PROPORÇÃO DE DOCENTES DA ESCOLA QUE PARTICIPOU DAS ATIVIDADES DE FORMAÇÃO CONTINUADA PROMOVIDAS POR VOCÊ NOS DOIS ÚLTIMOS ANOS	SEM PREMIAÇÃO		PREMIADA EM 2009/2010 E 2010/2011	
	2009	2011	2009	2011
Menos de 10%	3%	0%	2%	1%
Entre 11% e 30%	2%	4%	2%	2%
Entre 31% e 50%	6%	7%	2%	2%
Mais de 51%	37%	44%	43%	63%
Não sabe	2%	1%	2%	0%
Não respondeu	51%	44%	50%	33%
Total	100%	100%	100%	100%

Nota: Em 2011 a pergunta do questionário contextual foi levemente modificada, ganhando a seguinte redação: "*Qual foi a proporção de docentes da sua escola que participou das atividades de formação continuada promovidas por você nos últimos dois anos?*".

Fonte: Elaboração das autoras com base nos dados da SME/RJ, 2010; SME/RJ, 2011; Prova Brasil 2009 e 2011.

De maneira geral, a oferta de formação continuada nas escolas parece ter crescido, em particular nas unidades premiadas nos dois anos. A Tabela 4 mostra que a proporção de docentes que participa das atividades de formação continuada também aumentou em ambas as situações, mas ocorreu de forma mais acentuada nas escolas premiadas. Pode-se levantar a hipótese de que a articulação das políticas de avaliação e bonificação tenha incrementado tanto o interesse dos

professores em participar de atividades de formação, como dos gestores escolares em propiciá-las a sua equipe, como estratégia de consolidação ou garantia da continuidade do sucesso expresso no alcance das metas.

Outro aspecto afeto à gestão da escola são os critérios usados para admissão de alunos na época da matrícula (ou fora dela). Apesar do sistema informatizado de matrícula adotado pela SME/RJ,[6] que supostamente uniformizaria esses processos no âmbito da rede municipal, parece haver uma alteração nesse período, como indica a Tabela 5.

TABELA 5 – Critério de Admissão de Alunos

CRITÉRIO UTILIZADO PARA A ADMISSÃO DE ALUNOS NA ESCOLA	SEM PREMIAÇÃO		PREMIADA EM 2009/2010 E 2010/2011	
	2009	2011	2009	2011
Prova de seleção*	-	0%	-	0%
Sorteio*	-	1%	-	0%
Local de moradia	3%	4%	45%	2%
Prioridade por ordem de chegada	31%	14%	0%	23%
Outro critério	17%	33%	20%	45%
Não existe critério preestabelecido	47%	42%	32%	27%
Não respondeu	3%	6%	4%	3%
Total	100%	100%	100%	100%

Nota (*): Alternativas de resposta que só foram disponibilizadas no questionário de 2011.

Fonte: Elaboração das autoras com base nos dados da SME/RJ, 2010; SME/RJ, 2011; Prova Brasil 2009 e 2011.

6 As matrículas são efetuadas em duas etapas na rede pública municipal do Rio. A primeira corresponde à pré-matrícula e é realizada pela internet em um *site* exclusivo para essa finalidade. Nesse momento, são efetuadas pré-matrículas iniciais, transferências internas (dentro da rede) e externas (de outras redes para a rede do Rio). No ato da inscrição, é necessário indicar entre três e cinco escolas de interesse. Após a confirmação da inscrição, por SMS, e-mail ou carta, é preciso comparecer à escola com a documentação solicitada para confirmar a matrícula no período designado pela SME. Para aqueles que perderam a primeira etapa, são oferecidas vagas remanescentes em um momento posterior. Esse mecanismo de matrícula, adotado desde 2010, estabeleceu outra relação entre a escola e as famílias. Anteriormente, a matrícula ocorria presencialmente, em uma única etapa, em polos de matrícula, na presença de diretores de escolas que integravam uma determinada região do bairro e de Coordenadoria Regional de Educação, alocados numa escola de fácil acesso para a população. O que determinava a matrícula, até então, era, presumivelmente, a ordem de chegada.

Observam-se diferenças interessantes entre os critérios assinalados pelos diretores dos dois grupos de escolas nos anos analisados. Os diretores das escolas não premiadas afirmam majoritariamente não possuir critérios preestabelecidos e, secundariamente, passam da aparente preferência pelo critério "Prioridade por ordem de chegada" em 2009 para "Outro critério" em 2011.

Por outro lado, os diretores de escolas premiadas parecem estar no movimento de mudança em direção à adoção de "Outro critério" (de 20% para 45%). Ainda nessas escolas, o critério que praticamente deixa de ser considerado, segundo

as respostas de seus diretores, é o de proximidade da moradia dos alunos e famílias, enquanto que o ingresso por ordem de chegada passa a ser uma opção comum e atinge 23%. Esses resultados parecem reforçar a ideia de que a premiação – divulgada pela mídia e junto às famílias – tenha influído positivamente na percepção externa dessas escolas, que passam a ser mais atrativas, em especial para famílias bem informadas, o que já indica a posse de certo capital social e informacional.

Caberia questionar se o aumento expressivo da menção a "outro critério" poderia registrar a prevalência de uma maior seletividade não explícita, possivelmente contribuindo para a desigualdade de oportunidades educacionais analisadas por Costa e Koslinski (2011, p. 252), entre outros, na perspectiva de um "quase-mercado" educacional.

A Tabela 6, apresentada a seguir, mostra a percepção dos gestores escolares sobre a oferta de vagas em sua escola.

TABELA 6 – Oferta de vagas

SITUAÇÃO DA OFERTA DE VAGAS NA ESCOLA NESTE ANO	SEM PREMIAÇÃO		PREMIADA EM 2009/2010 E 2010/2011	
	2009	2011	2009	2011
Após a matrícula, a escola ainda tinha vagas disponíveis	42%	41%	37%	33%
A procura preencheu todas as vagas disponíveis	15%	21%	20%	12%
A procura foi maior que as vagas oferecidas	27%	22%	29%	33%
A procura por vagas superou em muito a oferta	11%	13%	11%	20%
Não respondeu	5%	3%	3%	2%
Total	100%	100%	100%	100%

Fonte: Elaboração das autoras com base nos dados da SME/RJ, 2010; SME/RJ, 2011; Prova Brasil 2009 e 2011.

Nos dois anos analisados, a oferta de vagas é maior do que a demanda em mais de 40% das escolas não premiadas, um sinal de baixa procura das famílias por essas escolas, o que pode ser interpretado como expressão também de uma percepção negativa ou não muito positiva da qualidade de seus serviços educacionais. É interessante observar que,

passados dois anos da implantação das políticas e da crescente divulgação de *rankings* com os resultados, bem como de sua socialização também com as famílias, parece ter havido pequenas alterações na demanda por vagas nessas escolas.

Já nas escolas premiadas, a proporção dos diretores que afirma que a demanda foi maior do que a oferta ("foi maior" e/ou "superou em muito") subiu 13 pontos percentuais (de 40% para 53%). Portanto, seja por influência direta da divulgação dos resultados e premiações, seja pela circulação dessas e outras informações relacionadas ao trabalho dessas escolas, parece tender a crescer uma seletividade maior por parte das famílias, dando preferência às escolas cujos resultados atingiram as metas estabelecidas pela SME/RJ.

Finalmente, a Tabela 7, apresentada a seguir, registra as respostas dos diretores dos dois grupos de escolas sobre os critérios utilizados na formação das turmas – tema de expressiva relevância na agenda da gestão pedagógica, que tem sido apontado pelas pesquisas nacionais e internacionais no campo da sociologia da educação como uma das ferramentas da gestão escolar que podem influenciar os resultados do alunado (ALVES; FRANCO, 2008).

TABELA 7 - Critérios usados na formação de turmas

CRITÉRIO UTILIZADO PARA A FORMAÇÃO DAS TURMAS NA ESCOLA	SEM PREMIAÇÃO		PREMIADA EM 2009/2010 E 2010/2011	
	2009	2011	2009	2011
Homogeneidade quanto à idade	79%	72%	78%	74%
Homogeneidade quanto ao rendimento escolar	3%	2%	3%	5%
Heterogeneidade quanto à idade	1%	4%	2%	1%
Heterogeneidade quanto ao rendimento	7%	12%	13%	15%
Não houve critério	4%	6%	1%	4%
Não respondeu	6%	3%	4%	2%
Total	100%	100%	100%	100%

Fonte: Elaboração das autoras com base nos dados da SME/RJ, 2010; SME/RJ, 2011; Prova Brasil 2009 e 2011.

Em 2009 e 2011, não foram observadas alterações relevantes nos processos de formação das turmas de alunos nos dois grupos de escolas. Mais de 70% dos diretores dos dois grupos, nos dois anos, assinalaram que o critério utilizado nas escolas que dirigem é o de "homogeneidade quanto à idade". Ainda assim, nota-se uma pequena queda dessa alternativa no período para os dois grupos, dando lugar, aparentemente, a critérios que valorizam também o rendimento escolar. Vale ressaltar que o ano de 2009 marca o fim da organização escolar por ciclos no município do Rio de Janeiro.

Finalmente, a oferta de programa de apoio ou reforço de aprendizagem para os alunos é declarada como existente e regular por mais de 90% dos diretores dos dois grupos. Observa-se, porém, uma pequena elevação desse percentual entre 2009 e 2011. Esse aumento poderia ser explicado pela adoção pela SME/RJ de uma política de incentivo ao estágio remunerado de universitários para o atendimento às crianças com dificuldade de aprendizagem na escola.

A breve incursão exploratória realizada nos dados dos questionários contextuais da Prova Brasil 2009 e 2011 desses dois conjuntos de escolas, consecutivamente premiadas (ou não) pelo Prêmio Anual de Desempenho que constitui a política de bonificação da SME/RJ, mostrou alguns traços indicativos de tendências de alteração nas práticas de gestão pedagógica, mais acentuados no grupo de escolas premiadas. Esses indícios, embora não se mostrem em geral estatisticamente significativos, apontam para o interesse no acompanhamento desses processos de implementação de políticas e para a avaliação periódica de suas consequências no nível das escolas e de sua gestão, que merecem ser aprofundados em novas pesquisas.

CONSIDERAÇÕES FINAIS

Este artigo apresentou três movimentos complementares e independentes de investigação sobre alguns dos efeitos das políticas de avaliação e responsabilização adotadas desde 2009 na rede municipal de ensino do Rio de Janeiro. Priorizou-se, nos três relatos de pesquisa, a discussão sobre as

demandas e desafios da gestão escolar frente às estratégias de avaliação externa e às políticas de responsabilização adotadas pela rede municipal do Rio de Janeiro.

Os três trabalhos ora integrados apresentam alguns dos desafios e tensões cada vez mais presentes na pauta da gestão escolar, alçando ao centro da cena as responsabilidades e as opções pedagógicas estratégicas dos gestores escolares, além das limitações administrativas e político-legais de sua intervenção.

O estudo de caso de Oliveira (2012b) mostra que o excesso de avaliações e os controles a elas relacionados sobrecarregam a gestão, configurando uma nova burocracia na rotina escolar. No caso do município do Rio de Janeiro, além das demandas da aplicação das avaliações nacionais, da Prova Rio e das Avaliações Bimestrais, a equipe de gestão da escola passa a controlar e reportar à SME as notas dos alunos nas avaliações bimestrais a que são também submetidos. Além desse trabalho extra, o papel mediador da equipe de gestão entre a SME/RJ e os professores se acentua e tende a tornar-se um fator determinante para a sustentação do foco na aprendizagem do aluno, que poderá ser também reforçado pela política de bonificação na medida de sua eficácia.

Restaria avaliar como essas demandas têm sido incorporadas por coordenadores e diretores nas escolas e acomodadas em suas condições de trabalho. No que tange à equipe de trabalho disponível na escola, os dados apresentados no segundo estudo denunciam a insuficiência de pessoal administrativo em parte das escolas da rede. A manutenção da precariedade da estrutura de pessoal de apoio administrativo nas escolas torna ainda mais complexo o trabalho da equipe gestora, pressionada a responder pelos resultados alcançados por seus alunos nos testes em larga escala. Em que pese a alteração para melhor do quadro docente, possibilitada pelo número expressivo de concursos públicos para lotação no cargo de professores, realizados no período 2009/2011, conforme assinalado, caberia indagar sobre sua integração efetiva ao trabalho das escolas no âmbito das políticas de avaliação e responsabilização vigentes.

Novas tensões são trazidas para as relações entre professores e equipe de gestão. Nesse contexto em que os resultados acadêmicos dos alunos são medida para a distribuição de prêmios e conferem visibilidade para a escola, corre-se o risco de que os profissionais se sintam pressionados a apresentarem resultados satisfatórios. Esse sentimento também é verdadeiro nas relações entre os professores (especialmente nas séries avaliadas) e seus coordenadores e diretores, considerando as expectativas em relação a seu trabalho. A disputa por melhores resultados muda a estrutura das relações pessoais, como sugere Wilmmott (1993, p. 222[7] apud BALL, 2001, p. 107): "exige-se dos empregados, individual e coletivamente, que simultaneamente reconheçam e se responsabilizem pela relação entre a segurança do seu posto de trabalho e a sua contribuição para a competitividade dos bens e dos serviços que produzem".

7 WILLMOTT, H. Strength is ignorance; slavery is freedom: managing culture in modern Organizations. *Journal of Management Studies*, v. 30, n. 4, p. 215-252, 1993.

Os dados levantados sobre os diretores das escolas premiadas ou não premiadas pela política de bonificação da SME/RJ revelaram algumas alterações nas respostas desses profissionais no período observado (2009-2011), o que parece indicar uma maior preocupação ou foco nos resultados alcançados por sua escola, expressa, por exemplo, no crescimento da oferta de atividades de formação para os docentes. Houve uma alteração significativa nos percentuais referentes à forma de ingresso do aluno na escola, especialmente nas unidades premiadas. Tal alteração, reduzindo a matrícula por "local de moradia" e aumentando a matrícula por "outro critério", poderia expressar um procedimento de seletividade de alunos, derivado da crescente demanda pelas escolas premiadas. A provisoriedade e o caráter exploratório dos resultados dos levantamentos que constituem o terceiro movimento de investigação apresentado neste artigo apontam a necessidade de seu aprofundamento futuro, sugerindo pistas instigantes para novas pesquisas.

Ainda que a título de conclusão provisória, pode-se inferir das análises apresentadas a necessidade de preparação destes profissionais – diretores e coordenadores escolares – para lidar com as demandas colocadas para a gestão escolar que passam a integrar seu cotidiano na última década. Entre elas estão a capacidade de leitura e interpretação de índices e resultados educacionais; a necessidade permanente

de intermediação entre os órgãos centrais e os profissionais da escola – especialmente os professores –; a interpretação das demandas oriundas das políticas avaliativas a fim de que possam ser utilizadas também como ferramentas de gestão e planejamento pedagógico; a manutenção de um clima escolar favorável, gerenciando as tensões que passam a surgir nas relações escolares em virtude das cobranças de resultados. Acredita-se que tais demandas devam ocupar um espaço cada vez mais relevante e estratégico tanto na pauta do campo acadêmico como no terreno da formulação, implementação e avaliação das políticas públicas de educação.

REFERÊNCIAS

ALVES, Fátima Cristina. Mapeamento das políticas de escolha de diretores da escola e de avaliação na rede pública das capitais brasileiras. *Revista Brasileira de Estudos Pedagógicos*, Brasília, v. 90, n. 224, p. 71-86, jan./abr. 2009.

ALVES, Maria Teresa G.; FRANCO, Creso. A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. In: BROOKE, Nigel; SOARES, José Francisco (Org.). *Pesquisa em eficácia escolar*: origem e trajetórias. Belo Horizonte: Editora UFMG, 2008.

ALVES, Maria Teresa G.; SOARES, José Francisco. Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos. *Educação em Revista*, Belo Horizonte, v. 45, p. 25-58, 2007.

ANDRADE, Eduardo de C. School Accountability no Brasil: experiências e dificuldades. *Revista de Economia Política*, v. 28, n. 3, p. 443-453, jul./set. 2008. Disponível em: <http://www.scielo.br/pdf/rep/v28n3/a05v28n3 >. Acesso em: 25 nov. 2011.

BALL, Stephen J. Diretrizes políticas globais e relações políticas locais em educação. *Currículo Sem Fronteiras*, v. 1, n. 2, p. 99-116, jul./dez. 2001. Disponível em: <http://www.curriculosemfronteiras.org/vol1iss2articles/ball.pdf >. Acesso em: 12 maio 2010.

BRASIL. Constituição Federal, 1988. Disponível em: <http://www.planalto.gov.br>. Acesso em: 12 jun. 2013.

_____. *Decreto n. 6.094, de 24 de abril de 2007*. Dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação, pela União Federal, em regime de colaboração com Municípios, Distrito Federal e Estados, e a participação das famílias e da comunidade, mediante programas e ações de assistência técnica e financeira, visando a mobilização social pela melhoria da qualidade da educação básica. Brasília, DF, 2007. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6094.htm>. Acesso em: 2 dez. 2013.

_____. *Lei de Diretrizes e Bases da Educação Nacional*, Lei 9.394/96. Brasília, DF, 1996. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9394.htm>. Acesso em 10 jun. 2013.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados Prova Brasil 2007*. Mídia eletrônica. Brasília, DF, 2007. Disponível em: <http://portal.inep.gov.br/basica-levantamentos-acessar>. Acesso em: 10 fev. 2014.

_____ *Microdados Prova Brasil 2009*. Mídia eletrônica. Brasília, DF, 2009. Disponível em: <http://portal.inep.gov.br/basica-levantamentos-acessar>. Acesso em: 10 fev. 2014.

_____. *Microdados Prova Brasil 2011*. Mídia eletrônica. Brasília, DF, 2011. Disponível em: <http://portal.inep.gov.br/basica-levantamentos-acessar>. Acesso em: 10 fev. 2014.

BROOKE, Nigel. Responsabilização educacional no Brasil. *Sísifo, Revista Iberoamericana de Evaluación Educativa*, v. 1, n. 1., p. 94-109, 2008. Disponível em: <http://www.rinace.net/riee/numeros/vol1-num1/art7_port.html>. Acesso em: 20 out. 2012.

COSTA, Marcio da; KOSLINSKI, Mariane C. Quase-mercado oculto: disputa por escolas "comuns" no Rio de Janeiro. *Cadernos de Pesquisa*, São Paulo, v. 41, n. 142, jan./abr. 2011.

FERNANDES, Reynaldo; GREMAUD, Amaury. Qualidade da educação: avaliação, indicadores e meta. In: VELOSO, Fernando et al. (Org.). *Educação básica no Brasil:* construindo o país do futuro. Rio de Janeiro: Elsevier, 2009. p. 213-238. Disponível em: <www.cps.fgv.br/ibrecps/rede/seminario/reynaldo_paper.pdf>. Acesso em: 3 set. 2012.

FUNDAÇÃO JOÃO GOULART. *Administração, concursos e cursos*. Disponível em: <http://concursos.rio.rj.gov.br/indexportal.htm>. Acesso em: 16 jul. 2014.

GOMES, Adailda; REGIS, André. Desempenho e infraestrutura: mapeamento das escolas públicas da região metropolitana do Rio de Janeiro. In: CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 3., 2012, Zaragoza, Espanha. *Anais*... Zaragoza, Espanha: Anpae, 2012. (Série Cadernos Anpae, v. 15). Disponível em: <http://www.anpae.org.br/iberoamericano2012/Trabalhos/AdaildaGomesDeOliveira_res_int_GT1.pdf >. Acesso em: 5 out. 2013.

LIMA, Maria de Fátima de M.; ALCÂNTARA, Guilherme de; ALMEIDA, Mônica A. de O. O que potencialmente afeta a gestão escolar? Percepções de diretores de duas redes públicas de ensino do estado do Rio de Janeiro. In: REUNIÃO NACIONAL DA ANPED, 36., 2013, Goiânia. *Anais*... Sistema Nacional de Educação e Participação Popular: desafios para as políticas educacionais. Goiânia: Anped, 2013.

OLIVEIRA, Ana Cristina P. *Política pública e prática docente*: quando e como dialogam no espaço escolar: a experiência de uma escola do Rio de Janeiro. 2012. 205 f. Dissertação (Mestrado em Educação) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Educação, Rio de Janeiro, 2012a.

_____. Política pública e gestão escolar: um estudo de caso no Rio de Janeiro. In: CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 3., 2012, Zaragoza. *Cadernos Anpae*... Timbaúba: Anpae, 2012b.

RIO DE JANEIRO (Município). Secretaria Municipal da Casa Civil. *Acordo de Resultados*. Disponível em: <http://www.rio.rj.gov.br/web/cvl/exibeconteudo?id=1820527>. Acesso em: 25 jul. 2014.

_____. *Acordo de Resultados*. 24 jun. 2013. Disponível em: <http://www.rio.rj.gov.br/web/cvl/acordo-de-resultados>. Acesso em: 25 jul. 2014.

RIO DE JANEIRO (Município). Secretaria Municipal de Educação. *Prefeito e Secretária de Educação anunciam as escolas que receberão o Prêmio Anual de Desempenho*. 2010. Disponível em: <http://www.rio.rj.gov.br/web/sme/exibeconteudo?id=1011528>. Acesso em: 30 nov. 2013.

_____. *Secretaria de Educação anuncia os resultados do IDE-Rio 2010*. 2011. Disponível em: <http://www.rio.rj.gov.br/web/sme/exibeconteudo?id=1772816>. Acesso em 30 nov. 2013.

SAMMONS, Pam. As características-chave das escolas eficazes. In: BROOKE, Nigel; SOARES, José Francisco (Org.). *Pesquisa em eficácia escolar*: origem e trajetórias. Belo Horizonte: Editora UFMG, 2008.

SOARES, José Francisco. Melhora do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de Pesquisa*, v. 37, n. 130, p. 135-160, jan./abr. 2007.

SOARES, Sergey; SÁTYRO, Natália. *O impacto da infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental* – 1998 a 2005. Rio de Janeiro: Ipea, maio 2008.

SOARES, Tufi M.; TEIXEIRA, Lúcia Helena G. Efeito do perfil do diretor na gestão escolar sobre a proficiência do aluno. *Estudos em Avaliação Educacional*, São Paulo, v. 17, n. 34, maio/ago. 2006.

SOARES, Tufi M. et al. A expectativa do professor e o desempenho dos alunos. *Psicologia: Teoria e Pesquisa*, Brasília, v. 26, n. 1, p. 157-170, jan./mar. 2010.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. Faculdade de Educação. GAME. *A avaliação externa como instrumento da gestão educacional nos estados*: relatório final. Fundação Victor Civita, ago. 2011.

CYNTHIA PAES DE CARVALHO
Doutora em Educação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Professora adjunta do Departamento de Educação da PUC-Rio, Rio de Janeiro, Rio de Janeiro, Brasil
cynthiapaesdecarvalho@puc-rio.br

ANA CRISTINA PRADO DE OLIVEIRA
Doutoranda do Programa de Pós-Graduação em Educação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Rio de Janeiro, Brasil
ana.prado.oliveira@gmail.com

MARIA DE FÁTIMA MAGALHÃES DE LIMA
Doutoranda do Programa de Pós-Graduação em Educação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Professora da Secretaria Municipal de Educação do Rio de Janeiro (SME/RJ), Rio de Janeiro, Rio de Janeiro, Brasil
fatima_mlima@superig.com.br

Recebido em: JULHO 2014
Aprovado para publicação em: NOVEMBRO 2014

COBRANÇAS, ESTRATÉGIAS E "JEITINHOS": AVALIAÇÕES EM LARGA ESCALA NO RIO DE JANEIRO

RODRIGO ROSISTOLATO
ANA PIRES DO PRADO
SILVINA JULIA FERNÁNDEZ

RESUMO

O objetivo deste artigo é descrever e analisar visões de gestores da rede municipal do Rio de Janeiro sobre as avaliações em larga escala e seus reflexos no cotidiano escolar. Os dados foram mapeados em quatro grupos focais, dois com diretores e dois com profissionais da gestão central e intermediária da Secretaria Municipal de Educação. Demonstraremos que, apesar de as avaliações externas ainda se apresentarem como eventos extraordinários ao cotidiano escolar, há uma tendência à utilização dos dados delas resultantes como norteadores das práticas de gestão, mesmo com o desconhecimento técnico das avaliações. Foi possível perceber um "efeito dominó". Trata-se de um conjunto de cobranças sucessivas aos agentes de acordo com sua posição na hierarquia da gestão, que induz estratégias para melhorar o desempenho institucional, assim como "jeitinhos" e estratagemas orientados pela expectativa de aumentar os índices das escolas.

PALAVRAS-CHAVE AVALIAÇÃO DOS SISTEMAS DE ENSINO • COTIDIANO ESCOLAR • PRÁTICAS DE GESTÃO • RIO DE JANEIRO.

RESUMEN

El objetivo de este artículo es describir y analizar visiones de gestores de la red municipal de Rio de Janeiro sobre las evaluaciones en larga escala y sus reflejos en el cotidiano escolar. Los datos fueron mapeados en cuatro grupos focales, dos con directores y dos con profesionales de la gestión central e intermedia de la Secretaría Municipal de Educación. Demostraremos que, a pesar de que las evaluaciones externas todavía se presentan como eventos extraordinarios al cotidiano escolar, hay una tendencia a la utilización de los datos resultantes como ejes de las prácticas de gestión, incluso con el desconocimiento técnico de las evaluaciones. Fue posible percibir un "efecto dominó". Se trata de un conjunto de exigencias sucesivas a los agentes de acuerdo a su posición en la jerarquía de la gestión, que induce estrategias para mejorar el desempeño institucional, así como "jeitinhos" *y estratagemas orientados por la expectativa de aumentar los índices de las escuelas.*

PALABRAS-CLAVE EVALUACIÓN DE LOS SISTEMAS DE ENSEÑANZA • COTIDIANO ESCOLAR • PRÁCTICAS DE GESTIÓN • RIO DE JANEIRO.

ABSTRACT

The objective of this article is to describe and analyze views of managers of the municipal network of Rio de Janeiro regarding the large-scale evaluations and their consequences in school daily life. The data were divided into four groups, two with directors and two with central and mid-level management professionals in the Municipal Department of Education. We will demonstrate that, despite external evaluations still being extraordinary events in school daily life, there is a tendency to use their resulting data as guidelines for management practices, even with technical lack of knowledge about the evaluations. It was possible to notice a "domino effect". This involves a set of successive demands on agents, according to their position in the management hierarchy, which induces strategies to improve institutional performance, as well as "jeitinhos" *and stratagems guided by the expectation of increasing the indices of the schools.*

KEYWORDS EVALUATION OF INSTRUCTIONAL SYSTEMS • SCHOOL DAILY LIFE • MANAGEMENT PRACTICES • RIO DE JANEIRO.

INTRODUÇÃO

A proposta deste texto é descrever e analisar visões de gestores da rede municipal do Rio de Janeiro sobre as avaliações em larga escala e seus reflexos no cotidiano escolar. Trabalharemos, especificamente, com três questões. Indicaremos as opiniões dos profissionais sobre as avaliações em larga escala, suas visões sobre as demandas presentes no cotidiano escolar após a consolidação das avaliações e as estratégias e as resistências relacionadas ao desempenho das escolas. Analisaremos as dissonâncias existentes entre as perspectivas apontadas pelos diferentes níveis de gestão.

A investigação faz parte das problemáticas desenvolvidas no Laboratório de Pesquisa em Oportunidades Educacionais (LaPOpE) e está diretamente relacionada ao debate nacional sobre avaliação de sistemas educacionais. Demonstraremos que, embora as avaliações em larga escala ainda sejam eventos extraordinários[1] ao cotidiano escolar, tendem a se consolidar como um dos elementos norteadores das práticas de gestão e docência. Há, no entanto, diferentes níveis de entendimento e ação. Ao mesmo tempo em que

1 Para um debate sobre a oposição entre eventos ordinários e extraordinários à vida cotidiana, ver Lévi-Strauss (1981).

mapeamos estratégias com vistas a ampliar o desempenho das escolas, percebemos "jeitinhos"[2] e estratagemas orientados pela expectativa exclusiva de aumentar os índices das escolas.

É possível dizer que o cenário plural que será descrito está associado, entre outros fatores, às dificuldades enfrentadas por gestores para entender os aspectos técnicos relacionados às avaliações externas e seus índices. Nesse sentido, podemos indicar que a demanda por utilização de dados para organização de processos de ensino e planejamento escolar cresceu. Por outro lado, a oferta de formação especializada manteve-se limitada ou quase inexistente. Talvez esse seja um dos principais dilemas relacionados às políticas de avaliação de sistemas educacionais no Brasil.

Em outros contextos, a percepção de que educadores devem entender dados educacionais produzidos em larga escala cresceu em concomitância com a ampliação dos dados disponíveis (MARSH; PANE; HAMILTON, 2006; MANDINACH; GUMMER, 2013). Nesses contextos, a discussão gira em torno da necessidade de ampliação da formação para uso de dados, inclusive na formação de educadores. No caso brasileiro, especificamente no Rio de Janeiro, demonstraremos que há desconhecimento, dúvidas e ausência de preocupações relacionadas à formação para uso dos dados, o que, de certa forma, pode estimular estratégias antirrepublicanas[3] e "jeitinhos" diversos para produzir índices artificiais de sucesso, sem efeito para uma abordagem fidedigna da situação dos sistemas educacionais no Brasil.

2 Para um debate sobre a "cultura do jeitinho" nas relações dos agentes sociais com o Estado republicano no Brasil, ver Da Matta (1983). Ver também Barbosa (1982).

3 Consideramos estratégias antirrepublicanas quaisquer práticas sociais que visem burlar as regras universais formalmente estabelecidas pelo Estado Democrático de Direito.

METODOLOGIA

Durante o ano de 2012, realizamos dois grupos focais com gestores da rede municipal de educação da cidade do Rio de Janeiro. Os grupos foram compostos por profissionais de escolas de alto desempenho e de baixo desempenho separadamente. A definição de alto e baixo desempenho levou em consideração as notas obtidas pelas escolas na Prova Brasil de 2009.[4] Como as notas vão de 0 a 10 em Língua Portuguesa e Matemática, dividimos as escolas em tercis.[5] Classificamos

4 Utilizamos a nota padronizada da Prova Brasil.

5 Todos os cálculos relacionados à amostra foram realizados pela socióloga Mariane Campelo Koslinski.

as escolas do primeiro tercil como de alto desempenho e as do terceiro tercil como de baixo desempenho. Na sequência, selecionamos as escolas e convidamos os gestores.[6] A escolha não foi totalmente aleatória porque priorizamos as escolas localizadas na região da quarta Coordenadoria Regional de Educação (CRE).[7] Dentre os convidados das escolas de alto desempenho, apenas três compareceram. Tivemos a participação de nove gestores no grupo focal das escolas de baixo desempenho.[8]

Após a realização dos grupos focais, iniciamos um curso de formação para gestores da rede municipal de ensino do Rio de Janeiro dedicado às avaliações em larga escala. Para a primeira turma do curso, selecionamos profissionais alocados na gestão central e na gestão intermediária da Secretaria Municipal de Educação (SME). Decidimos realizar, antes do início dos encontros, um grupo focal com a equipe central e outro com integrantes da gestão intermediária. A proposta era mapear suas percepções sobre as avaliações em larga escala antes do início das aulas e seguindo o mesmo roteiro do grupo focal realizado com gestores. Todos optaram por participar da seleção para o curso e foram voluntários para a realização do grupo focal. O critério de seleção dos profissionais foi sua presença em cargos da gestão central e/ou intermediária da SME. Selecionamos um profissional de cada CRE – são 11 ao todo – e quatro profissionais da gestão central.[9]

O roteiro dos grupos focais foi o mesmo utilizado com os gestores de escolas de alto e baixo desempenho. Ele foi construído com o objetivo de mapear as percepções, o conhecimento técnico, a experiência, a relação entre avaliação e funcionamento da escola, as consequências/mudanças no cotidiano das escolas e as opiniões sobre a divulgação pública dos resultados. Neste artigo, analisaremos os dados produzidos nos quatro grupos focais, considerando as convergências e as divergências presentes nas falas de gestores que trabalham em escolas, na gestão central e/ou intermediária.

6 Seguindo os critérios mencionados, foram selecionados para o grupo focal de alto desempenho 11 gestores e para o grupo de baixo desempenho, 16 gestores.

7 A rede municipal do Rio de Janeiro é dividida em coordenadorias regionais de educação (CREs). Priorizamos a quarta CRE em virtude da diversidade socioeconômica presente na região e da pluralidade de escolas com alto e baixo desempenho.

8 As justificativas para não participar das atividades da pesquisa foram, em geral, relacionadas a outro compromisso já agendado na escola ou à falta de pessoal que possibilitasse a ausência do gestor na unidade escolar.

9 Não podemos descrever cargos e funções porque colocaríamos o anonimato dos participantes em risco.

O DEBATE ACADÊMICO SOBRE AVALIAÇÃO EDUCACIONAL NO BRASIL

As avaliações em larga escala surgem, nos anos 1990, como uma das principais novidades no campo educacional brasileiro. Naquele momento, os debates sobre a universalização do acesso à educação básica estavam perdendo força em virtude da própria universalização, e as discussões sobre a equalização da distribuição de conhecimento escolar ganhavam território. Havia clareza sobre a necessidade de avaliar os sistemas educacionais, o que motivou as decisões técnicas e políticas que deram origem ao Sistema de Avaliação da Educação Básica (Saeb).

O Saeb oferecia uma visão panorâmica do desempenho dos sistemas educacionais. Tratava-se do primeiro passo da construção de bases de dados que permitissem a formulação de políticas educacionais e, simultaneamente, fomentassem o debate acadêmico sobre educação no Brasil. O sistema foi modificado com a criação da Prova Brasil, que ampliou o escopo da avaliação e, consequentemente, o volume de dados disponíveis. Essa mudança expandiu as possibilidades analíticas, além de exacerbar os debates públicos sobre as avaliações em larga escala no Brasil.

É possível afirmar que as avaliações transitaram, em um período curto de tempo, do total desconhecimento para o centro da arena pública de debates sobre a educação no Brasil (ROSISTOLATO; PRADO, 2014; BONAMINO, 2002, BONAMINO; SOUSA, 2012). Essa transição ocorreu em duas vias paralelas. Por um lado, há intenso debate sobre os possíveis impactos das avaliações no cotidiano das escolas, incluindo gestão e trabalho docente (BONAMINO; COSCARELLI; FRANCO, 2002; SOUSA, 2003; CARVALHO; MACEDO, 2011; BONAMINO; SOUSA, 2012). Na outra via, seguem as pesquisas que discutem os desenhos técnicos dos sistemas de avaliação e os dados que eles podem produzir (ALVES, 2007; SOARES, 2009).

Esses dois conjuntos de pesquisas agregam ramificações que envolvem estudos sobre efeito escola, perfis de gestores e modelos de gestão, avaliação institucional e implementação de políticas educacionais (MARTINS; SOUSA, 2012). Trata-se de um campo dinâmico em que investigações

sobre a concepção e a estrutura dos sistemas de avaliação (BONAMINO; SOUSA, 2012; FRANCO; BONAMINO, 2001) convivem com pesquisas claramente orientadas pela negação das avaliações externas como caminho para a equalização da distribuição de conhecimento escolar no Brasil (FREITAS, 2012). Essa pluralidade de interpretações revela o dissenso ainda presente nas reflexões sobre sistemas de avaliação em larga escala no Brasil.

Nesse conjunto de pesquisas, há duas temáticas convergentes. Trata-se do debate sobre o uso de indicadores educacionais no cotidiano escolar – e nas instâncias intermediárias e centrais de gestão – e da formação para o uso. Sobre a primeira temática, desde 2009, observa-se na produção acadêmica brasileira uma reflexão sobre a compreensão e o uso dos indicadores educacionais nas escolas e/ou nas redes de ensino. Argumenta-se que há um aumento das avaliações em larga escala no país, mas ainda são reduzidas sua compreensão e utilização pelos agentes educacionais.[10] Sobre a segunda temática, é possível afirmar que ainda é pouco explorada no Brasil, embora existam debates consolidados no cenário internacional.[11] Em tese, as discussões sobre avaliação em larga escala deveriam incluir propostas claras para a divulgação dos resultados e incentivo ao uso, incluindo cursos de formação e incorporação desse debate na graduação de gestores e professores. No entanto, essas reflexões não estão, necessariamente, articuladas.

10 Para a produção nessa temática, ver Silva *et al.* (2013), Mesquita (2012), Pimenta (2011) e Fernandes (2010).

11 Esse debate pode ser visto em Mandinach, Honey e Light (2006).

Brooke e Cunha (2011) apontam que na atualidade a maior dificuldade está no uso de indicadores para a orientação do trabalho de professores. No campo da gestão, porém, há ampliação de políticas de gestão orientadas e/ou fundamentadas por indicadores educacionais. O destaque dos autores é interessante porque permite perceber algum distanciamento entre professores e gestores quando o tema é a utilização dos indicadores produzidos pelos sistemas de avaliação educacional. Em princípio, gestores e professores seriam mutuamente dependentes, mas na prática essa interdependência pode não ocorrer.

Fontanive (2005) já indicava as possibilidades de uso pedagógico dos testes e alguns dilemas presentes na disseminação

dos resultados produzidos por avaliações em larga escala. A autora defendia explicitamente a divulgação pública e a popularização do uso das avaliações em larga escala. Ela frisava a possibilidade de ampliação de relações democráticas e equânimes na distribuição de conhecimento escolar com a divulgação e o consequente entendimento dos dados. Também apontava alguns problemas que poderiam ser causados pela ausência de conhecimento técnico, como a ênfase no treinamento de estudantes e a ampliação da competitividade negativa entre as escolas na mesma rede de ensino.

Em 2013, a Fundação Carlos Chagas publicou os resultados de uma pesquisa sobre os usos da avaliação externa por equipes gestoras e profissionais docentes (SILVA et al., 2013). Os dados indicam uma série de usos e também "abusos" relacionados à ênfase em simulados e preparação para provas, além de preocupações referentes à possibilidade de as avaliações ampliarem a ênfase em classificações hierarquizadas de escolas e alunos. O relatório também aponta a pluralidade de interpretações relacionadas aos índices educacionais e, principalmente, ao tipo de resposta que as escolas deveriam oferecer à política educacional.

Essa pluralidade é significativa porque permite perceber um tipo específico de desconhecimento com relação às políticas de avaliação educacional. Os gestores entendem que precisam responder à avaliação melhorando o fluxo e as notas das escolas – em alguns casos a qualquer custo. Essa interpretação diverge da proposta original. Se observarmos o desenho das políticas de avaliação, a expectativa é diferente. Espera-se que os dados permitam que gestores e professores tenham um panorama do desempenho de suas escolas em mãos para que possam produzir diagnósticos relacionados às causas do desempenho e, se for o caso, trabalhar para a ampliação de práticas pedagógicas e de gestão que possam melhorar a escola, o que, consequentemente, melhorará os índices.

Neste artigo, trabalhamos com a hipótese de que esse desconhecimento está relacionado à ausência de processos de formação articulados à implementação de políticas de avaliação em larga escala e divulgação de dados educacionais.

Não há, no Brasil, uma política de formação para o uso de dados, o que contribui para a pluralidade de interpretações. Tanto aquelas relacionadas ao que se deve fazer com eles, assim como aos tipos de respostas cabíveis quando é necessário explicar para as redes de ensino, ou para o público em geral, o desempenho de uma escola.

A utilização de dados para tomada de decisões relacionadas aos sistemas educacionais tem sido discutida – no contexto internacional – como um dos focos centrais da política e da prática educacional. A proposta é que os encaminhamentos referentes aos sistemas educacionais e às escolas sejam realizados com base em dados que permitam avaliações fidedignas sobre os problemas em foco (MARSH; PANE; HAMILTON, 2006). Por isso, há intensos debates sobre a necessidade de formação inicial e continuada para gestores e professores. A ideia é permitir que ambos conheçam os dados educacionais e consigam utilizá-los em suas tomadas de decisão (MANDINACH; GUMMER, 2013). A proposta é que essa formação seja sistemática e articulada à teoria e à prática dos novos educadores.

No Brasil, ao contrário, não há debates voltados especificamente para o treinamento de profissionais para o uso de indicadores educacionais. É importante considerar que nossos sistemas de avaliação e produção de dados são recentes, assim como o debate técnico e político sobre as possibilidades e os limites desse modelo de política educacional. Talvez por isso ainda não tenhamos avançado no planejamento para o uso.

A ausência de formação, no entanto, convive com as exigências relacionadas ao uso de dados para planejamento. Por mais que pareça incongruente, é possível dizer que se espera que gestores, mesmo sem conhecimento, usem os dados para o diagnóstico e a organização de atividades escolares. Cabe, portanto, analisar os reflexos desse modelo de política educacional no cotidiano dos gestores e pensar suas respostas às novas dinâmicas trazidas pelas avaliações em larga escala e pelos dados por elas produzidos.

Nosso contexto de análise é o Rio de Janeiro. Trata-se de um município que participa da Prova Brasil e possui um

sistema próprio de avaliação das escolas, também utilizado para bonificação de professores: a Prova Rio. É possível dizer que ambas – Prova Brasil e Prova Rio – causam efeitos específicos nas gestões central e intermediária, além de atingirem diretamente as escolas e os professores. Por isso, nossas questões de pesquisa estavam relacionadas às percepções sobre as avaliações e as possibilidades de usos efetivos dos indicadores para o planejamento educacional.

PERCEPÇÕES E EXPECTATIVAS

COMO GESTORES ANALISAM E COMPREENDEM OS DADOS

As avaliações em larga escala colocam o universo da escola e dos profissionais diretamente envolvidos com elas em xeque. São eventos extraordinários ao cotidiano escolar,[12] mas, paradoxalmente, têm influenciado as interações entre os agentes presentes na escola e as relações entre as escolas e os diferentes níveis de gestão.

A progressiva acomodação dessa novidade extraordinária no cotidiano ordinário da escola e das atividades dos gestores do nível central não se reflete, necessariamente, no conhecimento das avaliações externas, de suas propostas e de seus resultados. No nível central, o conhecimento sobre as avaliações externas é restrito a um grupo que trabalha no setor responsável pelos dados e análises das avaliações. Esse grupo também é responsável pela posterior transmissão dos resultados a outros setores da SME,[13] que geralmente têm dúvidas sobre os indicadores e os resultados. Os membros da gestão central reconhecem a dificuldade para compreender e explicar as análises das avaliações externas, principalmente pelo fato de envolverem uma lógica distinta da usual, como afirma a Entrevistada 1, do grupo focal de gestores da SME: "Esse negócio que o professor fala, uma acertou vinte e a outra acertou vinte e não ficaram com a mesma nota. Aí, explicar isso para o professor é difícil, que a lógica é outra, que não são as mesmas questões, enfim. Difícil explicar."

Na gestão intermediária, nas CREs, há uma dificuldade de compreender como se obtém determinado resultado e

12 A oposição entre os universos ordinário e extraordinário é um recurso analítico utilizado para destacar as diferenças entre as atividades mais cotidianas – da ordem e previsíveis – e as atividades que problematizam o cotidiano – extraordinárias e nada previsíveis (ROSISTOLATO; PRADO, 2014).

13 Neste artigo, optamos por não descrever detalhadamente o organograma de gestão da SME. O objetivo é preservar a identidade dos gestores que participaram dos grupos focais.

quais elementos entram no cálculo dos indicadores educacionais, conforme indicam as falas do grupo focal:

> *Entrevistada 2: Essa fórmula aí é complicado. A gente tenta decifrar.*
> *Entrevistada 3: Eu já tentei decifrar, mas é complicado. A gente sabe alguns elementos que compõem.*
> *Mediadora: Quais são esses elementos?*
> *[Muitas vozes ao mesmo tempo falando em "desempenho" e "reprovação", "retenção", "evasão".]*
> *Entrevistada 3: Entra o IDH também! A evasão, retenção, desempenho.* ***(Grupo focal, gestores CRE)***

Como alguns gestores do nível central e intermediário relataram, eles têm como atividade dialogar com as escolas sobre os resultados das avaliações. No entanto, o desconhecimento fez com que estudassem individualmente os indicadores para cumprir sua tarefa, algo que não era efetuado anteriormente. Os gestores justificam suas dificuldades de compreensão em virtude do desconhecimento de estatística.

> *Entrevistada 1: A gente vai analisar os números, cresceu onde? Cresceu como? Ou diminuiu? Onde e como?... Como que esse resultado caiu? Onde que você tem que investir e que tipo de retorno você pode querer? Então você se debruça sobre esse resultado. Agora, por exemplo, recebemos o resultado da Prova Rio, então eu paro para tentar entender isso. Então, por exemplo, eu fiz Pedagogia e a estatística para mim é muito complicada, então para a escola também. Então eu me forço a estudar aqueles dados para isso. Eu não tinha esse hábito. Mas eu quero entender e esmiuçar isso para poder conversar com o coordenador pedagógico.* ***(Grupo focal, gestores CRE)***

O desconhecimento das avaliações externas nos níveis central e intermediário tem reflexo direto nas escolas. Os diretores questionavam as avaliações externas, não por seus cálculos, mas por não saberem o que estava sendo avaliado e seus objetivos. Nos dois grupos focais com diretores, observamos silêncio e desconforto quando perguntávamos sobre o cálculo dos indicadores e seus significados.

> *Mediadora: Vocês sabem como se chega a esse índice?*
> *Entrevistada 4: O cálculo?*

Mediadora: É, o cálculo...
Entrevistada 4: É, eu já tive essa curiosidade... Até já estudei isso, não profundamente, mas para ter uma noção... Não é assim uma coisa que eu domino... Mas uma noção eu tenho... [...]. Não é só a avaliação não, envolve...
Entrevistada 5: É evasão e o desempenho da escola como um todo. ***(Grupo focal, gestores alto desempenho, Rio de Janeiro)***

Mesmo com o desconhecimento, há uma visão compartilhada entre os gestores dos diferentes níveis de que as avaliações externas estão criando novas demandas no cotidiano escolar e na própria gestão da rede. Para membros da secretaria, as avaliações externas centralizaram as ações da gestão, pois todas as atividades estão relacionadas a elas.

Entrevistada 6: Quem utiliza os resultados é quem está ligado ao processo de ensino-aprendizagem [...]. Mas, por exemplo, temos a coordenadoria de RH. Ela provavelmente não mexe diretamente com esses dados, mas talvez indiretamente [...].
Entrevistada 7: Eu acho que hoje se trabalha muito em função dos resultados das avaliações. ***(Grupo focal, gestores SME)***

Uma gestora da SME exemplifica a nova situação como um "efeito dominó": a secretaria exige da CRE, que exige da escola. A escola, por sua vez, demanda à CRE algumas ações e procura a secretaria para ajudá-la nessa função.

Entrevistada 7: Vira um efeito dominó porque as CREs estão mais interessadas e cobrando das escolas porque as escolas também estão cobrando delas coisas bacanas para melhorar as suas escolas. As diretoras estão em cima da CRE para melhorar a sua escola, a CRE em cima da gente, vira um efeito dominó. ***(Grupo focal, gestores SME)***

Seguindo essa lógica do "efeito dominó", as gestoras das CREs indicaram que mudaram suas funções: se antes fiscalizavam, agora são "parceiras" das escolas, até porque também são cobradas pelos resultados das avaliações externas.

Entrevistada 3: A partir de 2009, mais em 2010, pelo menos na nossa CRE, esse olhar fiscalizador, essas visitas fiscalizadoras, perderam lugar para esse olhar, para essa parceria.

Entrevistada 8: Porque a cobrança é grande, né gente?
Entrevistada 3: Ou seja, não para acusar e sim para nortear algumas coisas que precisam ser mudadas. A postura da CRE mudou.
[...]
Entrevistada 5: Quanto aos resultados, hoje a escola sente mais confiança na gente, nas parcerias, a gente senta com ela pra olhar esses índices. Porque eles estão ali, não tem como esconder. O rei está nu. Salvo uma ou outra, elas se sentem mais à vontade para conversar. ***(Grupo focal, gestores CRE)***

Há uma dissonância entre o que pensam diretores e as CREs sobre a mudança nas atividades e o cotidiano escolar com a introdução das avaliações externas. Por um lado, os diretores indicam que a gestão escolar e o trabalho de sala de aula estão direcionados para as avaliações em larga escala. Para eles, a Prova Rio e a Prova Brasil ainda são eventos extraordinários. No entanto, todo seu trabalho está relacionado às demandas das avaliações, seu treinamento, aplicação e resultado.

Por outro lado, para os gestores das CREs as avaliações externas têm produzido um fenômeno novo nas escolas: os professores e diretores têm se conscientizado da necessidade de mudança em sua prática pedagógica. A partir dos resultados das avaliações externas, as escolas começam a pensar em sua forma de avaliar, em suas práticas e nas maneiras de modificá-las. As avaliações externas, portanto, trouxeram algo novo para a organização escolar porque passaram a englobar todos que nela trabalham.

Entrevistada 3: Tem uma questão, do tempo que a escola ficou no seu mundo, uma colega usava uma expressão, era o "jeito doméstico de ser". Então, quando vem uma avaliação externa, [...] uma necessidade de universalizar ou de homogeneizar o ensino do país, [...] mas trouxe, eu acho, essa mexida, na escola, fabulosa. Quer dizer, eu não faço o que eu quero.
[...]
Entrevistada 8: É muito legal a gente ouvir colegas dizendo "eu fiz isso na minha sala". E ele achava, algum tempo atrás, que aquele conteúdo não dava. Ele falava que meu aluno da escola

pública não vai aprender isso, não. E ele viu que o aluno consegue aprender e que tem uma avaliação que cobra isso dele.
Entrevistada 3: Não é o jeito doméstico de ser.
[...]
Entrevistada 9: Muitas vezes a escola faz o seu planejamento, do que vai ser ensinado, do que vai ser cobrado, o que vai ser avaliado e, aí, passou pelas provas externas e tem consciência do que precisa ser trabalhado. Não importa se o aluno é da escola pública ou da escola privada, aquilo precisa ser trabalhado. Então, isso foi uma mudança para os alunos, para os professores e, a meu ver, foi uma mudança positiva.
Entrevistada 2: Hoje acabou aquela ideia de um cofre que fica dentro de outro cofre, que fica dentro de outro cofre. Porque a sala de aula tinha muito aquela coisa do professor de "fechei a porta, eu faço o que eu quero. Eu sou o dono do saber e esta é a minha sala". E a coisa ficava "tão minha", "tão minha", que não podia sair nem daquele espaço físico. E o gestor acaba fazendo a mesma coisa, "é a minha escola, dos meus alunos". Então, hoje essa visão tá um pouquinho diferente [...].
Entrevistada 3: Porque agora esse professor já não acha mais que ele é um ser unicelular, ele já começa a perceber que ele faz parte de um conjunto de células e a coisa tem que funcionar de outras formas. ***(Grupo focal, gestores CRE)***

Os gestores indicam outra dissonância entre diretores e gestão intermediária, que diz respeito às diferenças entre as avaliações externas – Prova Brasil e Prova Rio – e internas, conhecida como Prova Bimestral. Para os diretores, as avaliações realizadas pela rede são mais fáceis porque englobam o que foi trabalhado nos cadernos pedagógicos. Ao passo que as avaliações externas são mais difíceis e exigem outros conteúdos. Segundo eles, essas diferenças criam um problema para o gestor porque este tem que pedir aos docentes que trabalhem com o material dos cadernos pedagógicos ou com o conteúdo das avaliações externas.

A gestão da CRE analisa positivamente esse movimento. As diferenças fazem com que o diretor e o professor comecem a refletir sobre seu trabalho e a propor mudanças em sua prática pedagógica.

Entrevistada 8: Isso leva também a escola a pensar no que ela está fazendo, né? O cara viu que, quando pega uma prova que já passou, e ele viu que ele pode fazer aquilo. Porque a gente ficou muito tempo achando que tinha que ser aquele conteúdo, aquela atividade, só aquela, aquele formato, aquele formato para aluno da escola pública, que o aluno só podia aprender até ali, então, agora a gente vê que tem colegas que percebem que ele pode ir além, que ele tem condições [...]. E aquele cara que pega o resultado da avaliação externa, que se debruça em cima, ela também leva a refletir dessa forma. ***(Grupo focal, gestores CRE)***

Outra mudança significativa relaciona-se à divulgação pública dos resultados. Segundo os gestores da secretaria, esse é um dos objetivos centrais da atual gestão: aproximar os pais da escola, utilizando como meio de mobilização a imprensa e a divulgação dos resultados educacionais para a sociedade.

Mediadora: E nessa mudança vocês perceberam algo nas atitudes de alunos, nas famílias dos responsáveis?

Entrevistada 7: Tem até preocupação de pais ligando para saber resultado de escola tal, para saber onde vai matricular o filho. Está tendo até uma procura dos pais pelas escolas para decidirem onde vão matricular os filhos [...].

Entrevistada 6: Mas isso é um resultado de uma estratégia dessa gestão, que é realmente aproximar os pais das escolas, justamente com isso da prova, que é uma coisa que está na mídia, índice de escola, bota na porta, não bota na porta [...]. Eu acho que, com isso, você chama atenção para que os pais discutam a educação também. Eu acho que a gente caminhou para essa discussão estar na mídia e os pais se apropriarem. ***(Grupo focal, gestores SME)***

Os gestores das CREs valorizam a divulgação pública dos resultados e o trabalho mais próximo aos pais e às famílias dos alunos. No entanto, diferentemente dos gestores da SME, apontam alguns problemas dessa relação que possibilitaram a "faxina pedagógica" realizada em algumas escolas após o resultado positivo nas avaliações externas.

Entrevistada 10: Isso é uma coisa que eu sou preocupada, sim, dos pais chegarem na porta daquela escola e rejeitarem aquela

escola, rejeitarem aquela comunidade, abaixa mais ainda a autoestima daquela escola que está numa posição desprivilegiada, vamos dizer assim, que a gente não sabe que motivos são esses, será que é só isso mesmo, se é só mesmo problema de gestão, ou se falta participação desses pais, o que já falou aqui. Então, preocupa você ter uma escola lotada e outra do outro lado vazia porque tem uma placa do Ideb[14] na porta dizendo que é 3.2. Entrevistada 2: O que me incomoda muito não é nem a vazia, mas é a de nota alta fazer uma faxina pedagógica, uma higienização pedagógica. Isso me incomoda muito. ***(Grupo focal, gestores CRE)***

14 Índice de Desenvolvimento da Educação Básica.

Para os gestores, a divulgação pública dos resultados faz com que eles se preocupem com as avaliações e seus indicadores, como informa o Entrevistado 11: "É. No dia a dia do gestor, por causa disso: essa preocupação que ele tem, de como vai ser a sua escola, como ela vai ficar... de como ela vai ser vista. O retrato para a sociedade." (Grupo focal, gestores baixo desempenho, Rio de Janeiro).

Quando questionamos sobre o resultado das avaliações e se os índices retratavam as escolas de maneira fidedigna, encontramos uma visão uníssona nas coordenadorias e na secretaria: as avaliações externas não refletem totalmente a "nossa realidade".

No caso dos gestores da SME, há um questionamento das avaliações feitas pelo Ministério da Educação (MEC). Consideram importantes sua existência e sua aplicação, mas creem que elas não refletem o trabalho feito pela rede municipal. Sustentam sua posição apontando a ausência do olhar pedagógico nas questões, as avaliações centradas na Língua Portuguesa e na Matemática e, principalmente, a impossibilidade de comparar a rede carioca, segundo eles a maior do país, com outras redes.

Entrevistada 7: A gente pensa que essa rede é a maior de todas, então, a comparação já fica diferente. A quantidade de nossos alunos que faz a prova é grande, é uma rede grande e a gente incentiva que eles façam a prova. Além disso, a gente tem dado um salto maior de 2009 para 2011, especialmente num segmento que a gente sempre ia mais atrás. Os alunos, principalmente em Matemática, sempre nos deixavam mais atrás. No Brasil todo,

comparados com as outras capitais, nós não estávamos mal, levando em consideração o tamanho da nossa rede. Agora fomos lá pra cima.

Entrevistada 6: Temos que pensar também que as outras capitais dividem com o estado o ensino fundamental. O Rio não divide, é só ele. Então a gente vai da creche até o nono ano. Isso também é uma coisa que dificulta.

Gestora 1: Não tem como comparar. ***(Grupo focal, gestores SME)***

As gestoras das CREs consideram as avaliações da rede mais fidedignas do que as avaliações nacionais. Dizem que elas refletem o trabalho realizado nas escolas por darem subsídios para diagnósticos e intervenções mais rápidas.

Entrevistada 12: É essencial, e no caso do Rio, eu acho que a gente acaba subestimando a avaliação nacional, porque acaba favorecendo algumas questões e outras não. As regionais, eu gosto mais, eu acho mais fidedignas.

Entrevistada 9: É, quando se faz essa prova especificamente para o Rio, a gente consegue analisar o trabalho todo, ver o que está funcionando e intervir com mais rapidez.

Entrevistada 10: O interessante é que tudo é no caráter de comparação de resultados, então, por exemplo, ano passado tinham escolas que ou tiveram um desempenho muito baixo na avaliação nacional ou um IDE-Rio[15] baixo, só que na Prova Rio era diferente, ou para mais ou para menos. Se a escola está ruim em uma avaliação que só tem prova de Língua Portuguesa e Matemática, como que você fala que o desempenho da escola é ruim? A Prova Brasil, então, só tem dois anos, então a gente fica muito restrita nessa avaliação... Não dá para ser pontual. Há uma série de questões dentro da escola que interferem muito e que falam da escola. É meio por aí. Pela própria elaboração da prova e aplicação. Como que a gente deve olhar para esses resultados? É interessante, é importante, mas ocorreu uma coisa, que foi uma avaliação externa, mas a avaliação diagnóstica mostra resultados não tão esperados e isso se torna muito mais produtivo, melhor desenvolvimento do que a avaliação externa. Eu estou certa disso. Quando a gente vai tabular os resultados, acaba tendo elementos muito mais interessantes do que a avaliação externa. ***(Grupo focal, gestores CRE)***

15 Índice de Desenvolvimento da Educação do Município do Rio de Janeiro.

Ainda segundo os gestores das CREs, as avaliações nacionais, principalmente por terem aplicadores externos, possibilitam dribles de alguns diretores e produzem indicadores que não refletem o trabalho e a realidade da escola.

> *Mediadora: Vocês acham que esses resultados refletem a realidade das escolas das CREs?*
>
> *Entrevistada 10: Em algumas escolas até sim, mas outras nem sempre... Até porque, volto a dizer, as avaliações externas são pontuais em duas disciplinas somente. E aí pode ser um trabalho focado nessas áreas e no restante a coisa não aparece, não acontece, aí saímos do individual para o coletivo, a escola como um todo não é lá essas coisas. Nós sabemos que tem diretores que dispensam os alunos com conceitos baixos para não fazer as avaliações. Assim como os alunos incluídos das turmas são convidados a não comparecer nesse dia [muitas vozes ao mesmo tempo].* ***(Grupo focal, gestores CRE)***

Os gestores partem da ideia de que os resultados das avaliações externas não refletem a realidade da escola. Na concepção deles, respondem por um espelho social e não somente pelo trabalho pedagógico.

> *Entrevistada 13: O aluno que é excluído por abandono conta ponto no Ideb. Mas o que ocorre? [...] O que vai aparecer ali é o espelho do pedagógico de sua escola; agora, o que aparece de repente no Ideb é um espelho, não só o pedagógico; porque o que se deve cobrar da escola é o espelho pedagógico, né? Agora a gente acaba respondendo por um espelho social.* ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

Percebemos que as avaliações em larga escala têm transitado progressivamente para o mundo ordinário da escola e dos órgãos de gestão da rede municipal. Não houve, em nenhum grupo, visões que negassem as avaliações por princípio, mas há críticas relacionadas a sua aplicação e principalmente a sua compreensão. Os grupos focais indicam que os profissionais da educação ainda enfrentam dificuldades técnicas para compreender as avaliações externas, até mesmo aqueles responsáveis por analisá-las.

O QUE OS GESTORES ACREDITAM QUE DEVEM/PODEM FAZER?

As diversas visões a respeito das avaliações em larga escala no material trabalhado trazem também uma série de interpretações a respeito das finalidades às que essas avaliações estariam respondendo e, em consequência, provocam diversas expectativas, nos diferentes sujeitos escolares, sobre o que se espera que cada um deles faça *com* ou *com base* nelas em sua ação cotidiana em função dessas finalidades.

Nesse sentido, a implementação dessas avaliações é interpretada pelos sujeitos como uma *demanda de ação*, balizada por seus entendimentos sobre as formas, os conteúdos e as finalidades supostamente apresentados pelas avaliações no cotidiano escolar. Ou seja, os sujeitos buscam antecipar futuros possíveis não apenas por meio da extrapolação de dados do passado ou das condições de produção e ação atuais em cada realidade escolar, mas também considerando a suposta finalidade técnica e/ou política implícita nas avaliações e a também suposta atuação esperada para cada um deles, de acordo com seu lugar institucional, evidenciando diversos sentidos prospectivos mais ou menos estratégicos. Cabe perguntar, portanto, quais são as percepções de gestores e funcionários da SME a respeito dos desafios e das possibilidades de ação a partir da introdução das avaliações em larga escala no cotidiano escolar.

Participando de uma visão mais estratégica, a proposta de aproveitamento dos dados das avaliações externas para a realização de um diagnóstico escolar aparece claramente apenas na fala de Entrevistado 14, gestor de escola de alto desempenho: "Eu percebo que, mesmo tendo uma origem de prestação de contas [...], nós podemos, sim, aproveitá-las para diagnóstico dos defeitos ou das falhas e fazer correções futuras".

Essa visão é compartilhada pelos funcionários da SME, destacando-se a menção às CREs como instância dinâmica na utilização desses indicadores para melhorar a qualidade educacional:

> *Entrevistada 1: As CREs mexem diretamente. A CRE mexe um bocado. Eu acho que setores lá na secretaria, aquele pessoal do*

material, eles vão usar os dados, como? Vão comprar equipamento para x escolas, isso pode ser influenciado pelos resultados? Talvez possa. Vou equipar melhor, vou verificar. Acho que todos os setores podem usar os dados. A secretaria inteira tá voltada com esses dados. Lógico que quem está mais diretamente ligado ao ensino-aprendizagem vai estar mais junto desses dados. Se pensa muito nas escolas que precisam de maior apoio. ***(Grupo focal, gestores SME)***

Por isso, o interesse em facilitar a apropriação desses dados, para que possam ser estrategicamente aproveitados, é destacado na fala da Entrevistada 7: "Agora vai ter o boletim escolar, com todos os dados da escola, com análise com comentários, para a escola ter acesso ao boletim dela, com os pontos, os dados mais mastigadinhos para aquela escola." (Grupo focal, gestores SME).

Divergindo dessa visão, duas gestoras de escolas com baixo desempenho dizem que é necessário amadurecer o que fazer com os dados. Esse é o caso da Entrevistada 15, gestora de uma escola de baixo desempenho: "Eu acho que tem um processo de amadurecimento que passa pela nossa formação de aprender a lidar com esses instrumentos sem se descaracterizar e de incorporá-los, mas não substituí-los".

Compartilhando as dificuldades para lidar com esses dados, o Entrevistado 11, de escola de baixo desempenho, entende que eles são importantes para refletir sobre e motivar a equipe, mas isso em função das metas e do 14º salário, e atendendo aos fatores retenção e fluxo como variáveis para melhorar a nota da escola:

Entrevistado 11: Eu reuni o grupo e comentamos: "Olha, vamos nos esforçar para conseguir ou não?" Aí todo mundo concordou. Porque não passa só pela prova do Ideb só para o 9º ano. Então também influencia na aprovação, também, da escola [...]. Eu acho que também posso estar errado, mas eu acho que provocou nos professores que abraçaram a causa, em tentar se esforçar um pouco mais para ver se o aluno tem condições ou não de passar, porque antes, até o mesmo professor já passou uma prova, um teste, alguma coisa e não conseguiu, então deixa para lá, vai ficar com aquela nota mesmo. Porque se a escola reprovas-

se bastante, como reprovou em 2009 e 2010, a gente não iria conseguir, não iria ganhar. E como ganhamos o IDE-Rio, que é o 7º ano que faz a prova e eles sentiram o gostinho, que foi de 2010 para 2011, aí, no ano passado, em 2011, eles queriam de novo. ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

A motivação também aparece como uma alternativa de ação escolhida pelos funcionários das CREs, que afirmam ter que "esperar o tempo de o professor [diretor]" superar o choque de receber resultados muito baixos. No entanto, chama a atenção o fato de que eles tampouco saibam explicar como se calculam os indicadores e, portanto, uma análise mais técnica e detalhada dos números pode ficar comprometida, apesar de ser com base e em função desses indicadores que se espera que a escola realize seu planejamento.

Entrevistada 16: Quando eu vou para dentro das escolas, daquelas escolas que ficam lá embaixo no ranking, quando eu sento com elas, eu pergunto pelo plano de ação, o que está ocorrendo, aí a gente começa a entender algumas coisas e a fazer com que ele pense aquilo ali de outra forma. Houve uma escola que o diretor sentou na minha frente e não parava de chorar quando conversamos. Aí eu falei "a gente tem que agora largar o luto e ir à luta". E o nosso momento já passou por essas duas etapas, agora estamos mudando. ***(Grupo focal, gestores CRE)***

Contudo, o desconhecimento sobre as avaliações e seus resultados, assim como mobiliza em alguns a vontade de querer entender melhor ou mesmo de "dar um jeito" de melhorar os indicadores – por meio da motivação da equipe ou de outras variáveis de peso sobre as que se teria maior controle, como o fluxo escolar –, também gera movimentos de resistência ou até de certa denúncia. Dessa forma, a articulação da equipe escolar não se apresenta como tarefa fácil e, mesmo em escolas com alto desempenho, as resistências de alguns professores para trabalhar com os dados advindos das avaliações são destacadas pela gestão:

Entrevistada 4: Então, quando você organiza uma reunião da coordenação e tal, "gente, vamos pegar o retorno, vamos analisar"... É uma resistência que vocês não têm ideia, né, e assim,

lá na escola, a gente não trabalha para prova, não se trabalha para prova, a gente trabalha para o aluno, é a questão mesmo da formação. Se ele vai fazer a prova A, B, C ou se tem prova externa, é uma consequência, então a gente trabalha assim. ***(Grupo focal, gestores alto desempenho, Rio de Janeiro)***

A resistência de alguns professores se manifesta de diversas formas, chegando inclusive a ações que beiram a temeridade política (MATUS, 1996, p. 174), rejeitando a realização das provas na escola em que trabalham:

Entrevistada 9: Na CRE [citando uma das 11 CREs], acontece uma coisa, particularmente, que não foi citada até agora, que é a resistência dos professores da maneira "não vou colaborar com isso, não concordo com isso, os meus alunos não vão fazer essa avaliação". Aí, no dia da avaliação, não tem aluno, não tem turma, porque eles são contra e não participam. E quando alguns ganham 14º, eles batem no peito e falam "não ganho, não concordo com isso". Então é uma questão muitas vezes, simplesmente, que se leva pelo senso comum. Nós temos uma situação extrema de uma escola em que todos os professores se recusam a cumprir as provas bimestrais. ***(Grupo focal, gestores CRE)***

No extremo oposto, alguns gestores ficam apreensivos com relação à ação inercial, como resposta passiva, que pode ser escolhida por alguns professores e que parece estar acontecendo em algumas escolas:

Entrevistado 17: A gente tem que dizer, na minha escola também acontece, para o professor que ele não pode dar uma avaliação, botar a nota final só com a avaliação externa. Então isso, a culpa não é da avaliação externa, é culpa do profissional que eu não sei que palavra vou ter. ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

Assumir essa inércia a ponto de não realizar mais a própria prova, ação escolar com amplíssima trajetória e legitimidade dentro do sistema educacional brasileiro, e depender exclusivamente da avaliação de outra instância institucional, pode presumir tanto a delegação como a suposta expropriação do poder de avaliação, no entendimento de

que há uma perda substantiva de autonomia escolar, a partir do momento em que se deve "dividir" essa função com outras instâncias do sistema.

Entrevistada 18: A gente deixou de ser diretor e passou a ser gestor de algo que não nasce dentro da escola. E isso é muito sério.

Entrevistada 19: E isso que eu estou falando, está afetando o nosso PPP.[16] Eu senti isso. Eu acho que está afetando o nosso dia a dia.

Entrevistada 18: Por outro lado, também, tinha escola que nem tinha PPP.

Entrevistada 19: Mas, por exemplo, a gente tinha planejamento.

*Entrevistada 18: Quer ver uma coisa, quando a gente falava de avaliação, eu lembro que, quando surgiu, lá atrás, a resolução, a portaria que determinava como seria a avaliação, muitos professores quicaram: "Não! Está tirando da gente a autonomia, o direito de fazer a nossa prova". **(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

16 Projeto Político-Pedagógico.

A suposta autonomia absoluta anterior que, como fica evidente na fala da Entrevistada 18, possibilitava que algumas escolas não se preocupassem em realizar seu planejamento institucional, aparece novamente na fala da mesma gestora, preocupada com o fato de que o professor não abra mão do planejamento, o que implicaria também abrir mão da autonomia. Mas isso requer, ao mesmo tempo, competência profissional para poder assumir seu papel no sistema educacional.

*Entrevistada 15: Outro dia, eu estava conversando com uma diretora, ela estava falando para mim que ela discutiu com a professora dela porque a professora falou: "Eu não sei para que essa besteira de a gente ter que planejar tanto. Já vem tudo pronto. É só a gente digitar no sistema. As escolas 3.0. Já está tudo pronto". Quer dizer, ainda tem um equívoco aí do que é planejamento, do que é plano curricular. Então, eu acho que ainda tem algumas questões aí, semânticas e de concepções de trabalho que também ajudam a gente ser atropelados. **(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

E, mais à frente, a Entrevistada 15 completa:

Eu conheço pouquíssimos professores que sabem fazer isso: mediar apostila, livro didático, educopédia, entendeu? Mas...

Eu fiquei doente quando uma diretora falou para mim que a professora dela disse que não sabe o que fazer com tantas horas de centro de estudos. Olha! Olha o impacto que isso pode ter na boca de um político, daqui a pouco. ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

Por outro lado, as "suspeitas" a respeito de uma maior centralização por parte das instâncias superiores de governo institucional, assim como a implementação de projetos elaborados sem a participação dos sujeitos diretamente implicados nas escolas, mas apenas "aplicados" nesses espaços, aparecem confirmadas pela Entrevistada 1:

A relação com a SME mudou um pouco, porque é tudo muito descentralizado para as CREs. Agora é menos [...]. Hoje centralizamos mais. Hoje tentamos fazer o caminho contrário, tentar trabalhar com projetos maiores, sem contar que as CREs continuam fazendo o que acham necessário dentro dos seus espaços, com as suas caraterísticas, muito diferentes, com demandas diferentes. Mas, hoje, a gente acha que mesmo com as demandas, cada uma de um jeito, a gente deve ter um referencial pedagógico forte, mas, hoje, esse trabalho é maior do que era. Tem muita interferência da CRE no desempenho, porque ela gerencia os projetos para as escolas e tal. Mas a CRE tem que ter essa visão. Mas, ao mesmo tempo, lá vem a escola achando que a CRE a está perseguindo. ***(Grupo focal, gestor SME)***

A centralização de um planejamento atrelado à avaliação, mas que dispensa a participação dos sujeitos diretamente envolvidos no cotidiano escolar, no entanto, gera também alguns efeitos que desorientam e, inclusive, anulam a ação pedagógica competente de algumas escolas. Esse é o caso apresentado pelo Entrevistado 14:

No Rio, a gente tem uma realidade num dia, no outro já tem outra... Vem uma coisa em cima de você, é uma resolução, um decreto. É como essa verba do PDE,[17] *a gente previu, ficou um ano e pouco trabalhando, planejando: "Vai ser aplicado aqui, ali", aí você vê... Tudo mudou. Já vamos atender de outra forma em outro segmento no ano que vem, então, planejamos uma coisa que visava atender até o segmento que já não estará mais no ano que vem.* ***(Grupo focal, gestores alto desempenho, Rio de Janeiro)***

17 Plano de Desenvolvimento da Escola.

Paralelamente, outras situações similares, vinculadas à utilização padronizada de critérios da administração central, também surpreendem as gestões escolares.

> *Entrevistada 4: Não, a CRE fala assim: "você entrou no PDE, tua escola tá mal", porque o PDE é um plus, é um recurso a mais, né, para você dar conta do pedagógico, do que tá acontecendo na escola, eu vi de outra forma, eu falei para os professores: "Graças a Deus, então, que a gente agora vai ter dinheiro para comprar material para desenvolver os projetos, né?"*
>
> *Entrevistada 5: Às vezes eu queria ser a pior escola para ter dinheiro...*
>
> *Entrevistada 4: Pois é, não é, professora? Então, a gente recebeu por conta dessa questão da reprovação. Nós fomos "punidos" com PDE, com Mais Educação, é mais verba, é mais... Tá ótimo. A escola agora tá com um recurso.* ***(Grupo focal, gestores alto desempenho, Rio de Janeiro)***

A percepção e o conhecimento desses critérios decisionais, em especial em função da captação de recursos para as escolas e/ou seus funcionários, propiciam o desenvolvimento de outro tipo de ação, diferente da inércia passiva, mas, dissimuladamente estratégica, que podemos denominar de estratagema, ou seja, o engano calculado (MATUS, 1987, p. 103). Estratégia que o ator impõe a si mesmo como regra de jogo e que responde a um estilo político que "define as fronteiras do permitido e do proibido, e estrutura suas próprias normas de ética. Define com um perfil muito particular o modo de fazer e aceitar a política" (MATUS, 1996, p. 17). Vejamos:

> *Entrevistado 11: Porque a nota da [escola] L. era 3.0. Em 2007, foi 4.3 [...]. Aí, agora, a meta era atingir 3.4. Nós atingimos 4.2, 40%. Aí, os professores disseram: "Nós crescemos demais". Quer dizer, em 2013, de 4.2 vai ter que chegar a 4.6, aí, talvez, não vai conseguir e não vai receber o 14º.* ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

E, mais à frente, completa: "Vocês tinham que ter ensinado só um pouquinho para crescer só um pouquinho, para não crescer muito". Esse estratagema de controle do cresci-

mento do Ideb em função de ter chances de receber o 14º salário em mais de uma oportunidade, apesar de a escola ter condições de dar esse "salto de qualidade", associa-se a outros, que apresentamos nos parágrafos a seguir. Nesses casos, destinados a elevar os indicadores sem melhorar a eficácia escolar, solicitando, por exemplo, aos estudantes com maiores dificuldades de aprendizagem que não compareçam à escola no dia da prova, como explica o Entrevistado 17:

> *"Então, você tem histórias escabrosas. Gente que você tinha um conceito... 'Ó, hoje é a avaliação externa, quem tem conceito I fica em casa', aquele 'izinho' que... Então tem isso, então tem várias histórias em cima disso..."* ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

Da mesma forma, outras escolas parecem estar escolhendo o "atalho"[18] de focalizar um trabalho pedagógico mais eficiente apenas nas turmas que realizarão as provas, como explica a Entrevistada 15:

> *A gente tem escolas que preparam e que já pensam a sua organização escolar a partir de quem vai pegar as turmas que vão ser avaliadas porque são turmas que vão levar a nota e dizer como é que eu estou. É claro que esta avaliação não diz como é que a gente está. Também não digo que ela não tem que acontecer, não.* ***(Grupo focal, gestores baixo desempenho, Rio de Janeiro)***

18 Não podemos deixar de indicar que as práticas descritas – seleção de alunos, cálculo para controle dos índices, escolha de alunos para realização de provas – que estamos chamando de "atalhos" também são classificáveis como "jeitinhos", no sentido proposto por Roberto Da Matta (1983).

Por fim, apesar de ilegal, a seleção de alunos para ingresso e permanência na escola parece estar sendo um dos estratagemas das escolas, em especial daquelas com alto desempenho, como denunciam as falas a seguir.

> *Entrevistada 5: Aí, quando você vai transferir um aluno: "É um bom aluno? É um aluno I? É um aluno B?". Porque tá assim, né, nesse nível. Então é complicado.* ***(Grupo focal, gestores alto desempenho, Rio de Janeiro)***
>
> *Entrevistada 12: Nós temos bastante escola bem colocada, mas nós temos várias críticas sobre as escolas, são algumas escolas que selecionam alunos e tal. Então, são escolas que nós temos dez olhos para o outro lado. Então, às vezes a gente olha as bem colocadas e a gente sabe que são escolas assim.* ***(Grupo focal, gestores CRE)***

A diversidade de ações evidenciadas ao longo desses relatos nos mostra, então, os diferentes entendimentos dos sujeitos sobre as finalidades das avaliações externas e a produção de indicadores educacionais. Também evidencia as interpretações sobre sua função como agentes do sistema educacional, ou seja, sobre o desempenho de suas funções tanto como docentes quanto como funcionários públicos, questionando as diversas concepções sobre a profissão docente.

CONSIDERAÇÕES FINAIS

A análise dos dados descritos neste artigo permite indicar três questões associadas à recepção das avaliações no cotidiano escolar. Inicialmente, é possível apontar desconhecimento por parte de gestores com relação aos aspectos técnicos das avaliações. Encontramos, em todas as instâncias da gestão, falas truncadas sobre as avaliações externas e os índices por elas produzidos.

Outro aspecto é o conjunto de dissonâncias presentes nas falas de gestores. Foi possível perceber que as cobranças relacionadas ao desempenho das escolas seguem o que uma gestora classificou como "efeito dominó": um conjunto de cobranças sucessivas que afetam os agentes de acordo com a posição que ocupam na hierarquia da gestão e/ou das escolas. Conforme indicado, gestores centrais sofrem pressões, pressionam gestores intermediários, que cobram dos gestores de unidades escolares, que, ao final, exigem dos professores.

Finalmente, como resultado dessa conversa desafinada, os mais diversos "jeitinhos" e estratagemas são utilizados como forma de enfrentar os problemas considerados pontuais. Percebemos que os gestores entendem que determinadas ações – como impedir reprovações e/ou que alunos de menor desempenho façam as provas – resultam em aumentos artificiais nos índices. Essas estratégias acabam por criar uma visão equivocada sobre o desempenho das escolas e dos sistemas de ensino. É claro que também mapeamos falas que indicam reconhecimento e valorização dos dados, inclusive dos desafios trazidos por eles.

Ao final, não é possível afirmar, evidentemente, que a ampliação do conhecimento técnico eliminaria os "jeitinhos" e os estratagemas. Porém, a título de hipótese, podemos indicar que uma maior capacitação para o uso de dados reduziria as sombras e os espaços nebulosos associados aos sistemas nacionais e locais de avaliação externa.

REFERÊNCIAS

ALVES, Fátima. Qualidade da educação fundamental: integrando desempenho e fluxo escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 15, n. 57, p. 525-542, 2007.

BARBOSA, Lívia. *O jeitinho brasileiro*: a arte de ser mais igual que os outros. Rio de Janeiro: Campus, 1982.

BONAMINO, Alicia. *Tempos de avaliação educacional*: o Saeb, seus agentes, referências e tendências. Rio de Janeiro: Quartet, 2002.

BONAMINO, Alicia; COSCARELLI, Carla; FRANCO, Creso. Avaliação e letramento: concepções de aluno letrado subjacentes ao Saeb e ao Pisa. *Educação & Sociedade*, Campinas, v. 23, n. 81, p. 91-113, 2002.

BONAMINO, Alicia; SOUSA, Sandra Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, 2012.

BROOKE, Nigel; CUNHA, Maria Amália de A. A avaliação externa como instrumento da gestão educacional nos estados. *Estudos e Pesquisas Educacionais*, São Paulo, v. 2, p. 17-79, 2011. Disponível em: <http://www.fvc.org.br/estudos-e-pesquisas/avulsas/%20avaliacao-externa-instrumento-gestao-educacional-estados.shtml>. Acesso em: 08 jul. 2014.

CARVALHO, Gisele F. S.; **MACEDO**, Maria do Socorro A. N. Avaliação oficial: o que dizem os professores sobre o impacto na prática docente. *Educação e Pesquisa*, São Paulo, v. 37, n. 3, p. 549-564, 2011.

DA MATTA, Roberto. *Carnavais, malandros e heróis*. Rio de Janeiro: Guanabara, 1983.

FERNANDES, Maria José da S. As recentes reformas educacionais paulistas na visão dos professores. *Educação em Revista*, Belo Horizonte, v. 26, n. 3, p. 75-101, dez. 2010.

FONTANIVE, Nilma. O uso pedagógico dos testes. In: SOUZA, Alberto de M. (Org.). *Dimensões da avaliação educacional*. Petrópolis: Vozes, 2005.

FRANCO, Creso; BONAMINO, Alicia. Iniciativas recentes de avaliação da qualidade da educação no Brasil. In: FRANCO, Creso (Org.). *Avaliação, ciclos e promoção na educação.* Porto Alegre: Artmed, 2001. p. 15-28.

FREITAS, Luiz Carlos de. Os reformadores empresariais da educação: da desmoralização do magistério à destruição do sistema público de educação. *Educação & Sociedade*, São Paulo, v. 33, n. 119, p. 379-404, 2012. Disponível em: <http://www.scielo.br/pdf/es/v33n119/a04v33n119.pdf>. Acesso em: 21 jul. 2014.

LÉVI-STRAUSS, Claude. *Tristes trópicos.* Lisboa: Edições 70, 1981.

MANDINACH, Ellen B.; **GUMMER**, Edith S. A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, v. 42, p. 30-37, Jan./Feb. 2013.

MANDINACH, Ellen B.; HONEY, Margaret; LIGHT, Daniel. *A Theoretical Framework for Data-Driven-Decision Making*. Paper presented at American Educational Research Association, San Francisco, California, 2006.

MARSH, Julie; PANE, John F.; HAMILTON, Laura S. Making sense of data-driven decision making in education: evidence from recent RAND research. *RAND Occasional Papel* (OP-170-EDU). Santa Monica, CA: RAND, 2006.

MARTINS, Angela M.; SOUSA, Sandra Z. A produção científica sobre avaliação educacional e gestão de sistemas e de escolas: o campo da questão entre 2000 e 2008. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 20, n. 74, p. 9-26, 2012.

MATUS, Carlos. *Política, planificación y gobierno.* Caracas: Fundación Altadir, 1987.

_____. *Estratégias políticas*: Chimpanzé, Machiavello e Gandhi. São Paulo: Fundap, 1996.

MESQUITA, Silvana. Os resultados do Ideb no cotidiano escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 20, n. 76, p. 587-606, set. 2012.

PIMENTA, Cláudia O. Bons resultados no Ideb: fatores explicativos do município de Indaiatuba. In: SIMPÓSIO BRASILEIRO, 25.; CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 2., 2011. São Paulo. *Cadernos Anpae*. São Paulo: Anpae, 2011. v. 1. p. 1-13.

ROSISTOLATO, Rodrigo P. R.; PRADO, Ana P. do. Os profissionais da educação e as avaliações externas de aprendizagem: uma comparação entre Rio de Janeiro e Duque de Caxias. *Roteiro*, Rio de Janeiro, v. 39, p. 311-329, 2014.

SILVA, Vandré G. et al. Uso da avaliação externa por equipes gestoras e profissionais docentes: um estudo em quatro redes de ensino público. São Paulo: FCC/SEP, 2013. (Textos FCC, v. 38).

SOARES, José Francisco. O índice de desenvolvimento da educação de São Paulo – Idesp: bases metodológicas. *São Paulo em Perspectiva*, São Paulo, v. 23, n. 1, p. 29-41, 2009.

SOUSA, Sandra Z. Possíveis impactos das políticas de avaliação no currículo escolar. *Cadernos de Pesquisa*, São Paulo, n. 119, p. 175-190, 2003.

RODRIGO ROSISTOLATO
Professor do Programa de Pós-Graduação em Educação da
Universidade Federal do Rio de Janeiro (PPGE/UFRJ),
Rio de Janeiro, Rio de Janeiro, Brasil
rosistolato@hotmail.com

ANA PIRES DO PRADO
Professora da Faculdade de Educação da
Universidade Federal do Rio de Janeiro (FE/UFRJ),
Rio de Janeiro, Rio de Janeiro, Brasil
anapprado@yahoo.com

SILVINA JULIA FERNÁNDEZ
Professora da Faculdade de Educação da
Universidade Federal do Rio de Janeiro (FE/UFRJ),
Rio de Janeiro, Rio de Janeiro, Brasil
silvinajf@gmail.com

Recebido em: JULHO 2014
Aprovado para publicação em: NOVEMBRO 2014

ACCOUNTABILITY ESCOLAR: UM ESTUDO EXPLORATÓRIO DO PERFIL DAS ESCOLAS PREMIADAS

MARIANE CAMPELO KOSLINSKI
CAROLINA PORTELA DA CUNHA
FELIPE MACEDO DE ANDRADE

RESUMO

O artigo investiga o perfil das escolas que foram premiadas por um dos programas que compõem a política de accountability *escolar, recentemente adotada pela Secretaria Municipal de Educação (SME) do Rio de Janeiro. Discute-se uma bibliografia que sinaliza como desenhos de políticas de responsabilização podem emitir diferentes sinais para atores escolares e descreve-se, brevemente, a política de responsabilização escolar da SME do Rio de Janeiro. As análises utilizaram bases dos dados da SME do Rio de Janeiro e dos questionários contextuais da Prova Brasil, sendo identificadas algumas limitações do programa em premiar o esforço dos atores escolares. Nas considerações finais, são pontuadas algumas alterações na regulamentação que buscam corrigir alguns dos sinais/incentivos emitidos pelo programa foco do estudo.*

PALAVRAS-CHAVE ACCOUNTABILITY • AVALIAÇÃO DA EDUCAÇÃO • IDEB • IDE-RIO.

RESUMEN

El artículo investiga el perfil de las escuelas que recibieron premios por medio de uno de los programas que componen la política de accountability *escolar, recientemente adoptada por la Secretaría Municipal de Educación (SME) de Rio de Janeiro. Se discute una bibliografía que señala que proyectos de políticas de responsabilización pueden emitir diferentes señales para actores escolares y se describe brevemente la política de responsabilización escolar de la SME de Rio de Janeiro. Los análisis utilizaron bases de datos de la SME de Rio de Janeiro y de los cuestionarios contextuales de la* Prova Brasil, *identificándose algunas limitaciones del programa al premiar el esfuerzo de los actores escolares. En las consideraciones finales se subrayan algunas alteraciones en la reglamentación, que tratan de corregir algunas de las señales e incentivos emitidos por el programa foco del estudio.*

PALABRAS CLAVE ACCOUNTABILITY • EVALUACIÓN DE LA EDUCACIÓN • IDEB • IDE-RIO.

ABSTRACT

This article investigates the profile of schools that were rewarded for one of the programs that comprise the school accountability policy, recently adopted by the Municipal Department of Education (SME) of Rio de Janeiro. A bibliography indicating how the design of accountability policies may send different signals to school actors is discussed; and, the SME policy for school accountability in Rio de Janeiro is described briefly. The analyses used SME databases from Rio de Janeiro and the contextual questionnaires from the Prova Brasil*; some limitations of the program that rewards the efforts of school actors are identified. In the final considerations, some alterations in the rules that seek to correct some of the signals/incentives sent by the program focused on in this study are pointed out.*

KEYWORDS ACCOUNTABILITY • EDUCATIONAL EVALUATION • IDEB • IDE-RIO.

INTRODUÇÃO

Nas últimas décadas, sistemas educacionais em diversos contextos nacionais e subnacionais adotaram políticas de responsabilização (*accountability*) como forma de prover concepções mais refinadas de objetivos ou metas e para promover melhoria na qualidade e eficácia da educação (HOFFER, 2000; BROOKE, 2013; HOUT; ELLIOT, 2011). Algumas experiências no contexto internacional são relatadas em estudos realizados em países como Inglaterra e Estados Unidos (BROOKE, 2006), Chile (MCEWAN et al., 2008), México (VEGAS, 2005), Israel (LAVY, 2002, 2007) e Andhra Pradesh, na Índia (MURALIDHARAN; SUNDARAMAN, 2009). Grosso modo, a partir da divulgação de informações sobre os resultados e práticas da escola e da adoção de um sistema de consequências, governos pretendem promover mudanças na gestão escolar e nas práticas em sala de aula e, consequentemente, no desempenho das escolas.[1]

A experiência brasileira com *accountability* escolar ocorreu, sobretudo, por meio da recente implementação e difusão de sistemas de avaliação da educação em larga escala

1 Esses objetivos são, em sua maioria, respaldados pela ampla literatura de eficácia escolar, que estabelece que fatores internos às escolas, entre outros fatores, estão associados com sua eficácia no aprendizado dos alunos (REYNOLDS; TEDDLIE, 2008).

(BONAMINO; SOUZA, 2012). No entanto, da mesma forma que nos contextos anteriormente mencionados, não existe consenso nos estudos brasileiros sobre os possíveis impactos da implantação das políticas de responsabilização escolar. Mesmo após mais de uma década em que começaram a ser adotadas em alguns estados e municípios brasileiros, tais políticas ainda são foco de um acalorado debate acadêmico. Alguns autores defendem que o pagamento de bônus aos profissionais da educação pode gerar concorrência entre as escolas e provocar mais desigualdades escolares (SETUBAL, 2010). Outros apontam para a relação pouco saudável provocada pelo *accountability* vertical, que estabelece uma hierarquia entre quem cobra e quem é cobrado pelo sistema de *accountability*, e para os efeitos negativos dessas políticas no contexto norte-americano, que seriam reproduzidos no Brasil (FREITAS, 2013). Outros estudos argumentam que é legítimo realizar comparações entre as escolas, na tentativa de identificar as que apresentam dificuldade em desempenhar seu trabalho, bem como aquelas que devem servir como fonte de aprendizagem para as demais (BROOKE, 2013; PONTUAL, 2008). Também indicam que é necessário produzir evidências do resultado dessas políticas no Brasil, uma vez que os resultados de estudos internacionais mostram efeitos tanto positivos quanto negativos das políticas de responsabilização, não havendo como garantir, previamente, quais serão os impactos no contexto brasileiro (BROOKE, 2013).

No entanto, no Brasil, é recente a proliferação da responsabilização escolar, sendo poucos os estudos sobre seu impacto tanto no desempenho das escolas como em práticas escolares e/ou *gaming*.[2] O presente artigo pretende contribuir para essa discussão a partir de um estudo da política de responsabilização recentemente adotada pela Prefeitura do município do Rio de Janeiro. Mais especificamente, a análise focaliza o "Prêmio Anual de Desempenho",[3] programa que premia as escolas que alcançam metas de crescimento do Ideb (Índice de Desenvolvimento da Educação Básica) e IDE-Rio (Índice de Desenvolvimento da Educação do Município do Rio de Janeiro) com um bônus salarial para professores, direção e funcionários das escolas de ensino fundamental.

2 Estudos sobre políticas de *accountability* na área de educação observaram a proliferação de estratégias adotadas pelas escolas descritas como *gaming the system*, ou formas de trapaça ou de jogar com o sistema para atingir as metas estabelecidas.

3 O Prêmio Anual de Desempenho foi instituído pelo Decreto n. 30.860 de 1º de julho de 2009.

No entanto, o estudo não propõe identificar o impacto do referido programa, mas somente fazer um exercício exploratório para observar o perfil das escolas – com especial foco para gestão, nível socioeconômico do alunado, recursos, complexidade e desempenho prévio das escolas – que foram bem-sucedidas em atingir as metas estipuladas pela Secretaria Municipal de Educação (SME) do Rio de Janeiro em 2009 e 2011. Para tanto, parte-se do pressuposto de que programas de responsabilização escolar, ao avaliar escolas e estabelecer consequências, enviam sinais do comportamento esperado pelas Secretarias e/ou gestores dos Sistemas de Educação. Dessa forma, tal investigação exploratória pode trazer pistas sobre quais escolas são percebidas pelas demais como "bem-sucedidas" e, consequentemente, que tipos de práticas e/ou estratégias escolares têm sido/são, possivelmente, incentivados.

Assim, o artigo aborda, em sua primeira parte, estudos realizados nos contextos internacional e brasileiro que sinalizam a ênfase e o impacto da gestão escolar frente a políticas de responsabilização. Discute-se, ainda, como diferentes características dos desenhos de políticas de responsabilização podem premiar e/ou punir escolas com características que não necessariamente estão relacionadas às práticas/esforços dos atores escolares. Em seguida, é realizada uma breve discussão sobre a política de responsabilização escolar da Prefeitura do Rio de Janeiro, em especial a forma como estabelece metas para as escolas e a ênfase posta no protagonismo do diretor das escolas. Para analisar o perfil das escolas que receberam o Prêmio Anual de Desempenho em 2009 e 2011, são utilizadas as bases de dados da SME do Rio de Janeiro e dos questionários contextuais da Prova Brasil. As análises e as considerações finais identificam, de um lado, limitações do programa em premiar o esforço dos atores escolares e, de outro, algumas alterações na regulamentação que buscaram corrigir alguns dos sinais/incentivos emitidos pela SME por meio do Prêmio Anual de Desempenho.

ACCOUNTABILITY ESCOLAR, PRÁTICAS DE GESTÃO E MEDIDAS DE DESEMPENHO

Em geral, programas de *accountability* na educação fazem parte de reformas educacionais mais amplas com base no estabelecimento de padrões medidos a partir do desempenho dos alunos em avaliações externas. Para Ladd (2001), programas de responsabilização do tipo *top-down* podem focalizar diversas unidades de análise, desde instâncias intermediárias de gestão da educação, até escolas, professores e alunos, ou ainda uma combinação de tais unidades. Mesmo diante da possibilidade de múltiplos desenhos, as reformas educacionais de incentivos com base em resultados têm, em geral, a intenção de promover padrões de *performance* mais claros e mensuráveis, alinhar o currículo a tais padrões e encorajar a escola a melhorar os resultados. A ideia central é monitorar e incentivar certos comportamentos dos distritos, escolas, professores e/ou alunos para alcançar os padrões estabelecidos (FIGLIO; LOEB, 2011).

No entanto, para além do debate sobre a efetividade ou o impacto de políticas de responsabilização, diversos autores discutem qual seria a unidade mais efetiva para promover mudanças em políticas escolares e práticas dos professores e, consequentemente, sobre o desempenho dos alunos e da escola. Por exemplo, estudos argumentam que tomar as instâncias intermediárias (tais como os distritos escolares nos Estados Unidos) como foco dos programas de *accountability* não seria adequado, já que desenhos desse tipo poderiam maquiar a grande desigualdade de resultados entre escolas dentro de tais unidades (FIGLIO; LOEB, 2011; LADD, 2001).

Os programas de *accountability* com desenhos que focalizam a escola como um todo, ou "o processo de avaliação do desempenho escolar com base em medidas de desempenho dos alunos" (FIGLIO; LOEB, 2011, p. 383, tradução nossa), são os mais difundidos,[4] sendo defendidos por diversos estudos como os mais adequados. Mesmo que a *accountability* escolar tenha, em última instância, a intenção principal de modificar e incentivar certos comportamentos ou práticas dos professores, os argumentos a favor de políticas de responsabilização baseadas na escola apontam para a dificuldade

4 Por exemplo, a política *Race to the Top*, adotada em 2010 pelo governo federal dos Estados Unidos, incentivou a adoção não só de medidas de progresso de desempenho dos alunos, mas também de medidas de responsabilização voltadas para a escola como todo, já que os formuladores da política partiam do pressuposto de que seria mais fácil estimar a efetividade/eficácia dos diretores do que a dos professores. Isso porque as escolas contariam com amostras mais amplas e maior disponibilidade de dados (LIU et al., 2013).

de isolar a contribuição para a aprendizagem de um professor individualmente e/ou de mensurar a eficácia do professor. Além disso, desenhos de *accountability* que tomam como unidade de análise os professores poderiam incentivar a competição entre eles, em detrimento de cooperação para atingir uma meta em comum. Ladd (2001) ainda argumenta que, no contexto norte-americano, as escolas possuem mais autonomia e poder de decisão do que no passado e, portanto, teriam, mais do que os professores individualmente, possibilidades ou meios para promover mudanças na distribuição de recursos e práticas para melhorar o desempenho dos alunos, bem como maior potencial para promover mudanças na cultura da escola.[5]

A defesa de políticas de responsabilização com foco na escola encontra respaldo em diversas pesquisas. Ladd e Zelli (2002), por exemplo, observaram que o programa de responsabilização escolar do estado da Carolina do Norte trouxe, por um lado, maior controle e autonomia gerencial para os diretores e maior capacidade de alterar o comportamento do diretor na direção esperada pelo programa. Por outro lado, os diretores sentiram-se com maior poder para influenciar as ações/práticas dos professores em sala de aula com a implementação de incentivos financeiros. De forma similar, outros estudos recentes mostram que mudanças em práticas docentes podem ser influenciadas por pressões geradas pela relação entre gestores e professores e pelo próprio perfil do diretor, uma vez que políticas de responsabilização criam novas funções para esses profissionais. Tais estudos revelam que a percepção dos gestores em relação às políticas de responsabilização pode influenciar suas ações sobre os professores (como maior cobrança por resultados, avaliação do desempenho docente e adoção de estratégias para aumentar o desempenho dos alunos) e, consequentemente, gerar mudanças nas práticas docentes[6] (GAWLIK, 2013; DERRINGTON; CAMPBELL, 2013; KIMBALL; MILANOWSKI, 2009).

No entanto, mesmo os defensores dos programas de *accountability* com base na escola reconhecem os pontos fracos de tais desenhos, como o efeito carona [*free-rider*].[7] Além disso, alguns estudos encontraram evidências de maiores

5 A relação entre o papel dos gestores escolares e o aprendizado discente é amplamente abordada pela literatura em eficácia escolar. Sammons (2008), por exemplo, listou, entre 11 fatores considerados características-chave de eficácia escolar, o fator *liderança profissional*. De acordo com a autora, essa liderança deve ter os seguintes aspectos: firme e objetiva; um enfoque participativo; um profissional que lidera. A autora sinaliza que: "O impacto que os diretores têm no desempenho e no progresso de seus alunos provavelmente opera indiretamente, ao invés de diretamente, através da influência que ele exerce na cultura da escola e dos professores, nas atitudes e comportamentos, os quais, por sua vez, afetam as práticas de sala de aula e a qualidade do ensino e aprendizagem" (SAMMONS, 2008, p. 355).

6 De acordo com Gawlik (2013), os gestores escolares sofrem pressões externas das políticas de responsabilização adotadas pelo governo, mas eles também incorporam as mensagens trazidas por elas e servem como mediadores entre as políticas e as atividades dentro da escola. O estudo observou que os diretores atuam como formadores de opinião (*sense-maker*) dentro das escolas, o que mudaria a forma como cada escola lida com as políticas de responsabilização existentes.

7 Para Ladd (2001), nos contextos que adotam políticas de responsabilização com foco na escola, se a escola é considerada eficaz, todos os professores são premiados. Assim, professores mais eficazes poderiam se ressentir, uma vez que, mesmo sem esforço, professores menos produtivos também receberiam a premiação.

impactos de programas de responsabilização com foco nos professores. Por exemplo, Lavy (2007) estudou o desenho de dois programas de responsabilização de professores adotados em Israel. Os resultados da pesquisa revelaram, ainda que de forma não conclusiva, que incentivos por bonificação individual podem aumentar o desempenho docente, sobretudo se forem acompanhados de perto para não induzirem a práticas de *gaming*. Figlio e Kenny (2007), em estudo de políticas de responsabilização de professores no contexto norte-americano, encontraram associações positivas entre políticas de remuneração por desempenho docente e discente. Os autores observaram que, entre as escolas analisadas, as que possuíam políticas de remuneração variável para professores alcançavam maior pontuação nos testes do que as que não adotavam tal procedimento.

Por fim, além das disputas sobre qual seria o melhor foco para as políticas de responsabilização, existe uma ampla discussão sobre as formas de medir o desempenho da escola e de seu pessoal (professores e gestão escolar). Pesquisas realizadas nos Estados Unidos e no Reino Unido descrevem diversas formas/medidas possíveis para avaliar escolas, diretores e/ou professores a partir do desempenho dos alunos. Segundo tais estudos, as medidas utilizadas trazem diferentes sinais e são percebidas como mais ou menos injustas e, portanto, incentivam a adoção de práticas e estratégias distintas por parte das escolas[8] (LADD, 2001; LADD; ZELLI, 2002; FIGLIO; LOEB, 2011; LIU; CHARLEY; FULLER, 2013; TIMMERMANS; DOOLAARD; WOLF, 2011).

Ladd (2001) discute as vantagens e desvantagens e os possíveis comportamentos incentivados por sistemas de *accountability* escolar que utilizam medidas baseadas em médias ou *status*, medidas baseadas em taxa de aumento do desempenho e medidas de crescimento ou de valor agregado. Para a autora, os desenhos de *accountability* escolar que empregam a primeira abordagem consideram eficazes as escolas que atingem determinada pontuação média ou garantem que uma porcentagem de alunos atinja um nível específico de proficiência. No entanto, para a autora, essa abordagem traz um incentivo fraco para as ações das escolas, uma vez que

8 Por exemplo, alguns estudos enfatizam a importância de que o pessoal (professores e diretores) das escolas avaliadas perceba o sistema como justo e equitativo, caso contrário, pode tornar-se mais propenso a ignorar e/ou trapacear o sistema de avaliação (LIU; CHARLEY; FULLER, 2013; LADD, 2001; LADD; ZELLI, 2002; FIGLIO; LOEB, 2011).

reflete mais o perfil dos alunos do que as práticas específicas das professores e diretores. Além disso, tal desenho poderia incentivar uma saída de professores e diretores de escolas de certo perfil em busca de escolas com maior probabilidade de atingir as metas estabelecidas (ou que possuam uma entrada de alunos "mais favorável").

Desenhos que utilizam o segundo tipo de medida, baseada em taxas de aumento do desempenho da escola, consideram eficazes as escolas que conseguem aumentar o desempenho dos alunos de acordo com uma meta estabelecida.[9] Um grande entrave para abordagens com base em ganhos no desempenho é que muitas vezes as metas refletem a vontade dos formuladores das políticas e não necessariamente são factíveis para as escolas (LADD, 2001; FIGLIO; LOEB, 2011).

9 Ladd (2001) menciona o exemplo do estado de Kentucky, que estabeleceu a mesma meta para todas as escolas e depois calculou a porcentagem de aumento a cada dois anos, para que as escolas atingissem a meta em 20 anos. De acordo com esse desenho, as escolas de baixo desempenho precisariam alcançar um crescimento mais elevado do que aquelas de alto desempenho.

Considerando que professores ou a direção da escola podem responder às pressões de responsabilização de formas tanto positivas como negativas, Figlio e Loeb (2011) argumentam que as respostas negativas seriam mais frequentes, em especial quando as escolas possuem recursos insuficientes para afetar os resultados e não têm conhecimento ou habilidade para alcançar as metas e/ou quando os programas estabelecem metas que não sejam factíveis.[10] Na mesma direção, no contexto brasileiro, o estudo de Alves e Soares (2013) ilustra alguns dos problemas relacionados ao uso de indicadores como o Ideb para estabelecer metas para políticas de responsabilização escolar, uma vez que o indicador focaliza resultados finalísticos, sem levar em conta caraterísticas das escolas que influenciam e estão associadas a esse indicador, tais como nível socioeconômico dos alunos, condições de infraestrutura e complexidade das escolas.

10 De acordo com os autores, "a falta de receptividade potencial à *accountability* pode ser exacerbada pelo fato de os sistemas de *accountability* geralmente concentrarem-se em melhoras de resultados no curto prazo enquanto várias políticas e práticas que as escolas desejam implementar podem levar mais tempo para alcançar resultados" (FIGLIO; LOEB, 2011, p. 402, tradução nossa).

Novamente, desenhos ou abordagens que utilizam medidas baseadas em taxas de crescimento não seriam capazes de enviar sinais claros, correndo o risco de punir professores e diretores que fazem um bom trabalho dados os recursos disponíveis e o alunado que recebem, simplesmente porque a meta não é realista. Além disso, as metas são baseadas em grupos diferentes de alunos, isto é, a composição do alunado pode mudar ano a ano, outro fator sobre o qual os diretores, em especial de escolas de baixo desempenho e menor prestígio, possuem pouco controle.

Por fim, programas que usam medidas de valor agregado tentam avaliar "o quanto da aprendizagem do aluno pode ser atribuída à escola" (LADD, 2001, p. 393), ou quanto os alunos melhoram sua aprendizagem de um ano para outro ou durante determinado período[11] (LADD, 2001; FIGLIO; LOEB, 2011). Para Ladd (2001), a maior vantagem dessa abordagem consiste em tentar medir ganho de aprendizagem e possuir um controle da *performance* prévia da escola, fato que contribuiria para o estabelecimento de metas de incremento mais realistas para as escolas. No entanto, as medidas de valor agregado enfrentam críticas similares às abordagens anteriores, quando não separam a contribuição do pessoal da escola de outros fatores relacionados à aprendizagem dos alunos, tais como a composição do alunado e os recursos materiais disponíveis para a escola. Esse problema é parcialmente resolvido por análises que usam medidas de valor agregado contextualizado (*Contextualized Value Added – CVA*), ou seja, que empregam controles de diversas características dos alunos e/ou das escolas.[12]

Mesmo com a crescente adoção de políticas de responsabilização que utilizam medidas de valor agregado contextualizado, a abordagem não está livre de críticas, sendo que as mais recorrentes argumentam que os modelos de valor agregados são fortemente afetados por erros aleatórios.[13] As inconsistências das medidas seriam ainda maiores para escolas de pequeno porte, devido à menor quantidade de alunos avaliados. Tais escolas teriam maior probabilidade de apresentar variação de desempenho de um ano para o outro e, portanto, de serem consideradas eficazes (LADD, 2001; FIGLIO; LOEB, 2011; GORARD; HORDOSY; SIDDIQUI, 2013). Por exemplo, o estudo de Gorard, Hordosy e Siddiqui (2013), focalizando escolas secundárias na Inglaterra, observou que as medidas de valor agregado contextualizado das escolas não são constantes ao longo do tempo.[14] Além de diversos problemas de medida, como dados faltantes, o estudo verificou que os valores extremos de medidas de CVA, em especial positivos, concentravam-se entre as menores escolas da Inglaterra.[15] Para os autores, os achados sugerem que, ao menos em parte, os resultados do valor agregado são consequência da "volatilidade das medidas de pequeno número/porte".

11 Timmermans, Doolard e Wolf (2011) e Gorard, Hordosy e Siddiqui (2013) trazem uma discussão mais aprofundada da abordagem de valor agregado utilizada em diversos modelos de *accountability* escolar.

12 De acordo com Timmermans, Doolaard e Wolf (2011), cada vez mais políticas de responsabilização têm utilizado modelos de valor agregado incluindo variáveis de controle relacionadas a diferenças de entrada de alunos entre escolas, mesmo que esses modelos variem no que diz respeito aos controles utilizados. Como exemplo, os autores mencionam os sistemas de *accountability* adotados no Reino Unido e no estado do Tennessee, nos Estados Unidos.

13 Os outros modelos discutidos não estão livres de erros aleatórios, entretanto, o problema seria mais pronunciado em modelos de valor agregado, já que dependem de medidas de dois anos, sendo que ambos os resultados conteriam erros (FIGLIO; LOEB, 2011; GORARD; HORDOSY; SIDDIQUI, 2013).

14 As análises revelaram que, depois de cinco anos, não havia uma relação clara entre a pontuação de valor agregado inicial e a final de grande parte das escolas. Para Gorard, Hordosy e Siddiqui (2013), a falta de constância das medidas ao longo do tempo seria uma evidência de que os modelos de valor agregado não estariam medindo características constantes ou mais permanentes das escolas, como, por exemplo, o efeito da escola.

15 O estudo de Gorard, Hordosy e Siddiqui (2013) encontrou uma correlação de Pearson de -0,22 entre o tamanho da escola e o desvio absoluto do CVA.

Consequentemente, programas de responsabilização escolar que utilizam medidas de valor agregado também teriam uma capacidade limitada de incentivar e premiar práticas eficazes.

Por fim, formas de medir a *performance* da escola, dos diretores ou dos professores utilizadas em políticas de responsabilização escolar emitem sinais para escolas. Portanto, as medidas vistas como injustas ou pouco equitativas poderiam promover ou incentivar práticas de *gaming* amplamente observadas em estudos sobre impactos de programas de responsabilização, tais como estratégias de seleção/exclusão, foco do planejamento pedagógico da escola em parte dos alunos, limitação do currículo e aumento das desigualdades intraescolares (DIAMOND; SPILLANE, 2004, WEST; PETERSON, 2006, HEILIG; DARLING-HAMMOND, 2008; ALLENDE, 2010; HSIEH; URQUIOLA, 2006).

AS POLÍTICAS DE (CO)RESPONSABILIZAÇÃO NO BRASIL E NO MUNICÍPIO DO RIO DE JANEIRO

Bonamino e Souza (2012) identificam três gerações de avaliação educacional de larga escala no contexto brasileiro. As autoras descrevem o Sistema de Avaliação da Educação Básica (Saeb) como expoente da primeira geração. O caráter amostral da avaliação não permitia seu uso em políticas de responsabilização da escola. A segunda geração de políticas de avaliação é caracterizada pela ampla divulgação dos desempenhos das escolas, mas ainda sem a introdução de elevadas consequências às escolas. As metas de Ideb para sistemas educacionais e escolas estabelecidas a partir do Plano de Desenvolvimento da Educação (PDE)[16] podem ser compreendidas como um exemplo dessa segunda geração de responsabilização branda (*low-stakes*). Já a terceira geração é marcada pelo uso de sistemas de avaliação para a implementação de políticas de responsabilização de altas consequências *(high-stakes),* com recompensas e/ou sanções atribuídas às escolas com base nos seus resultados.

O governo federal não estabelece, por meio do PDE, diretriz ou parâmetros para adoção de políticas de responsabilização

16 Iniciado em 2007, o PDE, entre outras medidas, institui o Índice de Desenvolvimento da Educação Básica (Ideb) como indicador de qualidade da educação. O PDE, além de prever a divulgação do Ideb, o utiliza para estabelecer metas bianuais a serem atingidas pelas redes e escolas até 2021.

por governos subnacionais, permitindo que esses estabeleçam seus próprios desenhos de política. Dessa forma, Brooke e Cunha (2011) observam que os sistemas educacionais têm feito diferentes usos dos sistemas de avaliação educacional, que vão da não adoção de políticas de responsabilização até a aplicação de desenhos mais sofisticados com diversos controles.

Em termos gerais, Brooke (2006) e Brooke e Cunha (2011) argumentam que, no Brasil, as políticas de responsabilização na educação envolvem certo grau de descentralização de responsabilidades, em que a escola e o professor não são os únicos responsáveis pelos desempenhos dos estudantes nas avaliações em larga escala, incluindo, nesse rol, diversos atores responsáveis por alcançar certas metas e padrões. Tais desenhos compreendem o apoio das secretarias de educação e instâncias intermediárias para que as escolas alcancem suas metas, ou até mesmo o estabelecimento de metas para essas instâncias superiores de gestão da educação (KOSLINSKI et al., no prelo).[17]

No que diz respeito às medidas de desempenho, os estados e municípios brasileiros têm adotado políticas de responsabilização escolar que utilizam seus próprios sistemas de avaliação. Se uma primeira geração de políticas traçava metas a partir do desempenho das escolas, os programas mais recentes usam indicadores que conjugam desempenho e fluxo e, até mesmo, controles relacionados ao nível socioeconômico das escolas (KOSLINSKI et al., no prelo; BROOKE; CUNHA, 2011). Em geral os desenhos empregam medidas baseadas em taxa de aumento do desempenho das escolas e comportam metas variadas que levam em conta desempenho da escola em anos anteriores (KOSLINSKI et al., no prelo). Por fim, uma breve análise das políticas de responsabilização descritas por Brooke e Cunha (2011) aponta para a tendência de adoção de desenhos de alta consequência, implementados a partir de incentivos ou bonificação salarial.

A política de responsabilização escolar adotada pelo município do Rio de Janeiro segue características presentes em diversos desenhos usados por outros estados e municípios brasileiros. O conjunto de medidas que englobam o sistema

17 A diretriz de corresponsabilização está presente, por exemplo, no PDE: "Avaliação, financiamento e gestão se articulam de maneira inovadora e criam uma cadeia de responsabilização pela qualidade do ensino que abrange tanto os gestores, do diretor ou da diretora da escola ao Ministro da Educação, quanto a classe política, do prefeito ou da prefeita ao presidente da República" (BRASIL, 2007, p. 25-6). Também está presente no "Acordo de Resultados" do governo do estado de Minas Gerais, que determina que o resultado das escolas é composto por uma parte referente às notas da Secretaria (SEE) e da Secretaria Regional de Educação (SRE), bem como por resultados da escola. "Por meio dessa composição mostra-se claramente que a responsabilidade pelos resultados finais da escola é compartilhada pelas três instâncias" (BROOKE; CUNHA, 2011, p. 54).

de avaliação próprio do município (Prova Rio)[18] e o estabelecimento de metas para as escolas (Prêmio Anual do Desempenho) foram implementados em 2009. Seguindo o mesmo modelo de indicadores elaborados com base no desempenho médio dos estudantes mais o fluxo escolar, o Índice de Desenvolvimento da Educação do Município do Rio de Janeiro (IDE-Rio) é originado a partir da avaliação aplicada aos estudantes do 3º e 7º anos. O Prêmio Anual do Desempenho é o principal componente da política de responsabilização do município e tem por objetivo "recompensar os servidores que conjuguem esforços com vistas à melhoria na qualidade do ensino e aprendizagem dos seus alunos" (RIO DE JANEIRO, 2009, Art. 3º).

O conjunto de medidas que compõem a política de responsabilização da SME do Rio de Janeiro combina foco na escola e em instâncias intermediárias. De um lado, é concedido um prêmio, em forma de 14º salário, a todos os servidores de escolas que alcançam ou superaram as metas estabelecidas. A escola é considerada de forma global: por exemplo, o recebimento do prêmio por servidores das escolas que oferecem tanto os anos iniciais quanto os anos finais do ensino fundamental depende do desempenho da escola em ambos os segmentos.[19] No entanto, o valor recebido por servidor pode variar, uma vez que as ausências do trabalho são contabilizadas no cálculo do prêmio.[20] De outro lado, nos anos seguintes à implantação do Prêmio Anual do Desempenho, a SME adotou medidas que trouxeram um caráter maior de corresponsabilização entre as instâncias superiores e as escolas, ao fixar metas para as Coordenadorias Regionais de Educação (CREs) e para a SME. A prefeitura do Rio de Janeiro adotou, em 2010, um Plano Estratégico de Governança Pública, traduzido pelo Acordo de Resultados, que fixa metas para as secretarias. Dessa forma, o sistema de responsabilização que institui metas foi ampliado para a SME, CREs e demais unidades administrativas (RIO DE JANEIRO, 2010). Diante disso, SME e CREs também são responsabilizadas, visto que algumas de suas metas englobam o desempenho das escolas.[21]

Além da característica de corresponsabilização, o Prêmio Anual de Desempenho apresenta um desenho com metas estabelecidas a partir de medidas baseadas em taxa

18 A Prova Rio segue a mesma matriz da Prova Brasil, com um modelo de Teoria de Resposta ao Item (TRI). Essa avaliação é censitária, e sua primeira aplicação contou somente com alunos do 3º e 7º anos do ensino fundamental nas disciplinas de Matemática e Língua Portuguesa. A partir de 2010, a Prova Rio passou a ser aplicada aos alunos do 3º, 4º, 7º e 8º anos, nas mesmas disciplinas. Os estudantes do 5º e 9º anos são avaliados a partir da Prova Brasil, que ocorre somente nos anos ímpares.

19 Até 2013, os servidores lotados em escolas que ofereciam ambos os segmentos recebiam o prêmio caso a escola alcançasse a meta de um dos segmentos e ao menos 50% da meta fixada para o outro.

20 Por exemplo, de acordo com o Decreto n. 30.860, de 1º de julho de 2009, que regulamenta o Prêmio Anual de Desempenho de 2009: "[p]ara fins de cálculo, visando à concessão do Prêmio, as ausências do servidor serão contabilizadas na forma a seguir: I. 0 a 2 dias de ausências - 100% do valor do prêmio; II 3 ou 4 dias de ausências - 80% do valor do prêmio; III. 5 dias de ausências - 50% do valor do prêmio."

21 No início de 2013, a prefeitura do Rio de Janeiro, com a intenção de institucionalizar um modelo de Gestão de Alto Desempenho, definiu cargos estratégicos com metas específicas a serem alcançadas (RIO DE JANEIRO, 2013a). Para a SME, as funções de coordenador regional, diretor geral de escola e gerente de educação, entre outras, foram nomeadas como cargos estratégicos. Todos os ocupantes de cargos estratégicos da prefeitura que alcançarem 100% das metas individuais receberão o equivalente ao 15º e 16º salários.

de aumento do desempenho. Recebem o prêmio escolas que alcançam a meta de acréscimo do Ideb e/ou do IDE-Rio. A Tabela 1 apresenta as metas estabelecidas para 2009, de acordo com o Decreto n. 30.860, de 1º de julho de 2009.

TABELA 1 – Metas de acréscimo no IDEB e/ou IDE-Rio das escolas para o ano de 2009

IDEB OU IDE-RIO DO ÚLTIMO PERÍODO	METAS DE ACRÉSCIMO	
	ANOS INICIAIS	ANOS FINAIS
2,0 a 3,0	25%	17%
3,1 a 4,0	15%	12%
4,1 a 5,0	12%	9%
5,1 a 6,0	5%	3%
Maior ou igual a 6,1	2%	1%

Fonte: Rio de Janeiro (2009).

A tabela instituída pela regulamentação do Prêmio Anual de Desempenho em 2009 dividiu as escolas em cinco faixas de Ideb ou IDE-Rio, de acordo com seus respectivos desempenhos em avaliações em períodos anteriores, estabelecendo metas mais audaciosas para escolas com desempenhos mais baixos, se comparadas com aquelas das escolas que apresentam indicadores de partida mais altos. As metas sofreram diversas modificações nas edições subsequentes do Prêmio Anual de Desempenho[22] (KOSLINSKI et al., no prelo). Em suma, as medidas de desempenho utilizadas são baseadas em taxas de crescimento do desempenho e não incluem controles relacionados aos recursos da escola e/ou composição do alunado, podendo incorrer, portanto, no perigo discutido por Ladd (2001) e Figlio e Loeb (2011) de estabelecer metas que não sejam realistas para algumas escolas e/ou deixar de premiar escolas por fatores que estão fora de seu controle.

22 Por se tratar de uma política muito recente, muitos ajustes têm sido feitos, principalmente nos critérios estabelecidos para o recebimento da premiação. Assim, a tabela de meta de acréscimo que divide as escolas por grupos de desempenho e suas respectivas porcentagens sofreram modificações ao longo dos anos, tanto no número de grupos que compõem a tabela, como na taxa de crescimento para cada grupo.

METODOLOGIA E ALGUMAS EVIDÊNCIAS SOBRE O PERFIL DAS ESCOLAS PREMIADAS

As análises apresentadas a seguir utilizaram dados dos questionários contextuais da Prova Brasil de 2009 e 2011 (MEC/Inep),

do Ideb de 2005, 2007, 2009 e 2011 (MEC/Inep) e bases de dados da Secretaria Municipal de Educação do Rio de Janeiro com informações demográficas dos alunos e sobre a divulgação do resultado do Prêmio Anual do Desempenho de 2009 e 2011. As análises focalizaram somente as escolas que tinham metas estabelecidas pelo Prêmio Anual de Desempenho para os anos iniciais em 2009 e 2011. Assim, foram analisadas 707 escolas, para 2009, e 685, para 2011.

As análises utilizaram modelos de regressão logística para estimar a probabilidade de as escolas alcançarem as metas e, portanto, de receberem o Prêmio Anual de Desempenho em 2009 e 2011, considerando as seguintes características: perfil da gestão da escola; escolaridade dos pais dos alunos; percepção dos professores sobre carências de recursos da escola; desempenho prévio da escola (Ideb prévio e taxa de crescimento anterior); e outras características relacionadas à escola que podem influenciar sua capacidade de alcançar as metas estabelecidas pela SME, levando em conta o desenho do programa (número de alunos que realizaram a prova, complexidade da escola).

Como já explicitado, o foco das análises é observar quais das características descritas anteriormente estão associadas às chances de as escolas alcançarem as metas do Prêmio Anual de Desempenho, no intuito de identificar possíveis sinais ou incentivos que o programa envia para os atores escolares. No entanto, o impacto do programa nas práticas dos diretores, por exemplo, não é objeto do presente estudo.

O Quadro 1 e a Tabela 2 apresentam a descrição das variáveis utilizadas nos modelos de análise.

QUADRO 1 – Descrição das variáveis utilizadas nos modelos

VARIÁVEL	TIPO	DESCRIÇÃO
VARIÁVEIS DEPENDENTES		
Prêmio em 2009	dicotômica	Escola recebeu o prêmio em 2009 (1 = sim/ 0 = não)
Prêmio em 2011	dicotômica	Escola recebeu o prêmio em 2011 (1 = sim/ 0 = não)
VARIÁVEIS EXPLICATIVAS		
Liderança[1]	contínua	Análise fatorial com itens relacionados à percepção dos professores sobre a liderança do diretor (Alpha de Cronbach = 0,848)
Escolaridade dos pais (média da escola)	contínua	Porcentagem de alunos cujos responsáveis possuem alta escolaridade (ensino médio completo ou mais)
Percepção sobre carência de recursos[2]	contínua	Análise fatorial com itens sobre a percepção dos professores sobre a falta de recursos da escola (Alpha de Cronbach = 0,749)
Ideb anterior da escola		
baixo	dicotômica	Ideb da escola no ano anterior ao prêmio igual ou menor a 4,0 (1 = sim/ 0 = não)
médio	dicotômica	Ideb da escola no ano anterior ao prêmio entre 4,1 e 5,0 (1 = sim/ 0 = não)
Número de alunos participantes	numérica	Número de alunos do 5º ano da escola que participaram da Prova Brasil
Crescimento do Ideb anterior	contínua	Porcentagem de crescimento do Ideb entre 2005 e 2007 ou entre 2007 e 2009
Complexidade da escola	dicotômica	Escola oferece anos iniciais e anos finais do ensino fundamental (1 = sim/ 0 = não)

Fonte: Elaboração dos autores.

(1) A variável liderança foi construída a partir das questões do questionário de professores da Prova Brasil 2009 e 2011, que indagavam seu grau de concordância com as seguintes afirmativas: O(A) diretor(a) me anima e me motiva para o trabalho; Tenho plena confiança no(a) diretor(a) como profissional; O(A) diretor(a) consegue que os professores se comprometam com a escola; O(A) diretor estimula as atividades inovadoras; O(A) diretor(a) dá atenção especial a aspectos relacionados com a aprendizagem dos alunos; Sinto-me respeitado(a) pelo(a) diretor(a); Respeito o(a) diretor(a); O(A) diretor(a), professores e demais membros da equipe da escola colaboram para fazer esta escola funcionar bem.

(2) O indicador foi construído a partir de respostas às seguintes questões do questionário dos professores da Prova Brasil 2009 e 2011 sobre a ocorrência de problemas graves que dificultaram o funcionamento da escola: insuficiência de recursos financeiros; inexistência de professores para algumas disciplinas ou séries; carência de pessoal administrativo; carência de pessoal de apoio pedagógico (coordenador, supervisor, orientador educacional); falta de recursos pedagógicos; alto índice de faltas por parte de professores; alto índice de faltas por parte de alunos; problema disciplinares causados pelos alunos.

TABELA 2 – Medidas descritivas das variáveis utilizadas no modelo nos dois anos avaliados

	2009		2011	
VARIÁVEIS QUANTITATIVAS	**MÉDIA**	**D.P.**	**MÉDIA**	**D.P.**
Liderança	0,06	0,81	0,06	0,78
Escolaridade dos pais (média da escola)	33,1	15,8	35,3	15,7
Percepção sobre carência de recursos	-0,12	0,96	-0,10	0,88
Número de alunos participantes	68,5	38,3	69,5	35,1
Crescimento do Ideb anterior	7,5	13,9	11,8	14,2
VARIÁVEIS QUALITATIVAS	**PROPORÇÃO**		**PROPORÇÃO**	
Prêmio	0,41		0,52	
Ideb anterior da escola				
baixo	0,19		0,04	
médio	0,58		0,39	
alto	0,23		0,57	
Complexidade da escola	0,30		0,30	

Fonte: Elaboração dos autores.

A Tabela 2 indica que, na edição de 2009, em torno de 41% das escolas da rede que tinham metas para os anos iniciais receberam o Prêmio Anual de Desempenho, proporção que cresceu para cerca de 52%, em 2011. Entre 2009 e 2011, verificam-se, também, aumento de 23% para 57% das escolas com alto desempenho no ano de referência (Ideb dos anos iniciais maior que 5,0) e redução de 19% para 4% daquelas com baixo desempenho (Ideb dos anos iniciais igual ou inferior a 4,0.

As tabelas 3 e 4 apresentam os resultados dos modelos de regressão logística que estimam a probabilidade de as escolas receberem o Prêmio Anual de Desempenho em 2009 e 2011. Trazem os resultados da exponenciação do coeficiente (Exp(B)) associados às variáveis contidas em cada modelo que expressam a razão de chance (RC) [*odds ratio – OR*] ou a mudança nas chances de receber o Prêmio Anual de Desempenho associadas à alteração de categoria ou elevação de uma unidade das variáveis explicativas.[23]

23 Para uma descrição mais detalhada sobre os modelos de regressão logística e interpretação dos coeficientes, ver Powers e Xie (2000).

TABELA 3 – Estimativas (e erros-padrão) da Razão de Chances nos cinco modelos de regressão logística utilizados para estimar a probabilidade das escolas receberem o Prêmio Anual de Desempenho em 2009

VARIÁVEIS	MODELOS				
	1	2	3	4	5
Liderança	1,406*** (1,111)	1,285* (1,128)	1,280* (1,129)	1,373* (1,138)	1,385* (1,151)
Escolaridade dos pais (média da escola)		1,010+ (1,006)	1,016* (1,007)	1,007 (1,007)	1,018* (1,008)
Percepção sobre carência de recursos		0,597*** (1,119)	0,559*** (1,122)	0,566*** (1,132)	0,994 (1,154)
Ideb anterior da escola					
baixo			2,098* (1,378)	0,993 (1,424)	1,671 (1,506)
médio			0,980 (1,267)	0,659 (1,287)	0,761 (1,332)
Crescimento do Ideb anterior				0,958*** (1,009)	0,960*** (1,010)
Número de alunos participantes					0,999 (1,002)
Complexidade da escola					0,033*** (1,506)
Constante	0,623*** (1,086)	0,387*** (1,234)	0,272*** (1,434)	0,703 (1,505)	0,912 (1,648)

Fonte: Elaboração dos autores.

Nota: Cada célula da tabela contém a estimativa da Razão de Chances da variável presente no modelo e abaixo, entre parênteses, o erro-padrão dessa estimativa. Os símbolos ao lado da Razão de Chances indicam o nível descritivo (p-valor) em que essas estimativas foram significantes, sendo que $^{+}$ indica p<0,1; * indica p<0,05; ** indica p<0,01 e *** indica p<0,001 .

TABELA 4 – Estimativas (e erros-padrão) da Razão de Chances nos cinco modelos de regressão logística utilizados para estimar a probabilidade de as escolas receberem o Prêmio Anual de Desempenho em 2011

VARIÁVEIS	MODELOS				
	1	2	3	4	5
Liderança	1,058 (1,104)	1,023 (1,111)	1,029 (1,112)	1,038 (1,117)	1,046 (1,119)
Escolaridade dos pais (média da escola)		1,018*** (1,005)	1,029*** (1,006)	1,021*** (1,006)	1,020*** (1,006)
Percepção sobre carência de recursos		0,954 (1,098)	0,892 (1,101)	0,886 (1,105)	0,933 (1,110)
Ideb anterior da escola					
baixo			5,540*** (1,585)	2,263 (1,676)	1,499 (1,698)
médio			1,791*** (1,198)	1,037 (1,225)	1,014 (1,234)
Crescimento do Ideb anterior				0,958*** (1,007)	0,957*** (1,007)
Número de alunos participantes					0,986*** (1,003)
Complexidade da escola					0,858 (1,217)
Constante	1,072 (1,080)	0,567** (1,215)	0,289*** (1,295)	1,251 (1,357)	2,506* (1,474)

Fonte: Elaboração dos autores.

Nota: Cada célula da tabela contém a estimativa da Razão de Chances da variável presente no modelo e abaixo, entre parênteses, o erro-padrão dessa estimativa. Os símbolos ao lado da Razão de Chances indicam o nível descritivo (p-valor) em que essas estimativas foram significantes, sendo que $^{+}$ indica p<0,1; * indica p<0,05; ** indica p<0,01 e *** indica p<0,001.

O modelo 1 da Tabela 3 indica que quanto maior o indicador de liderança, ou seja, a intensidade com que os professores reconhecem a liderança do diretor, maior é a probabilidade de a escola ganhar o Prêmio Anual de Desempenho em 2009. O aumento de uma unidade na variável de liderança está associado a um incremento em torno de 41% das chances de ganhar o prêmio. Tal associação se mantém significativa nos modelos subsequentes, mesmo após a inserção de outras variáveis referentes a composição do alunado, carência de recursos, desempenho prévio, tamanho e complexidade das escolas (Tabela 3, modelos 2, 3, 4 e 5). No entanto, de acordo com os resultados obtidos no modelo 1 da Tabela 4, a associação entre liderança do diretor e a probabilidade de receber o prêmio não foi observada para 2011 (em todos os modelos a razão de chances é próxima a 1).

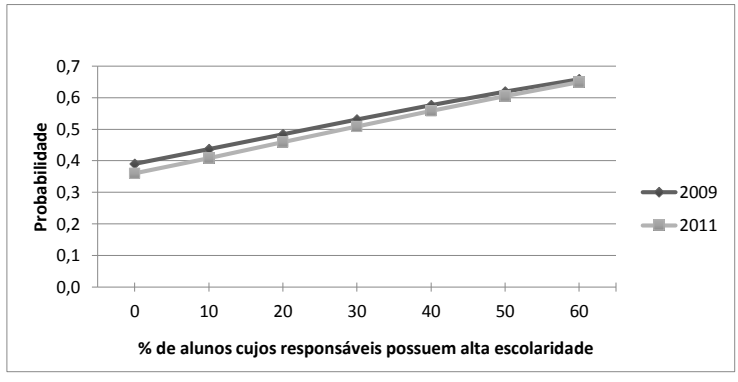
Em uma análise preliminar, os resultados sugerem que a liderança dos diretores deixa de ser recompensada em 2011. No entanto, os resultados precisam ser avaliados com cautela, tendo em vista as restrições da capacidade do questionário de medir com maior precisão o perfil da liderança da escola, uma vez que somente os professores das séries avaliadas respondem ao questionário.[24]

O modelo 2 das tabelas 3 e 4 mostram os resultados de regressões logísticas com a inserção de controles relacionados à composição do alunado (porcentagem de pais com alta escolaridade) e às percepções dos professores sobre carência de recursos da escola. Como esperado, em 2009, a porcentagem de alunos cujos responsáveis possuem alta escolaridade está positivamente associada às chances de receber o prêmio (Tabela 3, modelo 2). O modelo também estima uma associação negativa entre a percepção dos professores sobre carências de recursos da escola e as chances de ganhar o prêmio em 2009. As tendências observadas mantêm-se nos modelos subsequentes (Tabela 3, modelos 3, 4) com a inserção de outros controles referentes ao desempenho prévio das escolas. No entanto, no modelo 5, com a inclusão de controles relacionados ao número de alunos que participaram da Prova Brasil 2009 e à complexidade da escola, a associação entre a percepção de carências e a chance de ganhar o Prêmio não é observada.

Em 2011, verifica-se uma tendência similar à de 2009, no que diz respeito à composição do alunado: quanto maior a porcentagem de alunos com pais de alta escolaridade, maior é a probabilidade de a escola ganhar o prêmio (Tabela 4, modelo 2). Os resultados são mais facilmente interpretados a partir das probabilidades de ganhar o prêmio estimadas pelo modelo 5 em cada ano (Gráfico 1).

24 Paes de Carvalho *et al.* (2012) argumentam que esse é um dos principais entraves para a obtenção de uma medida mais precisa de liderança por meio dos questionários contextuais da Prova Brasil. Diretores podem designar seus professores "de confiança" para as séries avaliadas e, portanto, esses tenderiam a ter visões mais positivas em relação a seus diretores. Essa tendência pode ter sido aumentada com a introdução de políticas de responsabilização escolar. Novamente, tais especulações vão além do escopo do presente estudo e o uso dessa estratégia pela direção das escolas será foco de pesquisas posteriores.

GRÁFICO 1 – Probabilidade predita de as escolas receberem o Prêmio Anual de Desempenho em 2009 e 2011 de acordo com as variações na porcentagem de alunos cujos responsáveis possuem alta escolaridade (ensino médio completo ou mais)[25]



Fonte: Elaboração dos autores.

25 Os valores estimados das probabilidades foram calculados a partir do modelo 5 das tabelas 3 e 4 para escolas com Ideb base (2007 ou 2009) maior que 5,0, que oferecem somente os anos iniciais, e as demais variáveis (liderança, percepção sobre carência, crescimento prévio do Ideb e número de participantes) foram mantidas constantes, em seu valor médio.

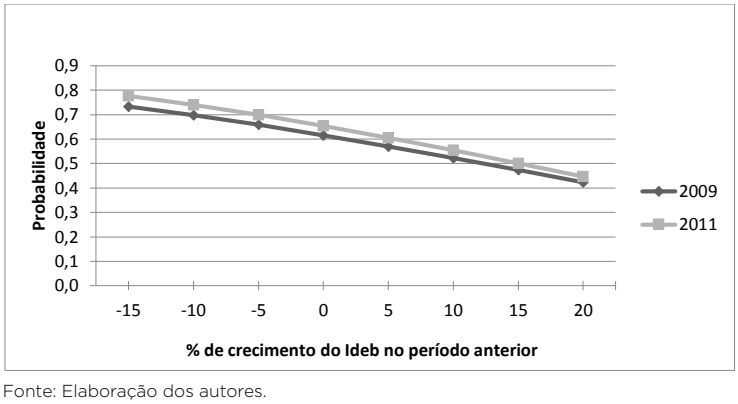
De acordo com o Gráfico 1, em ambos os anos, os modelos estimam um incremento gradual da probabilidade de receber o bônus para as escolas com maior concentração de alunos com pais de alta escolaridade. Por exemplo, em 2009, controlando pelas demais variáveis, a probabilidade predita de ganhar o prêmio era de 0,44 e de 0,62 para escolas com, respectivamente, 10% e 50% dos alunos cujos responsáveis apresentavam alta escolaridade. Para 2011, tais probabilidades correspondiam a 0,41 e 0,61, respectivamente, para esses grupos de alunos. No entanto, em 2011, os modelos estimados não mostram uma associação entre a percepção de carência de recursos pelos professores e a probabilidade de as escolas receberem o prêmio (Tabela 4, modelo 2).

Os modelos 3 e 4 (Tabela 3) incluem controles relacionados ao desempenho prévio das escolas. Considerando-se que escolas com diferentes pontos de partida de Ideb possuem metas de acréscimo distintas, partiu-se da hipótese de que as metas estabelecidas seriam mais difíceis de serem alcançadas para escolas com Ideb, em 2007, igual ou inferior a 4,0. Tais escolas teriam que alcançar um crescimento do Ideb equivalente a 15% ou mais (Tabela 1). Ao contrário do esperado, as escolas com Ideb de partida mais baixo apresentaram duas vezes mais chances de receber o prêmio em 2009 do que aquelas com Ideb maior que 5,0 (cuja meta de

acréscimo seria equivalente a 5% ou menos). No entanto, a variável relativa ao Ideb de partida deixa de ser significativa com a inserção da variável de aumento do Ideb da escola no período anterior ao prêmio (entre 2005 e 2007). O modelo 4 (Tabela 3) estima que, quanto maior a taxa de aumento do Ideb alcançado pelas escolas entre 2005 e 2007, menor é a chance de atingir a meta e ganhar o prêmio em 2009.

Tendência similar foi observada em 2011: o modelo 3 (Tabela 4) estima que escolas com Ideb em 2009 igual ou menor a 4,0 e entre 4,1 e 5,0 tinham, respectivamente, em torno de 5 vezes e quase 2 vezes mais chances de ganhar o prêmio em 2011 se comparadas àquelas com Ideb de partida superior a 5,0. Ou seja, no que diz respeito ao desenho do programa, as metas estipuladas pareciam mais difíceis de serem alcançadas, ou menos factíveis, para as escolas de alto desempenho. Novamente, esta associação não é observada com a inserção da variável relacionada à taxa de crescimento do Ideb das escolas entre 2007 e 2009. Aquelas com maior aumento do Ideb e que provavelmente ganharam o prêmio em 2009 apresentaram menor chance de receber o prêmio em 2011. O Gráfico 2 apresenta as probabilidades preditas de ganhar o prêmio em 2009 e 2011 para escolas com diferentes taxas de crescimento do Ideb nos períodos anteriores.

GRÁFICO 2 – Probabilidade predita de as escolas receberem o Prêmio Anual de Desempenho em 2009 e 2011 de acordo com as variações na taxa de crescimento do Ideb em períodos anteriores [26]



Fonte: Elaboração dos autores.

26 Os valores estimados das probabilidades foram calculados a partir do modelo 5 das tabelas 3 e 4 para escolas com Ideb base (2007 ou 2009) maior que 5,0, que oferecem somente os anos iniciais, e as demais variáveis (liderança, percepção sobre carência, escolaridade dos pais, número de participantes) foram mantidas constantes, em seu valor médio.

De acordo com o Gráfico 2, controlando por outras variáveis, a probabilidade predita de ganhar o prêmio em 2009 foi de 0,70, para as escolas que tiveram um crescimento negativo de 10% do Ideb (entre 2005 e 2007), e de 0,52, para aquelas com crescimento positivo de 10% do Ideb. As probabilidades preditas de ganhar o prêmio em 2011, para escolas com essas mesmas características de crescimento do Ideb (entre 2007 e 2009), foram equivalentes a 0,74 e 0,55, respectivamente.

Esta associação mostra que, mesmo que as metas estabelecidas não pareçam ser inalcançáveis, ou mais "difíceis" de serem atingidas para as escolas de baixo desempenho, talvez o programa, no formato implementado em 2009 e 2011, estipule metas "intangíveis" para escolas que aumentaram muito seu desempenho em períodos anteriores. As análises realizadas no presente trabalho não permitem observar se essas tendências são percebidas e se influenciam a conduta dos atores escolares. Mas, caso sejam percebidas, elas podem sinalizar ou motivar certo tipo de comportamento, como o de que não vale a pena fazer um grande esforço, ou esforço além do necessário, para atingir a meta em determinado ano; ou mesmo que a diminuição do Ideb ou do IDE-Rio em um ano pode aumentar as chances de alcançar as metas e ganhar o prêmio no período seguinte.

Por fim, no modelo 5 (Tabela 3 e 4) foram incluídas duas novas variáveis referentes ao número de alunos de 5º ano que participaram da Prova Brasil e à complexidade da escola (se oferece somente anos iniciais ou anos iniciais e finais). Os resultados observados são diferentes para o prêmio em 2009 e 2011.

Em 2009, controlando pelas demais variáveis incluídas no modelo, o número de alunos que participaram da Prova Brasil não parece exercer impacto na chance de as escolas receberem o prêmio (Tabela 3, modelo 5). Já em 2011, o modelo estima que, quanto maior o número de alunos participantes, menor é a chance de a escola ganhar o prêmio (Tabela 4, modelo 5). Uma análise das probabilidades preditas pelo modelo ajuda a compreender os resultados obtidos. Controlando por outras variáveis,[27] a probabilidade predita de ganhar o prêmio em 2011 para uma escola de pequeno porte, com 30

27 As probabilidades preditas foram calculadas a partir do modelo 5 da Tabela 3 para escolas com Ideb base (2007 ou 2009) superior a 5,0, que oferecem somente os anos iniciais, e as demais variáveis (liderança, percepção sobre carências, crescimento do Ideb anterior e escolaridade dos pais) na média.

alunos de 5º ano que realizaram a Prova Brasil, foi de 0,67. Essa probabilidade cai para 0,57 e 0,43 para escolas com 60 e 100 alunos participantes, respectivamente. A associação entre as escolas de menor porte e o recebimento do prêmio em 2011 pode ser decorrente de diversos fatores: desde o que Gorard, Hordosy e Siddiqui (2013) chamam de "volatilidade das medidas de pequeno número", até a menor complexidade da gestão e/ou maior possibilidade de manipulação de resultados a partir da exclusão de alunos por escolas de menor porte. Uma vez mais, novas investigações são necessárias para observar se essa associação se mantém em outros anos do prêmio e os mecanismos pelos quais ela opera.

Por fim, em 2009, a complexidade da escola foi determinante na probabilidade de as escolas receberem o prêmio: aquelas que possuíam somente o primeiro segmento tinham em torno de 30 vezes mais chances de receberem o prêmio se comparadas com as escolas que ofereciam os dois segmentos (Tabela 3, modelo 5). Apesar de muito forte em 2009, tal associação não foi observada em 2011 (Tabela 4, modelo 5).

Uma descrição da porcentagem de escolas que ganharam o prêmio em cada ano, de acordo com os segmentos que oferecem, facilita a interpretação dos resultados obtidos no modelo 5 (Tabelas 3 e 4). Em 2009, em torno de 55% (274 em 498) das escolas que só ofereciam os anos iniciais ganharam o prêmio e apenas 6% (12 entre 208 escolas) daquelas que dispunham dos anos iniciais e finais receberam o prêmio. O resultado pode ser interpretado a partir de distintos mecanismos, tanto pela complexidade da administração de uma escola que possui anos iniciais e finais, quanto pelo fato de que, para as escolas que oferecem ambos os segmentos do ensino fundamental, o Prêmio Anual de Desempenho estipula duas metas: uma referente ao Ideb de anos iniciais e outra para os anos finais. Em 2009, 75 escolas que possuíam ambos os segmentos conseguiram alcançar suas metas de Ideb para os anos iniciais, mas não atingiram ao menos 50% da meta dos anos finais.[28] Portanto, tal evidência permite supor que ter duas metas, e principalmente metas para os anos finais, foi o principal entrave para as escolas receberem o prêmio em 2009. Já em 2011, 51% das escolas que ofereciam somente

28 Tendência similar foi observada entre as escolas que só ofereciam anos finais: somente 6% delas (9 em 156) alcançaram suas metas e ganharam o Prêmio Anual de Desempenho em 2009.

anos iniciais e 53% das que tinham ambos os segmentos do ensino fundamental receberam o prêmio. A existência de duas metas não foi um entrave para as escolas que oferecem anos iniciais e finais receberem o prêmio em 2011.[29]

29 A descrição do crescimento do Ideb de anos finais das escolas municipais do Rio de Janeiro permite uma melhor compreensão das tendências observadas em 2009 e 2011. Em 2009, houve uma queda do Ideb de anos finais de 4,3 para 3,6, explicada principalmente pela diminuição do fluxo escolar. Já em 2011, o Ideb de anos finais aumentou para 4,2. No entanto, novos estudos precisam ser realizados para melhor compreender como as escolas conseguiram voltar a aumentar o Ideb em 2011, e se essa é uma tendência que se mantém para 2013.

CONSIDERAÇÕES FINAIS

O presente estudo realizou uma análise exploratória do perfil das escolas que alcançaram as metas e foram premiadas por uma política de responsabilização escolar adotada pela Secretaria Municipal de Educação do Rio de Janeiro. O objetivo da exploração foi identificar possíveis sinais emitidos às escolas por um programa com características de corresponsabilização entre instâncias intermediárias e a escola e com metas estabelecidas a partir de taxa de aumento do desempenho. As análises realizadas, como em estudos produzidos no contexto internacional, mesmo com as limitações que os dados disponíveis impuseram à presente investigação, identificaram dificuldades e limitações para premiar e, possivelmente, incentivar práticas escolares eficazes. Se, no primeiro ano do programa aqui estudado, o perfil da liderança das escolas parece ter exercido impacto sobre a probabilidade de as escolas ganharem o prêmio, outras características, muitas delas fora do controle da gestão escolar (ou de outros atores escolares), também tiveram peso na distribuição do prêmio, em ao menos uma das edições estudadas. Tais características abarcaram desde aspectos relacionados ao perfil do alunado, recursos e complexidade das escolas, até o número de alunos que realizaram a prova. Além disso, escolas com maior desempenho e, em especial, as que obtiveram crescimento mais acentuado do Ideb em períodos anteriores apresentaram menores chances de alcançar as metas e ganhar o prêmio em ambos os anos estudados.

Como argumentado anteriormente, o estudo partiu do pressuposto de que o desenho das políticas de responsabilização escolar emite diferentes sinas, incentivando diferentes tipos de estratégias por parte das escolas. Isto é, a percepção sobre o perfil das escolas e/ou sobre a justiça dos critérios para a premiação pode influenciar a forma como as escolas se adaptam às pressões de responsabilização.

Entretanto, se, de um lado, o estudo parte do pressuposto de que as escolas estão aprendendo e se adaptando aos incentivos das políticas de responsabilização escolar, de outro, as mudanças da legislação que regulamenta o programa indicam que os formuladores da política também se adaptam para lidar com os problemas e/ou consequências não intencionadas geradas pelo programa. Por exemplo, em edições mais recentes do Prêmio Anual de Desempenho, as escolas com Ideb de partida para os anos iniciais de 6,5 não têm meta de acréscimo, somente precisam manter o último índice alcançado (RIO DE JANEIRO, 2011, 2013b). Tais medidas podem amenizar o impacto do crescimento prévio sobre a chance de recebimento do prêmio em anos subsequentes, ao menos para as escolas de desempenho mais alto. Outra recente modificação foi o estabelecimento de uma nova condição para o recebimento do prêmio: as escolas precisam garantir uma taxa de participação dos alunos igual ou superior a 85% na Prova Brasil e/ou Prova Rio (RIO DE JANEIRO, 2013b). Tal medida pode coibir estratégias de seleção de alunos para a realização das avaliações externas de referência do Prêmio Anual de Desempenho. Por fim, a partir da edição de 2013, a escola deixa de ser tratada com uma unidade global e os professores e funcionários de escolas que possuem metas para anos iniciais e finais podem ganhar o prêmio referente a um segmento, independentemente do resultado do outro segmento (RIO DE JANEIRO, 2013b). Tal medida pode evitar a forte tendência, observada em 2009, de diminuição das chances de as escolas que oferecem ambos os segmentos de serem premiadas. No entanto, como argumenta Ladd (2001), ao não tratar a escola como uma unidade para a premiação, pode-se deixar de incentivar a cooperação entre professores de diferentes segmentos para atingir uma meta comum.

Novamente, essas são especulações sobre possíveis consequências de políticas de responsabilização adotadas no contexto brasileiro. Novos estudos são necessários para observar o impacto de tais políticas, e seus diferentes desenhos, tanto no desempenho quanto nas práticas/estratégias da gestão escolar e nas práticas docentes.

REFERÊNCIAS

ALLENDE, C. *The impact of information on academic achievement and school choice*: evidence from Chilean "traffic lights". 2010. Dissertação (Mestrado em Economia) – Instituto de Economia, Pontificia Universidad Católica de Chile, Santiago, 2010.

ALVES, M. T. G.; SOARES, J. F. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. *Educação e Pesquisa*, São Paulo, v. 39, n. 1, p. 177-194, jan./mar. 2013.

BONAMINO. A.; SOUZA, S. K. Três gerações de avaliação da educação básica no Brasil. *Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012.

BRASIL. Ministério da Educação. *O Plano de Desenvolvimento da Educação*: razões, princípios e programas. Brasília: MEC, 2007.

BROOKE, N. O futuro das políticas de responsabilização educacional no Brasil. *Cadernos de Pesquisa*, São Paulo, v. 36, n. 128, p. 377-401, maio/ago. 2006.

_____. Controvérsias sobre políticas de alto impacto. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 148, p. 336-347, jan./abr. 2013.

BROOKE, N.; CUNHA, M. A. A. A avaliação externa como instrumento de gestão educacional nos estados. *Estudos & Pesquisas Educacionais*, São Paulo, v. 2, p. 3-64, 2011.

DERRINGTON, M.; CAMPBELL, J. The changing conditions of instructional leadership: principal's perceptions of teacher evaluation accountability measures. In: BARNETT, B.; SHOHO, A. R.; BOWERS, A. J. (Ed.). *School and district leadership in an era of accountability*. Charlotte: Information Age Publishing, 2013. p. 231-251.

DIAMOND, J. B.; SPILLANE, J. P. High stakes accountability in urban elementary schools: challenging or reproducing inequality? *Teachers College Records*, v. 106, n. 6, p. 1145-1176, 2004.

FIGLIO, D.; KENNY, L. W. Individual teacher incentives and student performance. *Journal of Public Economics*, n. 91, p. 901-914, 2007.

FIGLIO, D.; LOEB, S. School accountability. In: HANUSHEK, E. A.; MACHIN, S.; WOESSMANN, L. (Ed.). *Handbook in economics of education*. North-Holland: Elselvier, 2011. v. 3.

FREITAS, L. C. Políticas de responsabilização: entre a falta de evidência e a ética. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 148, p. 348-365, jan./abr. 2013.

GAWLIK, M. Extending the leadership role from policy broker to sense maker: emerging evidence from charter schools. In: BARNETT, B.; SHOHO, A. R.; BOWERS, A. J. (Ed.). *School and district leadership in an era of accountability*. Charlotte: Information Age Publishing, 2013. p. 87-108.

GORARD, S.; HORDOSY, R.; SIDDIQUI, N. How unstable are 'school effects' assessed by a value-added technique? *International Education Studies*, v. 6, n. 1, p. 1-9, 2013.

HEILIG, J. V.; DARLING-HAMMOND, L. Accountability Texas-Style: the progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, v. 30, n. 2, 2008.

HOFFER, T. B. Accountability in education. In: HALINNAN, M. (Ed.). *Handbook of the sociology of education*. New York: Kluwer Academic, 2000. p. 529-543.

HOUT, M.; ELIOTT, S. (Ed.). *Incentives and test-based accountability in education*. Washington, D.C.: National Research Council, 2011.

HSIEH, C.; URQUIOLA, M. The effects of generalized school choice on achievement and stratification: evidence from Chile's Voucher Program. *Journal of Public Economics*, v. 90, p. 1477-1503, 2006.

KIMBALL, S.; MILANOWSKI, A. Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Education Administration Quarterly*, v. 45, n. 1, p 34-70, 2009.

KOSLINSKI, M. C.; CARRASQUEIRA, K.; ANDRADE, F. M.; CUNHA, C. P. Políticas de responsabilização educacional: modelos possíveis e uma análise preliminar da política do município do Rio de Janeiro. In: CAVALIERI, A. M.; SOARES, A. J. *Políticas educacionais no Estado do Rio de Janeiro*. No prelo.

LADD, H. F. School-based educational accountability systems: the promise and the pitfalls. *National Tax Journal*, v. 54, n. 2, 2001.

LADD, H. F.; ZELLI, A. School-based accountability in North Carolina: the responses of school principals. *Educational Administration Quarterly*, v. 38, n. 4, p. 494-529, 2002.

LAVY, V. Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy*, v. 110, n. 6, p. 1286-1317, 2002.

_____. Using performance-based pay to improve the quality of teachers. *The Future of Children*, v. 17, n. 1, 2007.

LIU, J.; CHARLEY, E.; FULLER, E. *Evaluating principals*: an analysis of current state strategies. In: UCEA CONVENTION, 2013, Indianapolis. *Paper presented*… Indianápolis: UCEA, 2013.

MCEWAN, P. J.; URQUIOLA, M.; VEGAS, E.; GALLEGO, F. A. School choice, stratification and information on school performance: lessons from Chile. *Economia*, v. 8, n. 2, p. 1-27, 2008.

MURALIDHARAN, K.; SUNDARAMAN, V. *Teacher performance pay*: experimental evidence from India. Cambridge, MA: National Bureau of Economic Research, 2009. (NBER working paper, 15323).

PAES DE CARVALHO, C.; WALHELM, A. P. S.; ALVES, F.; KOSLINSKI, M. *Gestão e desempenho escolar*: um estudo nas redes municipais da Região Metropolitana do Rio de Janeiro a partir dos resultados da Prova Brasil 2009. In: CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 3., 2012, Zaragoza. *Anais*... Zaragoza, ES: ANPAE, FEAE, FPAE, 2012.

PONTUAL, T. C. *Remuneração por mérito, desafio para a educação.* São Paulo: Fundação Lemman, 2008. Estudo comissionado pela Fundação Lemman.

POWERS, D. A.; XIE, Y. *Statistical methods for categorical data analysis*. San Diego, CA: Academic Press, 2000.

REYNOLDS, D.; TEDDLIE, C. O processo da eficácia escolar. In: BROOKE, N.; SOARES, J. F. *Pesquisa em eficácia escolar:* origem e trajetórias. Belo Horizonte: Editora UFMG, 2008.

RIO DE JANEIRO (Município). Decreto n. 30.860, de 1º de julho de 2009. *Diário Oficial do Município do Rio de Janeiro*, Ano XXIII, n. 71, 02 de julho de 2009.

_____. Decreto n. 32.214, de 4 de maio de 2010. *Diário Oficial do Município do Rio de Janeiro*, Ano XXIV, n. 33, 05 de maio de 2010.

_____. Decreto n. 33.399, de 16 de fevereiro de 2011. *Diário Oficial do Município do Rio de Janeiro*, Ano XXIV, n. 224, 17 de fevereiro de 2011.

_____. Decreto n. 36.673, de 1º de janeiro de 2013. *Diário Oficial do Município do Rio de Janeiro*, Ano XXVI, n. 192, 1º de janeiro de 2013a.

_____. Resolução SME n. 1.234, de 10 de abril de 2013. *Diário Oficial do Município do Rio de Janeiro*, Ano XXVII, n. 18, 11 de abril de 2013b.

SAMMONS, P. As características-chave das escolas eficazes In: BROOKE, N.; SOARES, J. F. *Pesquisa em eficácia escolar:* origem e trajetórias. Belo Horizonte: Editora UFMG, 2008.

SETUBAL, M. A. Meritocracia nos sistemas de educação. *O Estado de S. Paulo*, São Paulo, 2010. Disponível em: <http://www.estadao.com.br/noticias/impresso,meritocracia-nos-sistemas-de educacao,568390,0.htm>. Acesso em: 12 nov. 2013.

TIMMERMANS, A. C.; DOOLAARD, S.; WOLF, I. Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement*, v. 22, n. 4, p. 393-413, 2011.

WEST, M. R.; PETERSON, P. E. The efficacy of choice threats within school accountability systems: results from legislatively induced experiments. *The Economic Journal*, v. 116, n. 510, 2006.

VEGAS, E. (Ed.). *Incentives to improve teaching*: lessons from Latin America. Directions in Development. Washington, D.C.: The World Bank, 2005.

MARIANE CAMPELO KOSLINSKI

Doutora em Sociologia pela Universidade Federal do Rio de Janeiro (UFRJ). Professora Adjunta da Faculdade de Educação e do Programa de Pós-Graduação em Educação da UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil

mckoslinski@gmail.com

CAROLINA PORTELA DA CUNHA
Mestranda do Programa de Pós-graduação em Educação da
Universidade Federal do Rio de Janeiro (UFRJ),
Rio de Janeiro, Rio de Janeiro, Brasil
carolinaportela13@gmail.com

FELIPE MACEDO DE ANDRADE
Mestrando do Programa de Pós-Graduação em Educação da
Universidade Federal do Rio de Janeiro (UFRJ),
Rio de Janeiro, Rio de Janeiro, Brasil
felipema8@gmail.com

Recebido em: AGOSTO 2014
Aprovado para publicação em: OUTUBRO 2014

OUTROS TEMAS

AVALIAÇÕES FINAIS SOBRE O PNE 2001-2010 E PRELIMINARES DO PNE 2014-2024

DONALDO BELLO DE SOUZA

RESUMO

O presente artigo visa a discutir os resultados de algumas avaliações acadêmicas sobre a implantação do Plano Nacional de Educação (PNE) 2001-2010, assinalando aspectos de seu processo de elaboração e aprovação e das análises a respeito do plano sancionado, que apontam para avaliações preliminares do PNE 2014-2024. Tomando por base a revisão da literatura, conclui-se não haver consenso sobre o êxito daquele primeiro PNE, quer entre os estudos de maior abrangência analítica, quer entre esses e as avaliações que enfocam o exame das decisões relativas a determinadas etapas e modalidades da educação básica, essas consensualmente negativas em termos de resultados, impactos ou efeitos observados. Comparativamente ao primeiro plano, ressalta-se, ainda, que a literatura que versa acerca do novo PNE sugere estar mais atenta ao monitoramento do seu desenvolvimento inicial e futuro.

PALAVRAS-CHAVE PLANO NACIONAL DE EDUCAÇÃO • METAS EDUCACIONAIS • AVALIAÇÃO DE POLÍTICAS PÚBLICAS • GESTÃO DEMOCRÁTICA DA EDUCAÇÃO.

RESUMEN

El presente artículo tiene el propósito de discutir los resultados de algunas evaluaciones académicas sobre la implantación del Plano Nacional de Educação *(PNE) 2001-2010, así como de señalar aspectos de su proceso de elaboración y aprobación y de los análisis del plan sancionado, que apuntan hacia evaluaciones preliminares del PNE 2014-2024. Tomando como base la revisión de la literatura, se concluye que no hay un consenso sobre el éxito del primer PNE, sea entre los estudios de mayor amplitud analítica, sea entre éstos y las evaluaciones que enfocan el examen de las decisiones relativas a determinadas etapas y modalidades de la educación básica, éstas consensualmente negativas en lo que se refiere a resultados, impactos o efectos observados. En comparación con el primer plan, se destaca asimismo que la literatura acerca del nuevo PNE sugiere que está más atenta al monitoreo de su desarrollo inicial y futuro.*

PALABRAS CLAVE PLANO NACIONAL DE EDUCAÇÃO • METAS EDUCATIVAS • EVALUACIÓN DE POLÍTICAS PÚBLICAS • GESTIÓN DEMOCRÁTICA DE LA EDUCACIÓN.

ABSTRACT

The present article aims at discussing the results of some academic evaluations of the implementation of the Plano Nacional de Educação *(PNE) for 2001-2010. It highlights aspects of its process of elaboration and approval as well as aspects of the analyses of the endorsed plan that lead to the preliminary evaluations of the 2014-2024 PNE. Based on the revised literature, the conclusion is that there is no consensus on the success of the first PNE, neither among broader analytical studies, nor between these and the evaluations focusing on the examination of the decisions regarding the different stages and modalities of basic education, the latter being consensually negative in terms of results, impacts or effects. Compared to the first plan, it is worth noting that the literature concerning the new PNE suggests a more careful monitoring of its initial and future development.*

KEYWORDS PLANO NACIONAL DE EDUCAÇÃO • EDUCATIONAL GOALS • EVALUATION OF PUBLIC POLICIES • DEMOCRATIC MANAGEMENT OF EDUCATION

INTRODUÇÃO

É possível afirmar que a avaliação *stricto sensu* de planos de educação consiste em um importante componente do processo de planejamento (KIPNIS; ALGARTE, 2001) e, a um só tempo, em mecanismo que viabiliza o acompanhamento, a análise e o julgamento da execução de todas as suas fases. Isso significa considerar a avaliação das políticas educacionais como um processo que implica interatividade e negociações, marcado por certa progressibilidade (BONAMINO et al., 2006), expondo, portanto, dimensões políticas, pois, antes de tudo, resulta de uma construção social e, por conseguinte, da atribuição de valores (CORREIA, 2010; FARIA, 2005; FIGUEIREDO; FIGUEIREDO, 1986).

Nesses termos, a avaliação sistematizada (BONNIOL; VIAL, 2001) de planos de educação passa a ser entendida não apenas como instrumento articulado à correção dos seus rumos pelos governos, útil aos processos de regulação das políticas públicas (BARROSO, 2006), mas como forma política voltada à prestação de contas dos governos à sociedade, com vistas ao conhecimento de seu desempenho, resultados, impactos e efeitos (DRAIBE, 2001).

Todavia, a avaliação dessas políticas vem apresentando grandes desafios para a área educacional, a começar por sua complexidade teórica e, em especial, metodológica, passando pela diversidade de processos participativos que pode demandar e, ainda, em função da amplitude dos seus desdobramentos institucionais e, de uma forma mais geral, sociais (DOURADO, 2010). Além disso, parece que no Brasil não há uma tradição de valorização dos resultados dos processos avaliativos que, por exemplo, retroalimentem o planejamento educacional (BONAMINO et al., 2006), o que minora ainda mais a sua importância institucional e sociopolítica, configurando formal a previsão do acompanhamento e avaliação nos planos de educação nacionais e infranacionais.

No que tange ao Plano Nacional de Educação (PNE) 2001-2010, observa-se que a lei que o regulamentou – Lei n. 10.172, de 09 de janeiro de 2001 –, determinou que a "A União, em articulação com os estados, o Distrito Federal, os municípios e a sociedade civil" (BRASIL, 2001, Art. 3º, *caput*) deveria proceder à efetivação de avaliações periódicas relativas à implementação desse plano nacional, principiando-as no "quarto ano de vigência desta Lei" (BRASIL, 2001, Art. 3º, §2º), vindo a pressupor acompanhamento e avaliação sistemáticos, constituindo-se em subsídio estratégico e indispensável ao monitoramento e regulação do desenvolvimento das políticas públicas associadas a esse tipo de planejamento educacional, assim como em possibilidade de participação sociopolítica e transparência institucional.

No entanto, conforme examinado mais a diante, do ponto de vista acadêmico, as avaliações *stricto sensu* acerca da implementação do PNE 2001-2010 se mostram quantitativamente reduzidas e limitadas a algumas das etapas da educação básica e de suas modalidades de ensino, começando a se tornar mais evidentes no limiar de 2009, às vésperas da realização da I Conferência Nacional de Educação (Conae),[1] pautadas em dados resultantes de algumas das avaliações institucionais promovidas, por exemplo, pelo Ministério da Educação (MEC), embora essas tenham sido iniciadas em meados da vigência daquele plano por meio de importantes trabalhos, como o de Abicalil (2007), Brandão (2006) e Dourado (2006).

1 Conforme será visto mais adiante, a I Conae foi realizada de 28 de março a 1º de abril de 2010, para subsidiar a elaboração de um Projeto de Lei (PL) para o PNE 2011-2020 (BRASIL, 2010), atual PNE 2014-2020 (BRASIL, 2014).

Já no início da década de 2010, começaram a despontar as primeiras avaliações *lato sensu* concernentes ao novo PNE, de início visando ao Projeto de Lei (PL) n. 8.035, de 20 de dezembro de 2010 (BRASIL, 2011), originalmente articulado ao planejamento educacional para o período 2011-2020, para, algum tempo depois, deter-se na sua versão final aprovada por lei, cobrindo o decênio 2014-2024 – Lei n. 13.005, de 25 de junho de 2014 (BRASIL, 2014). Essas avaliações consistem em estudos analíticos preliminares que, embora não se pautem nos resultados, impactos ou efeitos desses instrumentos, mesmo porque, dada a sua incipiência, ainda não se consubstanciam nos sistemas educacionais, visam aos ajuizamentos e valorações do conteúdo político ou técnico do texto do plano propriamente dito, ou, prospectivamente, de suas potenciais repercussões, motivo pelo qual são aqui categorizadas como avaliações *lato sensu*, distinguindo-se, portanto, dos processos avaliativos sistematizados que caracterizam as avaliações *stricto sensu*, primordiais ao desenvolvimento das políticas públicas, em particular no campo do planejamento educacional.

Assim, o presente artigo pretende discutir os resultados de algumas das avaliações *stricto sensu*, acadêmicas – e não institucionais –, do PNE 2001-2010, também assinalando aspectos de seu processo de elaboração, aprovação e análises sobre o plano sancionado, apontando estudos avaliativos *lato sensu* preliminares, também acadêmicos, sobre o PNE 2014-2024 e implicações de seu acompanhamento e avaliação, tomando por base a revisão de uma extensa literatura direta e indiretamente relacionada a esses dois planos, publicada no Brasil no período de 1996 a 2014.[2]

2 Trata-se de um levantamento exaustivo realizado por Souza e Sousa (2012) acerca dos planos de educação no Brasil, inicialmente cobrindo o período de 1996 a 2010 e, mais recentemente, abarcando o quadriênio 2011-2014, em um de levantamento assistemático.

Para tanto, na próxima seção se realiza uma breve discussão conceitual acerca do binômio plano-planejamento, além da caracterização de alguns aspectos históricos relacionados ao planejamento educacional no Brasil. Na terceira seção, enfocam-se as análises de duas propostas elaboradas ao final da década de 1990 para o PNE 2001-2010, de autoria, respectivamente, da sociedade civil organizada e do Poder Executivo federal, para, nas duas seções seguintes, discutir, respectivamente, as reflexões teóricas sobre a versão final

desse plano aprovada por lei e a avaliação acadêmica de sua implantação propriamente dita. Na sequência, estabelecem-se algumas considerações acerca do processo de elaboração e aprovação do novo PNE, a vigorar no decênio 2014-2024, pontuando-se algumas das tendências analíticas que o enfocam, de modo a, na sétima e última seção, à guisa de conclusão, destacarem-se os resultados gerais das análises sobre aquele primeiro PNE e, ainda, os desafios postos ao acompanhamento e avaliação do novo plano.[3]

3 Registram-se os agradecimentos pelas importantes sugestões advindas da leitura atenta deste artigo por parte da colega professora Dra. Neusa Chaves Batista, docente da Faculdade de Educação da Universidade Federal do Rio Grande do Sul (UFRGS).

PLANEJAMENTO E PLANOS DE EDUCAÇÃO NO BRASIL

Diferentes concepções sobre o que é planejamento da educação, em que medida distingue-se dos planos de educação, assim como o lugar que estes ocupam no processo de planejamento propriamente dito, há muito coexistem, quer nas políticas públicas para a área, quer nas discussões estabelecidas no meio acadêmico, o que, em larga medida, condiciona tanto a ação do Estado e dos governos no campo educacional quanto propriamente sua análise científica (por exemplo, histórica, sociológica e política).

Para Cury (1998, p. 164), um plano pode ser considerado

> [...] um programa de realizações para ser cumprido e executado em um certo período (definição cronológica), dentro de objetivos a serem atingidos e para os quais se pleiteiam os meios, inclusive pecuniários, necessários para a implementação adequada.

perspectiva igualmente compartilhada por Bordignon (2014, p. 33), que o entende como:

> [...] o desenho, o projeto do caminho a seguir, das etapas (metas) e ações a empreender (estratégias) na caminhada para alcançar o futuro desejado. É o documento formal que consolida as decisões tomadas no processo de planejamento. O plano se configura, portanto, num registro escrito, apresentado sob a forma de um documento aprovado, na instância própria de competência legal.

Assim, um plano diferenciar-se-ia do planejamento propriamente dito na medida em que, enquanto o primeiro "coordena, racionaliza e dá unidade de fins à atuação do Estado, este garante a coerência entre meios e fins, sendo expressão da política geral do Estado" (MARTINS, 2012, p. 96), motivo pelo qual os planos podem cumprir diversas funções nas instituições políticas, econômicas e sociais, dependendo do seu contexto de elaboração e da correlação de forças sociais que os ensejou (BORDIGNON, 2014; VIEIRA, 2014; VIEIRA; ALBUQUERQUE, 2001). Nos tempos atuais, o planejamento passa a ser entendido, portanto, menos como a produção de um documento (o plano propriamente dito) e mais como um processo (MARTINS, 2012), afigurando-se, em última análise, em um ato de intervenção técnica e política (KUENZER; CALAZANS; GARCIA, 2003).

Em termos históricos, é praticamente consensual que a noção de plano de educação, relacionada à ideia de organização de um sistema de educação nacional, fez-se notar pela primeira vez no Brasil no Manifesto dos Pioneiros da Educação (MANIFESTO..., 2006), em 1932, inaugurando reflexões que se estendem até os dias atuais sobre a formulação e implantação de políticas públicas que se insiram em um quadro de maior racionalidade e sistematização, de modo a superar ações tendenciosamente experimentais, intuitivas, fragmentadas ou desarticuladas (VIEIRA, 2014; MARTINS, 2012; MENDONÇA, 2002; DIDONET, 2000, 2001; AZANHA, 1993, 1998).

Todavia, da década de 1930 até o limiar do século XX, sabe-se que não foram muitos os momentos de elaboração ou implementação, de fato, de planos de educação de abrangência nacional no Brasil, e que tais planos implicaram pouca mobilização sociopolítica. A vertente nacional desenvolvimentista que marcou a década de 1930 influenciaria as políticas públicas de educação até a década de 1970, em meio à ascensão do modelo de industrialização (1937-1955), ao longo de sua crise e substituição por um modelo "associado" de desenvolvimento econômico (1955-1968) (RIBEIRO, 1995), desaguando, até os anos iniciais da chamada Nova República, no tecnocratismo e economicismo que caracterizariam os Planos Setoriais de Educação e Cultura (PSECs),

derivados dos Planos Nacionais de Desenvolvimento (PNDs) (VIEIRA, 2014; VIEIRA; FARIAS, 2007; KUENZER; CALAZANS; GARCIA, 2003; CURY, 1998; HORTA, 1997, 1982), expressão de "uma estratégia de adequação entre o modelo político e o modelo econômico, de base capitalista" (VIEIRA, 2008, p. 117).

A partir do início da década de 1990, de um lado, em meio à reconstrução do federalismo e da democracia (ARRETCHE, 2002a; 2002b), e, de outro, sob os ventos da redefinição do papel do Estado e da própria educação no país (PERONI, 2003; SILVA Jr., 2002; VIEIRA, 2000), o processo de construção de um planejamento educacional sistematizado veio a dar lugar, primeiramente, em 1993, ao Plano Decenal de Educação para Todos (PDET) (BRASIL, 1993) – esquecido antes da metade de sua vigência (OLIVEIRA; ARAÚJO, 2014; CERQUEIRA, 2008; VIEIRA, 2000) –, para, em seguida, engendrar o PNE 2001-2010 (BRASIL, 2001) e, três anos e meio após o término de sua vigência, ao novo PNE (BRASIL, 2014).

ANÁLISES DAS PROPOSTAS PARA O PNE 2001-2010

As determinações legais quanto ao estabelecimento de um PNE encontram-se postas tanto na Constituição Federal (CF) de 1988 – "A lei estabelecerá o plano nacional de educação, de duração plurianual, visando à articulação e ao desenvolvimento do ensino em seus diversos níveis e à integração das ações do Poder Público [...]" (BRASIL, 1988, Art. 214) –, quanto na Lei de Diretrizes e Bases da Educação Nacional (LDBEN) n. 9.394, de 20 de dezembro de 1996 – "A União incumbir-se-á de: I – Elaborar o plano nacional de educação, em colaboração com os estados, o DF [Distrito Federal] e os municípios" (BRASIL, 1996, Art. 9º, I), e, ainda,

> A União, no prazo de um ano a partir da publicação desta Lei, encaminhará, ao Congresso Nacional, o Plano Nacional de Educação, com diretrizes e metas para os dez anos seguintes, em sintonia com a Declaração Mundial sobre educação para Todos (BRASIL, 1996, Art. 87, §1º).

Em paralelo, a Lei n. 9.131, de 24 de novembro de 1995 (BRASIL, 1995, Art. 7º, §1º), incumbiu o Conselho Nacional

de Educação (CNE) de "subsidiar a elaboração e acompanhar a execução do Plano Nacional de Educação".

O cumprimento dessas determinações implicou, ao final da década de 1990, a elaboração de duas propostas de PNE, submetidas ao Congresso Nacional por meio do PL n. 4.155, de 10 de fevereiro de 1998 (BRASIL, 1998a), e do PL n. 4.173, de 11 de fevereiro de 1998 (BRASIL, 1998b), de autoria, respectivamente, da sociedade civil organizada e do Poder Executivo federal, caracterizando-se, de acordo com Beisiegel (1999), no primeiro caso pela ampla participação social e pelo caráter de luta política e ideológica que exprimiu, enquanto que, no segundo caso, pelo comprometimento em relação à racionalidade administrativa das políticas educacionais do MEC, tratando-se, portanto, conforme salientado por Cury (1988, p. 162), de "duas formulações opostas quanto às concepções, diretrizes e metas que as norteiam", embora, na interpretação de Beisiegel (1999, p. 220), ambos os PLs tenham procurado "legitimar-se na ampla participação da sociedade no seu processo de elaboração".

Comparações entre essas duas propostas foram também realizadas, por exemplo, por Calderón e Borges (2014); Davies (2001, 2002, 2014), Souza e Duarte (2014), Oliveira (2005, 2009), Abicalil (2005), Valente e Romano (2002), Didonet (2000) e Beisiegel (1999), em que há consenso de que são propostas substantivamente avessas. Para Didonet (2000, p. 23), essas diferenças se fazem notar desde em aspectos formais, passando pelos estruturais, até aos conteudísticos dos planos:

> [...] como o diagnóstico dos problemas educacionais brasileiros, a análise das políticas educacionais vigentes, a identificação das necessidades sociais em termos de metas, a estimativa de recursos necessários para a educação.

Julgamento compartilhado por Beisiegel (1999, p. 229), que sintetiza as propostas em questão como expressão de três perspectivas básicas:

> [...] uma primeira, que predomina nas propostas do MEC, denominada por Cury (1998, p. 171) realista-conservadora, fortemente condicionada por limitações impostas pela política econômica do governo da União. Uma segunda

posição, igualmente presente no documento do MEC, propõe metas mais ambiciosas do que as compatíveis com as orientações atuais da política econômica da União. E, finalmente, uma terceira posição, presente no plano consolidado no II CONED, que estende consideravelmente o alcance das propostas e dos recursos financeiros para todas as áreas dos serviços educacionais e que, por falta de uma expressão mais adequada, poderia ser designada, provisoriamente, como político-utópica.

Para Calderón e Borges (2014) e Valente e Romano (2002), a dualidade estabelecida por esses PLs não consistia apenas em divergências sobre políticas nacionais de educação ou, de modo específico, sobre projetos de escola: "traduziam dois projetos conflitantes de país". De um lado, de caráter "democrático e popular, expresso na proposta da sociedade", e, de outro, comprometido com "a política do capital financeiro internacional e a ideologia das classes dominantes, devidamente refletido nas diretrizes e metas do governo" (VALENTE; ROMANO, 2002, p. 98).

O PL n. 4.155/1998 (BRASIL, 1998a), cujo primeiro signatário foi o deputado Ivan Valente, do Partido dos Trabalhadores (PT), e que foi subscrito por mais de 70 parlamentares e por todos os líderes dos partidos de oposição na Câmara dos Deputados e apresentado ao Congresso em 10 de fevereiro de 1998, enquanto que o PL n. 4.173/1998 (BRASIL, 1998b) o foi no dia seguinte, como anexo àquele primeiro plano.

ANÁLISES SOBRE O PNE 2001-2010 APROVADO

Para Didonet (2000), o PNE 2001-2010 entrou para a história da educação no Brasil distinguindo-se dos demais planos por seis qualificações básicas:

a) é o primeiro plano submetido à aprovação do Congresso Nacional, portanto, tem força de lei; b) cumpre um mandato constitucional (art. 2014 da Constituição Federal de 1988) e uma determinação da Lei de Diretrizes e Bases da Educação Nacional (LDB, art. 87, 1º); c) fixa diretrizes, objetivos e metas para um período de dez

> anos, o que garante continuidade da política educacional e coerência nas prioridades durante uma década; d) contempla todos os níveis e modalidades de educação e os âmbitos da produção de aprendizagens, da gestão e financiamento e da avaliação; e) envolve o Poder Legislativo no acompanhamento de sua execução; e f) chama a sociedade para acompanhar e controlar a sua execução. (DIDONET, 2000, p. 11)

Previsto pela Lei n. 10.172/2001 (BRASIL, 2001), o PNE definiu diagnósticos, diretrizes e 295 objetivos e metas relacionadas ao conjunto de níveis – educação básica (educação infantil, ensino fundamental e ensino médio) e educação superior – e modalidades de ensino – Educação de Jovens e Adultos, Educação a Distância e Tecnologias Educacionais, Educação Tecnológica e Formação Profissional, Educação Especial e Educação Indígena –, tratando também da Formação dos Professores e Valorização do Magistério, do Financiamento e Gestão, além do Acompanhamento e Avaliação do Plano, propugnando a efetivação de uma gestão descentralizada pela via do regime de colaboração entre os entes federativos.

No cômputo geral, considerando-se o trâmite inicial no Poder Legislativo das duas propostas de plano – PL n. 4.155/1998 (BRASIL, 1998a) e PL n. 4.173/1998 (BRASIL, 1998b) –, Souza e Duarte (2014) observam que a proposta do Governo Federal foi preponderante na versão final aprovada do PNE em causa, fazendo-se notar, de acordo com Bordignon (2014), entre outras dimensões, tanto na estrutura quanto em 75% dos seus objetivos e metas do plano, contra 25% da proposta da Conae. Tal hegemonia, independentemente das particularidades do contexto político da **época, é** explicado por Oliveira (2009, 2008, 2005) por meio da tese de que "o 'poder legislativo' para definição das políticas educacionais pela introdução de legislação, via processo legislativo, é preponderantemente do Executivo" (OLIVEIRA, 2009, p. 2). Com isso,

> O fenômeno ocorre não só porque o Poder Executivo usufrui de mecanismos no processo legislativo que lhe facultam controlar as principais instâncias decisórias do Legislativo, mas também porque o próprio Legislativo subordina suas propostas à aquiescência do Executivo, seja dependendo desse apoio para a aprovação dos seus projetos, seja referendando todos os vetos interpostos pelo Executivo, em muitos casos sem sequer apreciação pelo Congresso Nacional [...]. (OLIVEIRA, 2009, p. 2)

A par da predominância dos interesses do Poder Executivo na esfera do Legislativo, cabe ressaltar que, na época da sanção presidencial, esse PNE recebeu nove vetos que, no seu conjunto, implicaram importantes restrições à gestão e ao financiamento da educação no país, em especial no âmbito do ensino superior (DOURADO, 2010; OLIVEIRA Jr.; BERALDO, 2003; VALENTE; ROMANO, 2002; VALENTE, 2001). Em consequência, o PNE 2001-2010 era muito inconsistente, pois, embora previsse metas de expansão de todos os níveis e modalidades de ensino no país, não presumia custos e tampouco fontes de recursos adicionais para financiá-las (DAVIES, 2014).

Outro ponto a ser destacado é excesso de objetivos e metas, que impõem sérias dificuldades à sua implantação, seu acompanhamento e sua avaliação. Para Dourado (2010, p. 684), esse elevado conjunto de diretrizes e metas

> [...] retratam a carência de organicidade interna do Plano, na medida em que várias metas são reiteradas, por vezes superpostas, e, em outros casos, as metas não apresentam a devida articulação interna, especialmente no que se refere a concepções, financiamento e gestão.

Análises das metas de cada um desses níveis, etapas e modalidades, que anteveem as dificuldades que o plano em questão enfrentaria para o cumprimento dos objetivos e metas estabelecidos, podem ser encontradas, entre outros, nos trabalhos de Mendonça (2002), Davies (2002, 2001), Valente (2001) e, com idêntico vigor, em estudos de maior especificidade voltados, por exemplo, à educação infantil (BARRETO, 2003;

DIDONET, 2003; ARAÚJO, 2002), ao ensino fundamental (TEIXEIRA, 2002b, 2000), ao ensino médio (PINTO, 2004; TEIXEIRA, 2002a) e ao ensino superior (OLIVEIRA Jr.; BERALDO, 2003). No cômputo geral, esses estudos nitidamente convergem para a problemática do financiamento dos objetivos e metas do PNE 2001-2010 (BRASIL, 2001), sinalizando sua insuficiência e sua decorrente, inexequibilidade do plano. Nas palavras de Davies (2002, p. 117), o plano refletiria "muito discurso, nenhum recurso".

No caso da educação infantil, conforme sinalizado por Didonet (2003), as diretrizes e os objetivos e metas correlatos evidenciam a intenção de o PNE torná-la efetiva como primeira etapa da educação básica, porém, de acordo com Barreto (2003), isso somente seria possível se o seu atendimento contasse, de fato, com ações supletivas da União e dos estados, especialmente em municípios que apresentassem necessidades técnicas e financeiras, o que ocorria com acompanhamento da aplicação dos recursos por essas localidades e pelo CNE, conforme indicado por Araújo (2002). No âmbito do ensino fundamental, Teixeira (2002b), igualmente, destaca a importância central do financiamento para que os objetivos e metas do plano fossem alcançados, prevendo que os vetos presidenciais a dispositivos que implicavam ampliação de recursos e investimentos em educação afetariam o atendimento dessa etapa da educação básica. Esses mesmos postulados se repetem para o ensino médio e, sobretudo, para a educação superior, ou seja, no primeiro caso, Pinto (2004) e Teixeira (2002a) enfatizam que para o alcance da quantidade e qualidade de atendimento ao ensino médio definido pelo PNE, em paralelo à continuidade de expansão de sua oferta, em particular pelos estados, implicaria considerar recursos adicionais ao seu financiamento, enquanto que para a educação superior, conforme anteriormente mencionado, o nível mais afetado pelos vetos do Poder Executivo, Oliveira Jr; Beraldo (2003) sinalizam a baixa prioridade do seu atendimento neste PNE, reflexo da pouca importância atribuída ao papel da universidade, especialmente em relação à educação básica.

AVALIAÇÕES SOBRE A IMPLANTAÇÃO DO PNE 2001-2010

Avaliações sistematizadas sobre o PNE 2001-2010 vieram a ser promovidas inicialmente pelos Poderes Executivo e Legislativo, com vistas ao exame do cumprimento das suas decisões (objetivos e metas). Na esfera do Poder Executivo, o CNE disponibilizou um importante balanço desse plano sob o título "Indicações para subsidiar a construção do Plano Nacional de Educação 2011-2020", tornando-o público por meio da Portaria CNE/CP n. 10, de 6 de agosto de 2009 (BRASIL, 2009), enquanto que o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira do Ministério da Educação (MEC/Inep) difundiu, nesse mesmo ano, a série de três volumes intitulados "Avaliação do Plano Nacional de Educação 2001-2008" (BRASIL, 2009a, 2009b, 2009c). No âmbito do Poder Legislativo, esse PNE foi inicialmente avaliado, no ano de 2004, pela Comissão de Educação e Cultura da Câmara dos Deputados, dando origem ao documento "Avaliação técnica do Plano Nacional de Educação" (BRASIL, 2004), posteriormente, abarcando o período 2004-2006, sendo que os resultados encontram-se registrados no documento "Avaliação do PNE 2004-2006" (BRASIL, 2011).

É possível afirmar que as avaliações acadêmicas relacionadas à implementação do PNE 2001-2010 são quantitativamente reduzidas, mas, ainda sim, é possível agrupá-las em estudos de maior abrangência (AGUIAR, 2010; CURY, 2009; DOURADO, 2010; MARTINS, 2012) ou de maior especificidade, pois esses últimos objetivam exclusivamente determinadas etapas da educação básica – como a educação infantil (GANZELI, 2012b; VIEIRA, 2010) e o ensino médio (BRANDÃO, 2011; KUENZER, 2010) – ou determinadas modalidades de ensino – como a Educação de Jovens e Adultos (EJA) (DI PIERRO, 2010) –, trabalhos que, na sua quase totalidade, são subsidiados pelos resultados da avaliação levada a efeito pelo MEC/Inep 2001-2008 anteriormente aludida (BRASIL, 2009a, 2009b, 2009c).

Conforme já mencionado, os vetos presidenciais, à época da sanção do PNE 2001-2010, notadamente quanto ao seu financiamento, impuseram sérias restrições ao cumprimento das metas e objetivos que, ao lado de outros problemas

internos e externos ao documento do plano (BRASIL, 2009), acabaram por comprometer a sua exequibilidade, levando-o a transformar-se, conforme predicado por Kuenzer (2010), Cury (2009) e Sguissardi (2006), numa espécie de carta de intenções, coerentemente ao que Valente e Romano (2002), Azevedo (2001) e Monlevade (2001) haviam previsto e originariamente cognominado. Assim, na percepção de Cury, tais vetos configuraram o PNE em questão como um "Plano-anão: grande cabeça e pés pequenos" (2009, p. 26), enquanto que, para Kuenzer (2010), a situação se agravou ainda mais pelo fato de o plano não explicitar os mecanismos de financiamento que, de fato, lhe dariam suporte. Para essa autora, a não clareza quanto a esses parâmetros, em paralelo à ausência de indicadores associados às metas e objetivos, limitam, sobremaneira, a possibilidade de análise e, por conseguinte, de avaliação desse plano, já anteriormente afetada pelo fato de seu diagnóstico inicial não ter contemplado uma criteriosa análise da situação real da educação brasileira, discutindo-a com a sociedade civil, não apenas apresentando dados brutos, mas séries históricas, matrículas por idade, distorção idade-série, relação entre escolaridade, formação profissional e emprego, entre outros, de modo a mostrar como de fato evoluíram o acesso e a permanência nos diferentes níveis, etapas e modalidades de ensino e suas relações com a ocupação e com o acesso à cultura (KUENZER, 2010, p. 853).

Em uma perspectiva relativamente oposta, Martins (2012) registra o seu desacordo a respeito de apreciações mais negativas do êxito desse plano, em especial de que se tenha tornado uma carta de intenções, sinalizando que, enquanto lei, o PNE não pode ser considerado como tal:

> Uma lei, ainda que falha, está mais próxima de atingir a eficácia que uma mera carta de intenções, porque a lei, ao contrário desta, já rompeu a barreira da validade, já está em vigor. Seu cumprimento pode ser reivindicado judicialmente (MARTINS, 2012, p. 102).

Para esse autor, considerar o PNE 2001-2010 um fracasso significa "supervalorizar o caráter literal do 'plano como teoria', em

detrimento da dimensão processual do planejamento. Revela, assim, uma visão estática, não processual do planejamento" (MARTINS, 2012, p. 109). No seu exame, além de parte das metas do PNE terem sido alcançadas ou encaminhadas, foi por meio dele que se retomou o protagonismo da sociedade civil na formulação das políticas públicas de educação, implicando, portanto, aprendizagens sociopolíticas – nesse sentido, perspectiva igualmente compartilhada por Cury (2009) –, motivos que levam Martins (2012, p. 102) a advogar que esse PNE não se constituiu em fracasso, e tampouco em sucesso:

> Como o copo meio preenchido, que está parcialmente cheio e parcialmente vazio, o PNE falhou em alguns aspectos e foi bem sucedido em outros. E deixou seu legado, para ser ou não aproveitado pelo próximo plano.

Nessa mesma perspectiva, Aguiar (2010, p. 724) identifica avanços em relação às metas e objetivos desse plano, assim como equívocos, ressaltando que:

> [...] a ausência de cumprimento das metas não pode ser atribuída apenas à instância da União. Esta tem responsabilidades concretas, mas os estados, o Distrito Federal e os municípios são corresponsáveis pelos compromissos do Plano. Dependendo da forma como se efetivam as relações entre os entes federados, dos arranjos institucionais e das condições políticas existentes, as metas poderão ser ou não alcançadas.

Dourado (2010), embora reconhecendo o descompasso entre o que foi proposto pelo plano e o efetivado, converge para o mesmo entendimento de Aguiar (2010), esclarecendo que as avaliações que vêm sendo realizadas sobre o PNE 2001-2010 se pautam, predominantemente, no exame das ações do Governo Federal em face das metas e objetivos previstos, não considerando, portanto, as políticas, programas e demais ações relativas aos demais entes federados, completando que esse plano

> [...] não se constituiu como base e diretriz para políticas, planejamento e gestão da educação nacional nem foi acionado como tal pelos diferentes segmentos da sociedade civil e política brasileira (DOURADO, 2010, p. 685)

Não obstante o fato de se ter constituído "em um instrumento de luta em prol da construção das políticas" (DOURADO, 2010, p. 689). Para esse autor, tal limitação se apresenta como resultado dos problemas que marcaram o processo de tramitação do plano,[4]

> [...] da ausência de centralidade conferida a ele [ao PNE] na efetivação de políticas, dos limites político-econômicos vigentes à regulamentação da ação dos entes federados sob a ótica da cooperação e colaboração, como previsto na Constituição Federal de 1988 [...] (DOURADO, 2010, p. 693).

4 Sobre a tramitação do PNE 2001-2010 no Poder Legislativo ver, por exemplo, Peroni e Flores (2014), Oliveira (2009, 2005), Dourado (2010), Valente e Romano (2002), Didonet (2000) e Valente (2001), enquanto que no Poder Executivo, além desses mesmos estudos, os de Oliveira Jr. e Beraldo (2003) e de Mendonça (2002).

Já os estudos acadêmicos que visam a avaliações específicas sobre o PNE 2001-2010, no sentido de privilegiarem determinadas etapas e modalidades da educação básica, revelam o não cumprimento da maior parte das decisões associadas à educação infantil, ao ensino médio e à EJA.[5]

5 Vale enfatizar que o levantamento documental que resultou na identificação dos estudos aqui mencionados, indicado na seção introdutória deste artigo, não revelou avaliações acadêmicas sobre o PNE 2001-2010 que enfocassem o ensino fundamental, a educação superior e modalidades de ensino, a não ser referente à EJA (SOUZA; SOUSA, 2012).

Quanto à educação infantil, o estudo de Ganzeli (2012b), que contempla as metas vinculadas à sua oferta e gestão, destaca, entre outros aspectos, preocupações com a sua expansão e melhoria no padrão de atendimento, cujos problemas se fazem notar, por exemplo, na redução do atendimento de crianças em tempo integral nas creches (de 61,9% em 2000 para 60,3% em 2007) e na não elevação da taxa de escolarização bruta municipal em creche (em 2007, apenas 2,6% dos municípios apresentavam taxa superior a 50%) e em pré-escola (em 2007, apenas 16,4% dos municípios apresentavam taxa superior a 80%) para os 100% previstos no plano. Para Viera (2010), que além de ter tomado por base as metas relativas ao acesso à educação infantil e também considerou aquelas relativas à formação e valorização dos profissionais que atuam nessa etapa da educação básica e aos convênios com creches e pré-escolas, o crescimento do atendimento de crianças em creches no período 2001-2008, da ordem de 8,6%, não foi suficiente para alcançar os 50% de matrículas de crianças de 0 a 3 anos, previstos pelo PNE para 2010, alcançando apenas 18,1%, havendo ainda graves desigualdades de acesso associadas à renda, cor e ao meio rural. A um só tempo, Viera (2010) constata a predominância de professoras com Curso Normal atuando na educação infantil e uma

expressiva participação do setor privado nas suas matrículas que, em 2009, se situava em 52,1% nas creches conveniadas e em 28% na pré-escola, também conveniada.

No que remete ao ensino médio, a avaliação efetuada por Brandão (2011), pautada na análise de sua oferta e seu atendimento, além de também ponderar questões pedagógicas específicas, concluiu que o seu retrato "é pouco animador" (BRANDÃO, 2011, p. 203), uma vez que, em 2009, o número total de matrículas, ao invés de aumentar, decresceu, indicando que não mais da metade dos jovens entre 15 e 17 anos não estão cursando o ensino médio, efeito também associado às baixas taxas de conclusão e qualidade do ensino fundamental. Para Kuenzer (2010), que também analisa dados referentes ao acesso e à qualidade nessa última etapa da educação básica, "os problemas continuam os mesmos" e "a década foi perdida para o ensino médio" (KUENZER, 2010, p. 861), pois, a partir de 2007, passou a ocorrer crescimento negativo de matrículas, implicando, no período 2000-2008, redução de 8,4%.

Na avaliação realizada por Di Pierro (2010), que examinou os objetivos e metas do PNE 2001-2010 para a EJA, as taxas de analfabetismo no período 2001-2008 não apresentaram recuo satisfatório, ao lado de uma elevação também modesta da média de anos de estudos de escolaridade obrigatória, o que, para a autora, também comprometerá o alcance, em 2015, do "compromisso da iniciativa de Educação para Todos de reduzir pela metade o índice de analfabetismo de 13,6% registrado em 2000" (DI PIERRO, 2010, p. 946).

AVALIAÇÕES PRELIMINARES SOBRE O PNE 2014-2024

Ao término da vigência do PNE 2001-2010 (BRASIL, 2001), contabilizou-se a existência, em 2014, de 14 estados, além do Distrito Federal, sem Plano Estadual de Educação (PEE) (54% do total de 26) (SOUZA; MENEZES, 2014; SOUZA; DUARTE, 2014), e, em 2011, cerca de 2.181 municípios (39,2% de 5.565) sem Plano Municipal de Educação (PME) (BRASIL, 2012). Somou-se a esses déficits a ausência de um novo PNE que, originalmente, conforme previsto no PL n. 8.035/2010 (BRASIL, 2011), de autoria do Poder Executivo federal, deveria vigorar

no decênio 2011-2020, mas cujo trâmite legislativo consumiu três anos e meio desse período, tendo sido aprovado somente em 25 de junho de 2014, por meio da Lei n. 13.005 (BRASIL, 2014).

De forma similar aos debates que resultaram na proposta da Sociedade Civil para o PNE 2001-2010 (FÓRUM NACIONAL EM DEFESA DA ESCOLA PÚBLICA, 1997), as discussões sobre a elaboração de um novo PNE ocorreram, entre outros espaços, em 2010, na I Conae – cujos debates, além de abarcarem a educação básica e superior, também gravitaram em torno da criação de um Sistema Nacional de Educação –, evento resultante da deliberação da Conferência Nacional de Educação Básica (CONEB), ocorrida em 2008, ambas precedidas, de forma descentralizada, por conferências subnacionais (BORDIGNON, 2014; OLIVEIRA; ARAÚJO; 2014; PERONI; FLORES, 2014). Assim, a expectativa era a de que o Documento Final da Conae (BRASIL, 2010),[6] que incorporava diversas propostas da sociedade civil para a estruturação do próximo plano – notadamente orientadas por concepções de planejamento com vistas à articulação da gestão democrática das instituições educacionais, da redução da desigualdade e do desenvolvimento integrado à inclusão social (DUARTE; SANTOS, 2014) –, viesse a ser convertido em PL, não vindo a concretizar-se por conta da apresentação, pelo Governo Federal, do PL n. 8.035/2010 (BRASIL, 2011; OLIVEIRA; ARAÚJO, 2014).

6 Uma análise crítica pormenorizada a respeito do Documento Final da Conae (BRASIL, 2010) pode ser encontrada em Davies (2014, 2010), no qual esse autor, ao analisá-lo, identifica diversas fragilidades.

O PL n. 8.035/2010 (BRASIL, 2011) veio a ser encaminhado pelo Poder Executivo ao Congresso Nacional no dia 20 de dezembro de 2010, tendo sido votado pela Comissão Especial da Câmara dos Deputados (CEC), à qual estava subordinado, somente um ano e meio depois, em 13 de junho de 2012, acumulando, ao longo do seu trâmite, cerca de 2.905 emendas parlamentares, fruto da mobilização da sociedade civil (OLIVEIRA; ARAÚJO, 2014; PERONI; FLORES, 2014). Conforme sucedido anteriormente com o PNE 2001-2010, as questões relacionadas ao financiamento da educação pública constantes desse novo PNE parecem ter sido, mais uma vez, as mais polêmicas, engendrando algumas importantes disputas parlamentares. De acordo com Oliveira e Araújo (2014), as discussões ocorridas na Câmara dos Deputados, es-

pecialmente as realizadas por meio de audiências públicas, pautaram principalmente a problemática do financiamento da educação, delineando duas tendências, uma em defesa da aplicação de 10% do Produto Interno Bruto (PIB) em educação – coerente ao que previa o PL n. 8.035/2010 –, enquanto que a outra defendia a destinação apenas de 7%. Para Peroni e Flores (2014), tal centralidade decorreria justamente da experiência social pregressa relativa ao PNE 2001-2010, que, conforme visto, veio a ter a sua previsão de recursos vetada, o que determinou a consolidação de uma frente de luta em torno de uma clara destinação de recursos para o financiamento do plano, especialmente pautada nos 10% do PIB.

Logo após a aprovação do PL n. 8.035/2010 (BRASIL, 2011) pela Comissão de Constituição e Justiça da Câmara dos Deputados, ocorrida em 16 de outubro de 2012 – incluindo a proposta de destinação de 10% do PIB –, houve o seu encaminhamento ao Senado Federal pela Mesa Diretora da Câmara dos Deputados, no dia 25 desse mesmo mês, onde, na forma de Projeto de Lei da Câmara (PLC) n. 103 (BRASIL, 2012), passou a aguardar a sua avaliação,[7] cuja votação veio a ser realizada apenas no dia 17 de dezembro de 2013 (PERONI; FLORES, 2014).

Retornando à Câmara dos Deputados, o PL foi finalmente votado no dia 3 de junho de 2014, e, alguns dias depois, em 25 de junho do mesmo ano, sancionado sem vetos pela Presidência da República. No texto sancionado pelo Governo Federal, o novo PNE determina, na vigésima meta, a ampliação do "investimento público em educação de forma a atingir, no mínimo, o patamar de 7% do Produto Interno Bruto (PIB) do País no quinto ano de vigência desta Lei e, no mínimo, o equivalente a 10% do PIB no final do decênio" (BRASIL, 2014).

Embora a aprovação desse novo PNE seja recente, alguns estudos já principiaram discussões a seu respeito, sinalizando aspectos da proposta final da Conae (BRASIL, 2010) que deixaram de ser incorporados a esse plano, ou enfocando questões polêmicas atinentes aos seus princípios ou objetivos e metas como, por exemplo, a criação de um Sistema Nacional de Educação, as vinculações do plano com a gestão democrática da educação e, ainda, o financiamento da educação (DAVIES, 2014, 2010; PERONI; FLORES, 2014).

7 Informações circunstanciadas relacionadas a configurações assumidas pelo PLC n. 103/2012 (BRASIL, 2012) ao longo do seu trâmite no Senado, assim como a respeito desse próprio processo legislativo, antes da sua apreciação final pela Comissão de Educação, Cultura e Esporte e pelo Plenário, encontram-se em Silva *et al.* (2013), enquanto que o apontamento das fragilidades desse PLC são formuladas por Davies (2014).

Outros trabalhos, elaborados ao longo do trâmite do PL n. 8.035/2010 (BRASIL, 2011) no Poder Legislativo, alguns ainda referendando as diretrizes do Documento Final da Conae (BRASIL, 2010), trazem à vista análises relacionadas a questões-chave que o novo PNE, para dar conta das demandas educacionais do país, deveria ter contemplado, configurando-se em importantes referências para as análises atuais sobre os direcionamentos impressos nesse novo plano, por exemplo, para a educação infantil (GANZELI, 2012b), a EJA (DI PIERRO, 2010) e o ensino médio (BRANDÃO, 2011; PINTO; AMARAL; CASTRO, 2011; KUENZER, 2010). A par dessas etapas da educação básica, outros estudos vêm estabelecendo ajuizamentos acerca dos problemas que marcam as decisões concernentes à avaliação da educação básica nesse novo plano (OLIVEIRA, 2010), à valorização do magistério (BRITO, 2013), ao financiamento da educação (OLIVEIRA, 2011; AMARAL, 2010) e, por fim, à gestão da educação – privilegiando o planejamento educacional (MARTINS, 2012), a constituição de um Sistema Nacional de Educação – no qual o PNE teria papel articulador estratégico (FERNANDES; BRITO; PERONI, 2012; SAVIANI, 2010a, 2010b; ARAUJO, 2010) –, o regime de colaboração (GANZELI, 2012a; ARAUJO, 2010) e o acompanhamento e controle do plano propriamente dito (OLIVEIRA, 2011).[8]

8 Cabe destacar a coletânea organizada por Dourado (2011), que apresenta dez importantes estudos sobre os encaminhamentos que vinham sendo dados a diversos temas constantes do PL n. 8.035/2010.

Nesses termos, em que pese o fato de o novo PNE assentar-se apenas em vinte decisões – distintamente do primeiro plano, que veio a ser constituído por 295 –, no cômputo geral continuam a vigorar os desafios relativos à universalização do conjunto de etapas e modalidades da educação básica, articuladamente à melhoria da sua qualidade; à ampliação, igualmente, da oferta com qualidade do ensino superior, incluindo-se o segmento de pós-graduação; passando pela qualificação e valorização dos profissionais da educação e culminando tanto no asseguramento de condições para a efetivação da gestão democrática quanto na ampliação do investimento público em educação, agora para os almejados 10% do PIB.

CONSIDERAÇÕES FINAIS

Circunscritamente ao universo de referências até aqui inventariadas, é possível afirmar que as análises acadêmicas sobre o PNE 2001-2010 tenderam a concentrar-se mais sobre a sua elaboração e aprovação do que propriamente a respeito da avaliação da sua implantação, expondo consensualidade quanto ao fato de que seu trâmite no Poder Legislativo veio a ocorrer em meio à hegemonia dos interesses do Poder Executivo, interesses esses igualmente manifestos quando na sanção presidencial, consubstanciados por meio de vetos incidentes no seu financiamento.

A um só tempo, os estudos em questão permitem depreender que, mesmo antes da sua implantação, diversas avaliações *lato sensu* pautadas no texto do plano propriamente dito anteviram dificuldades que o PNE 2001-2010 enfrentaria para o cumprimento das decisões fixadas, convergindo para o entendimento de que os recursos para o seu financiamento se mostravam insuficientes, em larga medida agravado pela ausência da explicitação dos mecanismos de financiamento que dariam suporte ao plano.

Já em termos da implantação do PNE, constata-se que as poucas avaliações *stricto sensu* levadas a efeito divergem quanto à sua efetividade. No âmbito dos exames de maior abrangência analítica, o êxito do PNE 2001-2010 não é consensual, havendo posicionamentos, de um lado, que advogam o fracasso do plano, inclusive destacando a impossibilidade técnica de avaliá-lo, e, de outro, que ressaltam o mérito de sua expressão jurídica, significado sociopolítico e valor pedagógico que exprime para o processo social e institucional de construção de novos planos de educação. Na esfera dos estudos que pontualmente enfocam o exame das metas e objetivos de determinadas etapas e modalidades da educação básica, tomando como referência resultados das avaliações institucionais realizadas pelo MEC, sua avaliação se revela consensualmente negativa em termos dos resultados, impactos ou efeitos observados.

Comparativamente ao PNE 2001-2010, a literatura que versa sobre o novo PNE sugere estar mais atenta ao monitoramento do seu desenvolvimento inicial, agora levando em

conta um plano que, em função das grandes fragilidades do seu antecessor, mostra-se tecnicamente mais aperfeiçoado, embora em termos político-educacionais ainda reflita conservadorismos e decisões privatistas da educação.

O acompanhamento e avaliação do PNE 2014-2024, não apenas do ponto de vista da sua versão nacional, mas também ao nível infranacional, consiste em um dos grandes desafios a serem enfrentados na sua implementação. Na esfera nacional, esse plano estabelece que a execução e o "cumprimento de suas metas serão objeto de monitoramento contínuo e de avaliações periódicas" (BRASIL, 2014, Art. 5º, *caput*), a serem institucionalmente realizados pelo MEC, pela Comissão de Educação da Câmara dos Deputados e Comissão de Educação, Cultura e Esporte do Senado Federal, pelo CNE e, ainda, pelo Fórum Nacional de Educação (FNE) (BRASIL, 2014, Art. 5º, I a IV). No âmbito infranacional, dispõe que "Os sistemas de ensino dos Estados, do Distrito Federal e dos Municípios criarão mecanismos para o acompanhamento local da consecução das metas deste PNE e dos planos previstos no art. 8º" (BRASIL, 2014, art. 7º, § 3º), ou seja, esses entes federados, ao elaborarem, no prazo de um ano, os respectivos PEEs, Plano de Educação do Distrito Federal (PEDF) e PMEs, deverão observar o seu acompanhamento e avaliação em "consonância com as diretrizes, metas e estratégias previstas neste PNE" (BRASIL, 2014, Art. 8º).

Por fim, cabe ressaltar que, de modo distinto do primeiro PNE, o novo plano se encontra assentado em objetivos e metas com quantitativos e prazos mensuráveis, trazendo à vista as intenções e ações das políticas educacionais a serem adotadas para a década, contando, ainda, com a indicação de algumas das etapas a serem cumpridas ao longo da sua implementação, o que, em larga medida, viabiliza os processos de acompanhamento e avaliação, quer pelos Poderes Públicos, quer pela sociedade civil, em particular por meio do importante papel a ser cumprido pela agenda das pesquisas em políticas públicas em educação.

REFERÊNCIAS

ABICALIL, Carlos Augusto. *Plano Nacional de Educação na República Federativa do Brasil*: instrumento de retórica ou política pública para a realização do direito à Educação Básica? 2005. 185 f. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade de Brasília, Brasília, Distrito Federal, 2005.

_____. *PNE*: limites e desafios; uma avaliação necessária. Brasília, DF: Câmara dos Deputados, 2007.

AGUIAR, Márcia Ângela da Silva. Avaliação do Plano Nacional de Educação 2001-2009: questões para reflexão. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 707-727, jul./set. 2010.

AMARAL, Nelson Cardoso. Financiamento da educação básica e o PNE 2011-2020. *Retrato de Escola*, Brasília, DF, v. 4, n. 6, p. 123-141, jan./jul. 2010.

ARAÚJO, Gilda Cardoso de. Constituição, federação e propostas para o novo plano nacional de educação: análise das propostas de organização nacional da educação brasileira a partir do regime de colaboração. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 749-768, jul./set. 2010.

ARAÚJO, Regina Magna Bonifácio de. A educação infantil e o Plano Nacional de Educação. In: TEIXEIRA, Lúcia Helena Gonçalves. (Org.). *LDB e PNE*: desdobramentos na política educacional brasileira. São Bernardo do Campo: Unesp, 2002. p. 55-66. (Cadernos Anpae, 1).

ARRETCHE, Martha. Federalismo e relações intergovernamentais no Brasil: a reforma de programas sociais. *Dados: Revista de Ciências Sociais,* Rio de Janeiro, v. 45, n. 3, p. 431-458, 2002a.

_____. Relações federativas nas políticas sociais. *Educação & Sociedade*, Campinas, v. 23, n. 80, p. 25-48, set. 2002b.

AZANHA, José Mário Pires. Política e planos de educação no Brasil: alguns pontos para reflexão. *Cadernos de Pesquisa,* São Paulo, n. 85, p. 70-78, maio, 1993.

_____. Planos e políticas de educação no Brasil: alguns pontos para reflexão. In: BARROS, Roque Spencer Maciel et al. *Estrutura e funcionamento da educação básica*. 2. ed. São Paulo: Pioneira, 1998. p. 102-123.

AZEVEDO, Neroaldo Pontes de. Política e gestão na educação municipal. In: CONFERÊNCIA NACIONAL DE EDUCAÇÃO, CULTURA E DESPORTO, 1., 2001, Brasília, DF. *Anais*... Brasília, DF: Câmara dos Deputados, 2001. p. 71-74.

BARRETO, Ângela Maria Rabelo Ferreira. A educação infantil no contexto das políticas públicas. *Revista Brasileira de Educação*, Rio de Janeiro, v. 8, n. 24, p. 53-65, set./dez. 2003.

BARROSO, João. Introdução: a investigação sobre a regulação das políticas públicas de educação em Portugal. In: _____. (Org.). *O Estado e a educação*: a regulação transnacional, a regulação nacional e a regulação local. Lisboa: Educa, 2006. p. 41-70.

BEISIEGEL, Celso de Rui. O Plano Nacional de Educação. *Cadernos de Pesquisa*, São Paulo, v. 29, n. 106, p. 217-231, mar. 1999.

BONAMINO, Alicia; FRANCO, Creso; SOUSA, Sandra Zákia; MAIA, Maurício. Avaliação de políticas educacionais. In: BRASIL. Ministério da Educação. Secretaria de Educação Básica. *Pradime*: Programa de Apoio aos Dirigentes Municipais de Educa. Brasília, DF: MEC/SEB, 2006. p. 133-176. (Caderno de Textos, 1).

BONNIOL, Jean-Jacques; VIAL, Michel. *Modelos de avaliação*: textos fundamentais. Porto Alegre: Artmed, 2001.

BORDIGNON, Genuíno. Caminhar da educação brasileira: muitos planos, pouco planejamento. In: SOUZA, Donaldo Bello de; MARTINS, Ângela Maria (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 29-54.

BRANDÃO, Carlos da Fonseca. *PNE passo a passo (Lei n. 10.172/2001)*: discussão dos objetivos e metas do Plano Nacional de Educação. São Paulo: Avercamp, 2006.

_____. O ensino médio no contexto do Plano Nacional de Educação: o que ainda precisa ser feito. *Cadernos CEDES*, Campinas, v. 31, n. 84, p. 195-208, maio/ago. 2011.

BRASIL. Constituição [de 1988] da República Federativa do Brasil. *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 5 out. 1988.

BRASIL. Lei n. 9.131, de 24 de novembro de 1995. Altera dispositivos da Lei n. 4.024, de 20 de dezembro de 1961, e dá outras providências. *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 25 nov. 1995.

_____. Lei n. 9.394, de 20 de dezembro de 1996. Estabelece as Diretrizes e Bases da Educação Nacional. *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 23 dez. 1996.

_____. Lei n. 10.172, de 09 de janeiro de 2001. Aprova o Plano Nacional de Educação (PNE). *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 10 jan. 2001.

_____. Lei n. 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 26 jun. 2014.

BRASIL. Projeto de Lei n. 4.155, de 10 de fevereiro de 1988. Aprova o Plano Nacional de Educação [Proposta da Sociedade Brasileira]. *Diário da Câmara dos Deputados*, Brasília, DF, v. 53, n. 42, 12 mar. 1998a.

_____. *Projeto de Lei n. 4.173*, de 11 de fevereiro de 1998. Institui o Plano Nacional de Educação [Proposta do Poder Executivo ao Congresso Nacional apensada ao PL 4.155/1998]. Brasília, DF: Câmara dos Deputados, 13 fev. 1998b.

_____. *Projeto de Lei do Plano Nacional de Educação (PNE 2011-2020)*: Projeto em tramitação no Congresso Nacional – PL n. 8.035/2010. Brasília, DF: Câmara dos Deputados, Edições Câmara, 2011.

BRASIL. Câmara dos Deputados. *Projeto de Lei n. 103*, de 25 de outubro de 2012, (Projeto de Lei n. 8.035, de 2010, na Casa de origem), de iniciativa da Presidência da República. Aprova o Plano Nacional de Educação – PNE e dá outras providências. Brasília, DF: Congresso Nacional, 2012.

BRASIL. Câmara dos Deputados. Comissão de Educação e Cultura. *Avaliação técnica do Plano Nacional de Educação*. Brasília, DF: Câmara dos Deputados, 2004. (Série Ação Parlamentar, 294).

_____. *Avaliação do PNE 2004-2006*: avaliação do cumprimento das metas do Plano Nacional de Educação (PNE). Brasília, DF: Centro de Documentação e Informação, Edições Câmara, 2011. (Ação Parlamentar, 351).

BRASIL. Conselho Nacional de Educação. *Portaria CNE/CP n. 10*, de 6 de agosto de 2009. Indicações para subsidiar a construção do Plano Nacional de Educação 2011-2020. Brasília, DF: CNE/CP, 2009.

BRASIL. Instituto Brasileiro de Geografia e Estatística. Diretoria de Pesquisas. Coordenação de População e Indicadores Sociais. *Perfil dos municípios brasileiros 2011*: pesquisa de informações básicas municipais. Rio de Janeiro: IBGE, 2012.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Avaliação do Plano Nacional de Educação 2001-2008*: níveis de ensino. Brasília, DF: MEC/INEP, 2009a. v. 1.

_____. *Avaliação do Plano Nacional de Educação 2001-2008*: modalidades de ensino. Brasília, DF: MEC/INEP, 2009b. v. 2.

_____. *Avaliação do Plano Nacional de Educação 2001-2008*: magistério da educação básica e financiamento e gestão educacional . Brasília, DF: MEC/INEP, 2009c. v. 3.

BRASIL. Ministério da Educação. Secretaria Executiva. *Conferência Nacional de Educação 2010 – Construindo o sistema nacional articulado de educação*: o plano nacional de educação, diretrizes e estratégias de ação. Documento Final. Brasília, DF: MEC, 2010.

BRASIL. Ministério da Educação e do Desporto. *Plano Decenal de Educação Para Todos*. Brasília, DF: MEC, 1993.

BRITO, Vera Lúcia Alves de. O Plano Nacional de Educação e o ingresso dos profissionais do magistério da educação básica. *Educação & Sociedade*, Campinas, v. 34, n. 125, p. 1251-1267, out./dez. 2013.

CALDERÓN, Adolfo Ignácio; BORGES, Regilson Maciel. Construção dos planos nacionais de educação no Brasil: os grupos de articulação de interesses em ação. In: SOUZA, D. B. de; MARTINS, Â. M. (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 99-123.

CERQUEIRA, Nádia Dorian Souza de. *Valorização do magistério:* contradições entre os anúncios de políticas governamentais e as medidas efetivadas para a melhoria do trabalho docente no período de 1993-2003 no Município de Alagoinhas-Bahia. 2008. 76 f. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal da Bahia, Salvador, Bahia.

CORREIA, José Alberto. Paradigmas e cognições no campo da administração educacional: das políticas de avaliação à avaliação como política. *Revista Brasileira de Educação*, Rio de Janeiro, v. 15, n. 45, p. 456-592, set./dez. 2010.

CURY, Carlos Roberto Jamil. O Plano Nacional de Educação: duas formulações. *Cadernos de Pesquisa*, São Paulo, v. 28, n. 104, p. 162-180, jul. 1998.

_____. Por um Plano Nacional de Educação: nacional, federativo, democrático e efetivo. *Revista Brasileira de Política e Administração da Educação*, Porto Alegre, v. 25, n. 1, p. 13-30, 2009.

DAVIES, Nicholas. Plano Nacional de Educação: muito discurso, nenhum recurso. *Universidade e Sociedade*, São Paulo, v. 11, n. 25, p. 29-39, dez. 2001.

_____. Plano Nacional de Educação: muito discurso, nenhum recurso. In: TEIXEIRA, L. H. G. (Org.). *LDB e PNE*: desdobramentos na política educacional brasileira. São Bernardo do Campo: Unesp, 2002. p. 117-143.

_____. A Conferência Nacional de Educação (Conae) e suas fragilidades. *Revista HISTEDBR [online]*, Campinas, v. 10, n. 39, p. 321-345, set. 2010.

_____. Fragilidades e desafios do financiamento em planos de educação: 10% do PIB são a salvação? In: SOUZA, Donaldo Bello de; MARTINS, Ângela. Maria. (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 183-205.

DI PIERRO, Maria Clara. A educação de jovens e adultos no plano nacional de educação: avaliação, desafios e perspectivas. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 939-959, jul./set. 2010.

DIDONET, Vital. *Plano Nacional de Educação (PNE)*. Brasília, DF: Plano, 2000.

_____. *O Plano Nacional de Educação e os planos estaduais e municipais*: uma conquista histórica. *Gestão em Rede*, Curitiba, n. 33, p. 14-19, nov./dez. 2001.

_____. A educação infantil no Plano Nacional de Educação. *Amae Educando*, Belo Horizonte, v. 36, n. 316, p. 6-10, 2003.

DOURADO, Luiz Fernandes. Plano Nacional de Educação: avaliações e retomada do protagonismo da sociedade civil organizada na luta pela educação. In: FERREIRA, N. S. C. (Org.). *Políticas públicas e gestão da educação*: polêmicas, fundamentos e análises. Brasília, DF: Liber Livro, 2006. p. 21-50.

_____. Avaliação do Plano Nacional de Educação 2001-2009: questões estruturais e conjunturais de uma política. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 677-705, jul./set. 2010.

_____. (Org.). *Plano nacional de educação (2011-2020)*: avaliação e perspectivas. Goiânia: Ed. UFG; Belo Horizonte: Autêntica, 2011.

DRAIBE, Sônia Miriam. Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, Maria Cecília Roxo Nobre; CARVALHO, Maria do Carmo Brant de. (Org.). *Tendências e perspectivas na avaliação de políticas e programas sociais*. São Paulo: IEE/PUC-SP, 2001. p. 13-42.

DUARTE, Marisa Ribeiro Teixeira; SANTOS, Maria Rosimary Soares. Educação e desenvolvimento: planejamento de ações governamentais e novos modos de regulação social? In: SOUZA, Donaldo Bello de; MARTINS, Ângela Maria. (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 73-98.

FARIA, Carlos Aurélio Pimenta de. A política da avaliação de políticas públicas. *Revista Brasileira de Ciências Sociais*, São Paulo, v. 20, n. 59, p. 97-109, out. 2005.

FERNANDES, Maria Dilnéia Espíndola; BRITO, Silvia Helena Andrade de; PERONI, Vera Maria Vidal. *Revista Brasileira de Estudos Pedagógicos*, Brasília, DF, v. 93, n. 235, p. 565-578, set./dez. 2012.

FIGUEIREDO, Marcus Faria; FIGUEIREDO, Argelina Maria Cheibub. Avaliação política e avaliação de políticas: um quadro de referência teórica. *Análise & Conjuntura*, Belo Horizonte, v. 1, n. 3, p. 107-127, set./dez. 1986.

FÓRUM NACIONAL EM DEFESA DA ESCOLA PÚBLICA. *Plano Nacional de Educação*: proposta da sociedade brasileira [Consolidado na plenária de encerramento do II CONED]. In: CONGRESSO NACIONAL DE EDUCAÇÃO, 2., 1997. *Anais*... Belo Horizonte: FNDEP, nov. 1997.

GANZELI, Pedro. Plano nacional de educação no Brasil: reflexos sobre o regime de colaboração. In: CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 3., 2012, Zaragoza, Espanha. *Anais*... Recife: ANPAE, 2012a.

_____. Plano nacional de educação: implicações para a educação infantil. *Revista Exitus*, Belém, v. 2, n. 2, p. 77-102, jul./dez. 2012b.

HORTA, José Silvério Baia. Liberalismo, tecnocracia e planejamento educacional no Brasil. São Paulo: Cortez; Campinas: Autores Associados, 1982.

_____. Plano Nacional de Educação: da tecnocracia à participação democrática. In: CURY, Carlos Roberto Jamil. *Medo* à *liberdade e compromisso democrático*: LDB e Plano Nacional de Educação. São Paulo: Pioneira, 1997. p. 137-206.

KIPNIS, Bernardo; ALGARTE, Roberto. Planejamento e avaliação educacionais. In: WITTMANN, Lauro Carlos; GRACINDO, Regina Vinhaes (Org.). *O estado da arte em política e gestão da educação no Brasil 1991-1997*. Campinas: Autores Associados; São Paulo: Anpae, 2001. p. 151-171.

KUENZER, Acácia Zeneida. O ensino médio no Plano Nacional de Educação 2011-2020: superando a década perdida? *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 851-873, jul./set. 2010.

KUENZER, Acácia; CALAZANS, Maria Julieta; GARCIA, Walter. *Planejamento e educação no Brasil*. 6. ed. São Paulo: Cortes, 2003. (Coleção Questões da Nossa Época, 21).

MANIFESTO dos Pioneiros da Educação Nova [1932]. *Revista Histedbr On-line*, Campinas, n. especial, p. 188-204, ago. 2006. Disponível em <http://www.histedbr.fae.unicamp.br/revista/edicoes/22e/doc1_22e.pdf> Acesso em: 06 nov. 2014.

MARTINS, Paulo de Sena. Planejamento e plano nacional de educação. *Cadernos Aslegis*, Brasília, DF, n. 39, p. 91-118, jan./abr. 2012.

MENDONÇA, Erasto Fortes. Plano Nacional de Educação: desdobramentos na política nacional. In: TEIXEIRA, Lúcia Helena G. (Org.). *LDB e PNE*: desdobramentos na política educacional brasileira. São Bernardo do Campo: UMESP, 2002. p. 13-54. (Cadernos Anpae, n. 1).

MONLEVADE, João Antônio Cabral de. Avaliação e perspectivas do financiamento da educação pública no Brasil. In: CONFERÊNCIA NACIONAL DE EDUCAÇÃO, CULTURA E DESPORTO, 1, 2001, Brasília, DF. *Anais*... Brasília, DF: Câmara dos Deputados, 2001. p. 43-48.

OLIVEIRA, Dalila Andrade de; ARAÚJO, Heleno. Educação entre os planos de governo e as políticas de Estado: o foco no financiamento e a questão docente. In: SOUZA, Donaldo Bello de; MARTINS, Ângela Maria (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 167-182.

OLIVEIRA, João Ferreira de. A educação básica e o PNE/2011-2020: políticas de avaliação democrática. *Revista Retratos da Escola*, Brasília, v. 4, n. 6, p. 91-108, jan.-jun. 2010.

OLIVEIRA Jr., Lourival Batista de; BERALDO, Antonio Fernando. Vetar é preciso, educar não é preciso: os vetos presidenciais ao Plano Nacional de Educação. *Revista Brasileira de Política e Administração da Educação*, Piracicaba, v. 19, n. 1, p. 55-80, 2003.

OLIVEIRA, Romualdo Portela de. O Plano Nacional de Educação: algumas questões para debate. *Jornal de Políticas Educacionais*, Curitiba, n. 9, p. 3-10, jan./jun. 2011.

OLIVEIRA, Rosimar de Fátima. *O papel do Poder Legislativo na formulação das políticas educacionais*. 2005. 262 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2005.

_____. O papel do poder legislativo na formulação das políticas sobre financiamento da educação. *Revista Brasileira de Política e Administração da Educação*, Porto Alegre, v. 24, n. 2, p. 235-247, maio/ago. 2008.

_____. Revisando os mecanismos de formulação do Plano Nacional de Educação: considerações sobre o processo decisório. In: REUNIÃO ANUAL DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM EDUCAÇÃO, 32., 2009, Caxambu. *Anais*... Rio de Janeiro: ANPEd, 2009.

PERONI, Vera Maria Vidal. *Política educacional e papel do estado no Brasil dos anos 1990*. São Paulo: Xamã, 2003.

PERONI, Vera Maria Vidal; FLORES, Maria Luiza Rodrigues. Sistema nacional, plano nacional e gestão democrática da educação no Brasil: articulações e tensões. In: SOUZA, D. B. de; MARTINS, Â. M. (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 147-166.

PINTO, José Marcelino de Rezende. Financiamento do ensino médio no Brasil: uma abordagem inicial. *Trabalho, Educação e Saúde*, Rio de Janeiro, v. 2, n. 1, p. 135-152, mar. 2004.

PINTO, José Marcelino de Rezende; AMARAL, Nelson Cardoso; CASTRO, Jorge Abrahão de. O financiamento do ensino médio no Brasil: de uma escola boa para poucos à massificação barata da rede pública. *Educação & Sociedade*, Campinas, v. 32, n. 116, p. 639-665, jul./set. 2011.

RIBEIRO, Maria Luísa Santos. *História da educação brasileira*: a organização escolar. 14. ed. Campinas: Autores Associados, 1995.

SAVIANI, Dermeval. Organização da educação nacional: Sistema e Conselho Nacional de Educação, Plano e Fórum Nacional de Educação. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 769-787, jul./set. 2010a.

_____. Sistema nacional de educação articulado ao Plano Nacional de Educação. *Revista Brasileira de Educação*, Rio de Janeiro, v. 15, n. 44, p. 380-412, maio/ago. 2010b.

SGUISSARDI, Valdemar. Prefácio. In: BRANDÃO, Carlos da Fonseca. *PNE passo a passo (Lei n. 10.172/2001)*: discussão dos objetivos e metas do Plano Nacional de Educação. São Paulo: Avercamp, 2006. p. 13-16.

SILVA, Fernando Mariano da; MONLEVADE, João Antonio Cabral de; QUEIROZ, José Edmar de; BRITTO Tatiana Feitosa de. *O Plano Nacional de Educação no Senado*: estágio atual. Brasília, DF: Senado Federal, nov. 2013. (Textos para Discussão, 143).

SILVA Jr., João dos Reis. *Reforma do estado e da educação no Brasil de FHC*. São Paulo: Xamã, 2002.

SOUZA, Donaldo Bello de; DUARTE, Marisa Ribeiro Teixeira. Planos de educação no Brasil: projeções do sistema nacional de educação e suas variantes subnacionais. In: SOUZA, Donaldo Bello de; MARTINS, Ângela Maria (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 209-232.

SOUZA, Donaldo Bello de; MENEZES, Janaína Specht da Silva. *Estudo crítico-exploratório sobre os Planos Estaduais de Educação (PEEs) no Brasil*: contribuições para formulações e reformulações decorrentes do novo Plano Nacional de Educação (PNE) – Relatório de pesquisa. Rio de Janeiro: Nephem/Uerj; Neephi/Unirio, jun. 2014.

SOUZA, Donaldo Bello de; SOUSA, Gustavo José Albino de. Planos nacionais, estaduais e municipais de educação no Brasil: balanço quali-quantitativo sobre a literatura acadêmica (1996-2010). *Agenda Social*, Campos dos Goytacases, v. 6, p. 50-70, 2012.

TEIXEIRA, Beatriz de Bastos. Plano nacional de educação impõe PCNs às escolas. In: CAVALCANTE, A. L.; PEREIRA, J. R.; LIMA, M. J. R.. *Plano Nacional de Educação*: algumas considerações. Brasília, DF: Núcleo de educação, Cultura, Desporto, Ciência & Tecnologia da Bancada do Partido dos Trabalhadores na Câmara dos Deputados, 2000.

_____. Ensino médio: perspectivas e limites contidos no PNE. In: TEIXEIRA, L. H. G. (Org.). *LDB e PNE*: desdobramentos na política educacional brasileira. São Bernardo do Campo: Unesp, 2002a, p. 83-91.(Cadernos Anpae, n. 1).

_____. O ensino fundamental no Plano Nacional de Educação. In: TEIXEIRA, L. H. G. (Org.). *LDB e PNE*: desdobramentos na política educacional brasileira. São Bernardo do Campo: Unesp, 2002b. p. 67-82. (Cadernos Anpae, n. 1).

VALENTE, Ivan. *Plano Nacional de Educação*. Rio de Janeiro: DP&A, 2001.

VALENTE, Ivan; ROMANO, Roberto. PNE: Plano Nacional de Educação ou carta de intenções? *Educação & Sociedade*, Campinas, v. 23, n. 80, p. 96-107, set. 2002.

VIEIRA, Lívia Maria Fraga. A educação infantil e o Plano Nacional de Educação: as propostas da CONAE 2010. *Educação & Sociedade*, Campinas, v. 31, n. 112, p. 809-8031, jul./set. 2010.

VIEIRA, Sofia Lerche. *Política educacional em tempos de transição (1985-1995)*. Brasília: Plano, 2000.

_____. *Desejo de reforma*: legislação educacional no Brasil Império e República. Brasília, DF: Liber Livro, 2008.

_____. Planos e políticas educacionais: das concepções às práticas no Brasil. In: SOUZA, Donaldo Bello de; MARTINS, Ângela Maria. (Org.). *Planos de educação no Brasil*: planejamentos, políticas, práticas. São Paulo: Loyola, 2014. p. 55-72.

VIEIRA, Sofia Lerche; ALBUQUERQUE, Maria Glaucia Menezes. *Política e planejamento educacional*. Fortaleza: Edições Demócrito Rocha, 2001.

VIEIRA, Sofia Lerche; FARIAS, Isabel Maria Sabino de. *Política educacional no Brasil*: introdução histórica. Brasília, DF: Líber Livro, 2007.

DONALDO BELLO DE SOUZA

Professor Associado do Departamento de Políticas Públicas, Avaliação e Gestão da Educação na Faculdade de Educação da Universidade do Estado do Rio de Janeiro (DEPAG/ UERJ), Rio de Janeiro, Rio de Janeiro, Brasil

donaldosouza@hotmail.com

Recebido em: NOVEMBRO 2014
Aprovado para publicação em: DEZEMBRO 2014

AVALIAÇÕES EDUCACIONAIS E SEUS REFLEXOS EM AÇÕES FEDERAIS E NA MÍDIA ELETRÔNICA

JOÃO LUIZ HORTA NETO

RESUMO

O objetivo deste trabalho é discutir os usos dos resultados das avaliações educacionais brasileiras por algumas ações desenvolvidas pelo governo federal no Plano de Desenvolvimento da Escola, seus possíveis reflexos no trabalho da escola e como a mídia eletrônica escrita vem tratando do tema. Discute-se o direito à aprendizagem e o processo de accountability, *eixos das ações governamentais nos últimos anos. A seguir, analisa-se o desenvolvimento dos testes aplicados aos alunos do ensino fundamental entre 1990 e 2012, identificando seus instrumentos e principais resultados. Devido à atual importância dos testes como um indicador de qualidade educacional, investigou-se, ainda, como essa qualidade é tratada pelos meios de comunicação eletrônicos. Conclui-se que os testes têm sido utilizados basicamente como instrumento para definir desempenhos das escolas, com reflexos indesejáveis sobre o trabalho que elas e seus profissionais realizam, na medida em que o foco das ações tem sido buscar melhorar o desempenho nos testes.*

PALAVRAS-CHAVE AVALIAÇÃO EDUCACIONAL • PROVA BRASIL • PROVINHA BRASIL • PLANO DE DESENVOLVIMENTO DA ESCOLA.

* Este trabalho, fruto da pesquisa de doutorado elaborado pelo autor sob a orientação da Professora Silvia Cristina Yannoulas, recebeu apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) dentro do programa Observatório da Educação.

RESUMEN

El objetivo de este trabajo es discutir los usos de los resultados de las evaluaciones educacionales brasileñas por intermedio de algunas acciones desarrolladas por el gobierno federal en el Plano de Desenvolvimento da Escola, *sus posibles reflejos en el trabajo de la escuela y la forma en que los medios electrónicos escritos vienen tratando el tema. Se discute el derecho al aprendizaje y el proceso de* accountability, *ejes de las acciones gubernamentales en los últimos años. A seguir se analiza el desarrollo de las pruebas aplicadas a los alumnos de la educación fundamental entre 1990 y 2012, identificando sus instrumentos y principales resultados. Debido a la actual importancia de las pruebas como un indicador de calidad educativa, se investigó asimismo cómo dicha calidad es tratada por los medios de comunicación electrónicos. Se concluye que las pruebas han sido utilizadas básicamente como instrumentos para definir desempeños de las escuelas, con reflejos indeseables sobre el trabajo que ellas y sus profesionales realizan, ya que el foco de las acciones ha sido intentar mejorar el desempeño en las pruebas.*

PALABRAS CLAVE EVALUACIÓN EDUCATIVA • PROVA BRASIL • PROVINHA BRASIL • PLANO DE DESENVOLVIMENTO DA ESCOLA.

ABSTRACT

The purpose of this study is to discuss the uses of the results of Brazilian educational evaluations for some actions developed by the Plano de Desenvolvimento da Escola, *their possible consequences in the work of schools and how the electronic written media has been approaching this topic. The right to learning and the process of accountability, axes of governmental actions in recent years, are discussed. Next, the development of the tests given to elementary school students between 1990 and 2012, identifying their instruments and principal results, are analyzed. Due to the current importance of the tests as an indicator of educational quality, how electronic communications media treat this quality is also investigated. It is concluded that the tests have been used, basically, as an instrument to define school performance, with undesirable consequences for the work that they and their professionals are doing, such that the focus of the actions has been to seek to improve the performance on the tests.*

KEYWORDS EDUCATIONAL EVALUATION • PROVA BRASIL • PROVINHA BRASIL • SCHOOL DEVELOPMENT PLAN.

INTRODUÇÃO

Avaliar é um conceito em si mesmo polissêmico, algumas vezes confundido com os instrumentos usados para medir resultados educacionais, geralmente por meio de testes. Como a atividade de avaliar vai além de simplesmente medir e divulgar os resultados, neste trabalho será usada a expressão "teste avaliativo" em referência ao instrumento de medida utilizado em avaliações educacionais.

Discutir o desempenho dos sistemas educacionais a partir de testes exige clareza sobre todo o processo de sua realização e aplicação e de seu sentido. Essa não é uma discussão trivial, pois discute a função social da escola e as modificações pelas quais ela vem passando ao longo do tempo e o papel do Estado nesse processo (HORTA NETO, 2010).

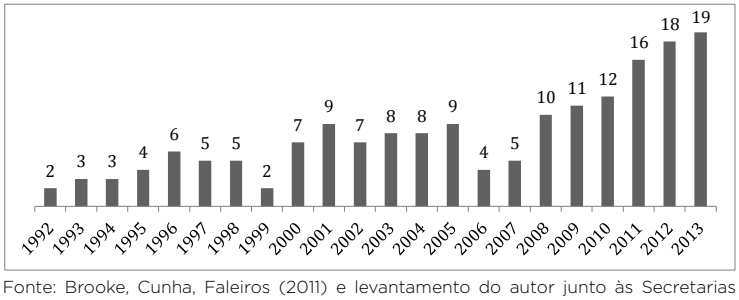
O desenvolvimento deste trabalho baseou-se em quatro hipóteses: tem havido um excesso de testes avaliativos, na expectativa de que sua aplicação e os resultados obtidos impulsionem desempenhos crescentes; apesar de os testes terem a capacidade de oferecer informações importantes para a melhoria da gestão dos sistemas de ensino, a ênfase nos

últimos anos tem sido na sua utilização para melhorar o desempenho do aluno; pode estar havendo um estreitamento curricular, na medida em que o ensino esteja voltado para o que é demandado pelos testes; os resultados dos testes não têm sido utilizados como balizadores das ações desenvolvidas pelo governo federal, mas sim como balizadores para as transferências voluntárias da União. Essas hipóteses serão discutidas ao longo do texto.

O CRESCENTE AUMENTO NO NÚMERO DOS TESTES

No Brasil, nos últimos anos, tem-se ampliado o número de governos municipais e estaduais e de gestores das redes privadas de ensino que utilizam testes elaborados externamente para medir o desempenho dos estudantes. Isso pode ser observado no Gráfico 1, que apresenta o avanço do número de governos estaduais que contam com avaliações próprias no período entre 1992 e 2013.

GRÁFICO 1 – Número de governos estaduais, por ano, que contavam com testes avaliativos próprios – 1992 a 2013



Fonte: Brooke, Cunha, Faleiros (2011) e levantamento do autor junto às Secretarias Estaduais de Educação.

Em 1992, os estados de Minas Gerais e Ceará começaram a aplicação dos testes avaliativos. Em 1995, além dos dois estados, desenvolveu-se o teste no Mato Grosso do Sul e em São Paulo (BROOKE; CUNHA; FALEIROS, 2011). Em 2012, 18 dos 27 estados da federação aplicavam testes cognitivos a seus alunos. Eram os seguintes os estados, e os nomes dos seus respectivos testes: Acre – Sistema Estadual de Avaliação da Aprendizagem Escolar (Seape); Alagoas – Sistema de Avaliação

Educacional de Alagoas (Saveal); Amazonas – Sistema de Avaliação do Desempenho Educacional do Amazonas (Sadeam); Bahia – Sistema de Avaliação Baiano de Educação (Sabe); Ceará – Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece); Espírito Santo – Programa de Avaliação da Educação Básica do Espírito Santo (Paebes); Goiás – Sistema de Avaliação Educacional do Estado de Goiás (Saego); Minas Gerais – Sistema Mineiro de Avaliação da Educação Pública (Simave); Mato Grosso do Sul – Sistema de Avaliação da Rede Pública de Mato Grosso do Sul (Saems); Paraíba – Sistema de Avaliação da Educação da Paraíba (Avaliando IDEPB); Paraná – Sistema de Avaliação da Educação Básica do Paraná (Saep); Pernambuco – Sistema de Avaliação Educacional de Pernambuco (Saepe); Piauí – Sistema de Avaliação Educacional do Piauí (Saepi); Rio de Janeiro – Sistema de Avaliação da Educação do Estado do Rio de Janeiro (Saerj); Rio Grande do Sul – Sistema Estadual de Avaliação Participativa (Seap); Rondônia – Sistema de Avaliação Educacional de Rondônia (Saero); São Paulo – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp); Tocantins – Sistema de Avaliação da Educação do Tocantins (Salto). Em 2013, o Pará anunciou a criação do Sistema Paraense de Avaliação Educacional (Sispae). Estão em processo de desenvolvimento o Sistema Maranhense de Avaliação do Desempenho Escolar (Simade), no Maranhão, e um teste avaliativo no Distrito Federal, que já havia criado, depois descontinuado, um sistema próprio anos atrás. Seguindo esse ritmo, é possível que os estados que ainda não possuem testes próprios anunciem os seus em curto espaço de tempo.

Analisando-se o Gráfico 1, percebe-se, desde 1992, uma tendência de crescimento do número de estados constituindo seus sistemas próprios de avaliação. Entretanto, antes de 2008, o crescimento apontado sofre algumas interrupções e quedas, firmando-se a partir de então. Cabe apontar dois momentos importantes. Em 2006, um ano após a primeira edição da Prova Brasil, o número de estados cai a menos da metade, o que faz sentido, pois, se o teste federal passou a ser censitário, muitos estados deixaram de aplicar os seus. O interessante é que, em 2008, ano seguinte à criação do

Índice de Desenvolvimento da Educação Básica (Ideb), dobra o número de estados constituindo sistemas próprios de avaliação, e, a partir de então, o número vem crescendo ano a ano, sugerindo que a criação do indicador de alguma forma estimulou os estados a desenvolverem seus testes.

A maior parte desses testes segue o mesmo modelo daquele instituído pelo governo federal para o Sistema de Avaliação da Educação Básica (Saeb)/Prova Brasil, tanto em termos das áreas de conhecimento e anos escolares testados, bem como na maior parte das habilidades cognitivas escolhidas para compor as Matrizes de Referência dos testes. Se já existem dados coletados pelo governo federal de forma censitária, questiona-se a razão de obtê-los novamente pelos governos subnacionais.

Obtidos os desempenhos das escolas, alguns governos começaram a utilizar estratégias de premiação com o objetivo declarado de estimular resultados escolares cada vez melhores. Baseados também nos resultados dos testes, os órgãos gestores passam a instituir programas de incentivos e pagamento de bônus para a equipe escolar por atingir metas, metas essas normalmente fixadas de forma unilateral. Outras estratégias conhecidas são a distribuição de prêmios e o sorteio de brindes aos alunos que mais se destacam nos testes.

No caso do governo federal, desde 1988, ano dos primeiros estudos para a aplicação do Sistema de Avaliação da Educação Básica (Saeb), realizado regularmente desde 1990, vem aumentando a quantidade de testes aplicados e o número de alunos testados. Depois do Saeb, criaram-se a Prova Brasil, em 2005, a Provinha Brasil, em 2008, e a Avaliação Nacional de Alfabetização (ANA) em 2013. Outros exames foram sendo criados e aprimorados: o Exame Nacional dos Estudantes do Ensino Médio (Enem) e o Exame Nacional de Certificação de Competências da Educação de Jovens e Adultos (Encceja). Como se percebe, diversos instrumentos que o governo federal tem a sua disposição, gerando grande volume de dados sobre o desempenho dos alunos, todos produzidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Contando apenas o Saeb/Prova Brasil, foram aplicados quase 19 milhões de testes aos alunos desde 1990 (HORTA NETO, 2013).

Além de criar testes avaliativos próprios, o Brasil vem participando de alguns estudos internacionais, como o Programme for International Student Assessment (Pisa), patrocinado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), e o Laboratório Latino Americano de Evaluación de la Calidad de la Educación (LLECE), programa coordenado pela Oficina Regional de Educación para América Latina y El Caribe (Orealc), entidade ligada à Unesco.

Nessa profusão de testes nacionais e internacionais, pode acontecer de uma mesma escola ter seus alunos testados por diversas vezes ao longo do ano, envolvendo pré-testes, utilizados para avaliar a qualidade dos itens que comporão os testes, e testes municipais, estaduais, nacionais e internacionais.

Junto com essa quantidade de testes, cria-se, também, um mercado da avaliação educacional que está em franco desenvolvimento no Brasil, oferecendo livros e consultorias com o objetivo de preparar as escolas para a participação nesses testes, assim como empresas e instituições desenvolvendo instrumentos para testar o desempenho das escolas e seus alunos e vendendo seus serviços para as secretarias estaduais e municipais ou para as escolas, nesse caso, privadas.

Para compreender melhor esse movimento na direção do incremento do número de testes, é importante discutir dois conceitos que se entrecruzam: o direito à aprendizagem e o *accountability*.

O DIREITO À APRENDIZAGEM E O PROCESSO DE *ACCOUNTABILITY*

A educação, como um direito social, está presente na Constituição brasileira de 1988, por ser inerente à cidadania e aos direitos humanos, tendo sido declarada como direito do cidadão e dever do Estado. No entanto, não é possível haver garantia do direito à educação sem assegurar que sua oferta atenderá às necessidades dos beneficiados, e isso implica o papel ativo e responsável do Estado, formulando políticas e obrigando-se a oferecer ensino em igualdade de condições para todos. Aliás, a própria Constituição prevê que essa

igualdade de condições se materialize, garantindo-se um padrão de qualidade.

Apesar de presente em nossa Constituição, o termo qualidade nunca foi explicitado. No Brasil, dado à sua polissemia, pode-se observar que, com o passar do tempo, vários foram os significados que lhe foram atribuídos: associada à ampliação da oferta de vagas, acarretando, a partir da década de 1940, a construção de prédios escolares; a partir do final da década de 1970 e durante a década de 1980, associada à permanência com sucesso na escola; a partir da década de 1990, associada à melhoria do desempenho cognitivo dos alunos, medida pela aplicação de testes avaliativos (OLIVEIRA; ARAÚJO, 2005). Nesse último período, como pode ser verificado pelos dados apresentados no Gráfico 1, o centro das atenções dos governos tem sido aumentar a quantidade de testes avaliativos. Uma das hipóteses que explicaria esse aumento seria a de que, a partir dos seus resultados, toda a comunidade escolar se mobilizasse por si, ou por pressão das famílias dos alunos, em busca de desempenhos cada vez melhores. Como consequência, a atenção do poder público estaria centrada em estratégias que garantissem melhorias nos resultados dos testes, e não em aprimorar o processo educacional. Dessa forma, a avaliação, em seu sentido amplo, perde seu sentido como instrumento da prática pedagógica e apequena-se.

Como o desempenho ganha a centralidade, o próprio direito à educação, envolvendo a garantia de acesso e permanência para todos e ensino de qualidade uniforme também para todos, trazendo em seu bojo o potencial de emancipar o indivíduo e de nivelar as desigualdades dos que frequentam a escola, acaba também se apequenando. Nos últimos anos, tem-se ouvido muito falar sobre o direito à aprendizagem, expressão que parece ter sido primeiramente utilizada pelo Unicef em duas publicações (FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA; BRASIL; UNIÃO DOS DIRIGENTES MUNICIPAIS DE EDUCAÇÃO, 2008; BRASIL; FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA, 2007), e que pouco tempo depois, passou a ser utilizada pelo Ministério da Educação (MEC) quando lançou documento sobre os direitos de

aprendizagem para o ciclo de alfabetização (BRASIL, 2012). Dessa forma, algo que no espírito da legislação educacional se relaciona com a formação básica do cidadão, em íntima relação com o direito à educação, transmuta-se em direito à aprendizagem da leitura, da escrita e do cálculo, conforme se depreende da leitura do documento citado.

A aprendizagem não pode ser compreendida fora do direito à educação, assim como do ensino de qualidade, de prédios escolares adequados e de outros fatores que permitem garantir esse direito. Reduzir o direito à educação, um direito tão amplo e fundamental à construção da cidadania, a algo restrito apenas à aprendizagem, é um equívoco conceitual e político. Conceitual, porque o direito à educação é muito mais amplo do que a aprendizagem; político, porque ainda inexistem condições iguais para todas as escolas brasileiras, e retirar o foco disso é como aceitar a imensa desigualdade existente. Mesmo que o objetivo fosse usar uma expressão forte para dar visibilidade a um problema que merece ser atacado – as dificuldades na aprendizagem –, isso não justificaria a interpretação reducionista dada à Lei.

Independentemente das razões que levaram à formulação do termo direito à aprendizagem, o fato é que por tal formulação se justificaria, pelo menos em parte, o aumento dos testes. Isso porque, se a aprendizagem da leitura, da escrita e do cálculo é um direito, os testes para medir a aprendizagem seriam os instrumentos para verificar o cumprimento desse direito. Assim, a aprendizagem escolar se restringiria àquilo que pode ser medido pelos testes, novamente uma redução de algo muito maior, pois para que a aprendizagem ocorra é necessário que se desenvolvam também diversas outras habilidades que nem sempre podem ser medidas pelos testes avaliativos.

Ao medir, por testes, se o direito à aprendizagem está ou não sendo atendido, o passo seguinte nessa lógica seria expor publicamente os resultados, premiando aquelas instituições e seus profissionais que realizam bem seu trabalho, ou apontar aqueles que não o fazem, na esperança de que nas próximas medições, principalmente por conta da pressão dos resultados, eles sejam melhores do que os anteriores.

Nesse processo, muitas vezes deixa-se a escola e seus profissionais expostos à própria sorte, como se não necessitassem de apoio para desenvolver suas atividades.

Essa lógica perversa faz parte de uma série de políticas que tomam por base o *accountability* e que se espalha pelo mundo. A esse fenômeno, Sahlberg (2011) dá a sugestiva alcunha de *Germ* (germe), abreviatura de *Global Educational Reform Movement* (Movimento Global de Reforma Educacional). Segundo o autor, esse germe viaja com os especialistas, a mídia e os políticos, espalhando-se pelo mundo, trazendo como consequências mal-estar entre os professores e menor aprendizagem para as crianças. Entre os sintomas indicados pelo autor como causas das doenças ocasionadas por esse germe, duas se destacam: a maior competição na educação, baseada na crença de que a educação se aprimora quando as escolas competem entre si, criando o que Afonso (1999, 2005) chamou de Quase-Mercado; a forte responsabilização das escolas e a busca da efetividade do professor, ambos tendo por base os testes avaliativos aplicados aos estudantes.

Diversos autores sustentam que o aumento dos testes avaliativos têm causado efeitos danosos e indesejados, como o aumento do ensino focado nos testes, o estreitamento do currículo para priorizar a leitura e a matemática e a aproximação da pedagogia a uma instrução mecânica (RAVITCH, 2010, 2013; SAHLBERG, 2011; HORTA NETO, 2013).

Sahlberg (2011) afirma que existe outra forma de aprimorar a aprendizagem e cita o exemplo da Finlândia. Segundo o autor, o país se fixa em um ensino e aprendizado customizados, ao invés de modelos estandardizados; tem como foco o aprendizado criativo, envolvendo, além do estudo das diferentes áreas do conhecimento humano, as habilidades necessárias para o desenvolvimento da personalidade do indivíduo, em vez de um ensino com foco em leitura, matemática e ciências; estimula um currículo baseado na escola e desenvolvido pelos professores, em que os riscos e as incertezas quanto ao ensino e à aprendizagem são aceitáveis, em vez de um currículo escolar prescritivo; baseia suas ações na responsabilidade compartilhada e na confiança em vez de *accountability* baseado em testes e controle.

No caso brasileiro, o governo federal tem ampliado a cada ano a abrangência e a quantidade de testes avaliativos voltados para a educação básica, uma ação que vem recebendo destaque em relação a outras, como a formação de professores para atuar nesse nível de ensino, sua atribuição legal. Para justificar essa afirmação, optou-se por acompanhar, tomando por base o ensino fundamental, a evolução do Saeb, a criação de novos testes, algumas das ações desenvolvidas pelo governo federal e seus reflexos sobre os sistemas de ensino e as escolas.

A EVOLUÇÃO DOS TESTES PARA AFERIR O DESEMPENHO DO SISTEMA EDUCACIONAL: OS OBJETIVOS INICIAIS E SUAS ALTERAÇÕES

Entre as décadas de 1960 e 1980, uma série de acontecimentos no Brasil contribuiu para o desenvolvimento de e*xpertise* na área de avaliação educacional. O primeiro deles, sem dúvida, foi o surgimento dos vestibulares unificados na década de 1960 em São Paulo.

Esses vestibulares unificados alteraram a lógica anterior, que estipulava uma nota mínima para o ingresso na faculdade. Pela nova lógica, a seleção de candidatos é feita até o limite do número de vagas e, para tanto, novos instrumentos de medida começam a ser desenvolvidos. O sucesso desse modelo e a utilização dos testes objetivos fizeram com que os itens de múltipla escolha ganhassem destaque nos livros didáticos, passando a ser amplamente utilizados pelos educadores (GATTI, 2002).

Para elaborar instrumentos cada vez mais sofisticados e confiáveis, capazes de dar conta do desafio de selecionar, entre milhares de candidatos, aqueles mais bem preparados para enfrentar a educação superior, foi necessário o desenvolvimento de equipes cujo objetivo fosse criar instrumentos de medida sofisticados e aprimorar a elaboração dos testes. Criaram-se, assim, as bases de conhecimento para as etapas seguintes.

SAEB E PROVA BRASIL

Pode-se classificar o Saeb pelas três fases que o modificam substancialmente, apesar de outros autores proporem diferentes divisões (BONAMINO; SOUZA, 2012). A primeira fase, a dos primeiros estudos até sua implantação, entre 1980 e 1993, começou com as pesquisas iniciadas durante o Projeto Edurural, passou pela aplicação piloto de 1988, e envolveu as aplicações que ocorreram nos dois ciclos seguintes, de 1991 e 1993. A segunda fase, entre 1995 e 2003, envolvendo o esforço para aprimorar seus instrumentos, foi iniciada com o uso da Teoria da Resposta ao Item (TRI), no ciclo de 1995; a construção das Matrizes de Referências, no ciclo de 1997 e a consolidação desse modelo metodológico nos três próximos ciclos, entre 1999 e 2003. A terceira fase se iniciou no ciclo de 2005, perdurando até os dias de hoje, com a transformação do teste amostral em censitário.

O que chama a atenção nesse processo é que não houve descontinuidade no desenvolvimento e no amadurecimento do Saeb, processo iniciado na época do regime militar, durante o governo de João Figueiredo (1979-1985), e que atravessou os governos de José Sarney (1985-1990), Fernando Collor (1990-1992), Itamar Franco (1992-1995), os dois governos de Fernando Henrique Cardoso (1995-2003), os dois governos de Luiz Inácio Lula da Silva (2003-2011) e o governo de Dilma Rousseff. Foram governos com diferentes orientações políticas, mas em todos se privilegiou a realização de medidas de desempenho cognitivo como uma das estratégias para aprimorar a qualidade da educação. O que mudou ao longo do tempo foi a abrangência da avaliação e os usos que se fizeram de seus resultados.

Na primeira fase, com o envolvimento de Jacobo Waiselfitz, Heraldo Vianna e Bernadete Gatti, a maior preocupação era envolver os professores na elaboração dos itens do teste e em uma discussão coletiva dos resultados obtidos. Por causa disso, apesar de Heraldo Vianna reconhecer a importância do uso da TRI, utilizou-se a Teoria Clássica dos Testes para apresentar os resultados dos testes, pois essa forma de apresentação, dada em função do número de acertos neles, facilitaria a compreensão dos professores (VIANNA, 1990).

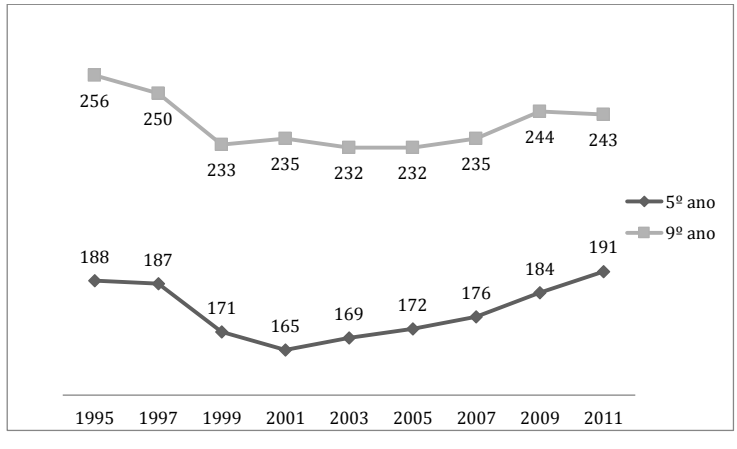
Além disso, a escolha dos anos escolares a serem testados (1ª, 3ª, 5ª e 7ª séries) recaiu sobre aqueles em que, reconhecidamente, os alunos apresentavam maiores problemas, e que, por isso, precisariam ser mais bem investigados (KORTE; NASH, 1995).

A segunda fase inaugura a busca pela excelência da medida, abandonando progressivamente a preocupação com a compreensão dos resultados, que passam a ser apresentados em uma escala que vai de 0 a 500 pontos, permitindo comparações a cada ciclo. Outra característica nessa fase foi a alteração dos anos escolares testados (4ª e 8ª séries) com a justificativa de que dessa forma seria possível garantir uma maior homogeneidade nos conteúdos escolares ensinados, algo que seria muito disperso na escolha dos anos escolares anteriores.

O processo de *accountability* se aprofunda a partir da terceira fase, iniciada em 2005, ano em que os testes, agora chamados de Prova Brasil, são aplicados de forma quase censitária às escolas públicas, e seus resultados passam a compor um índice destinado a aferir a qualidade da educação básica, o Ideb.

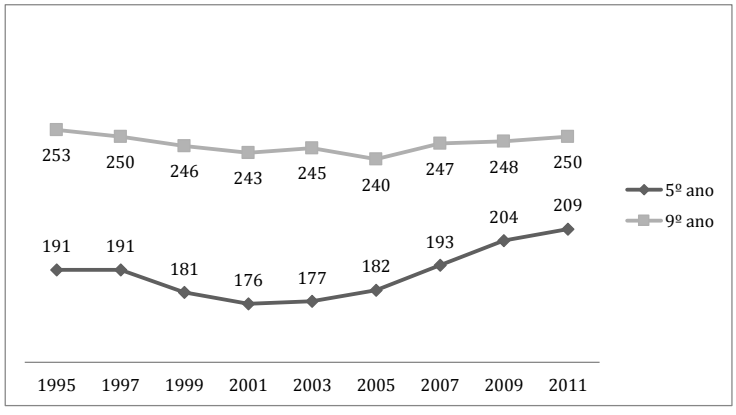
Os Gráficos 2 e 3 mostram como variaram as proficiências médias em Língua Portuguesa e Matemática, respectivamente, no Saeb e na Prova Brasil entre 1995 e 2011.

GRÁFICO 2 – Evolução da proficiência média em Língua Portuguesa no Saeb/Prova Brasil – escolas públicas – 1995 a 2011



Fonte: Elaboração do autor com base nos dados do Inep (BRASIL, 2011).

GRÁFICO 3 – Evolução da proficiência média em Matemática no Saeb/Prova Brasil – escolas públicas – 1995 a 2011



Fonte: Elaboração do autor com base nos dados do Inep (BRASIL, 2011).

Nos dois gráficos, é apresentada a variação da proficiência obtida nos testes. Para começar a análise das proficiências no contexto em que foram colhidas, é preciso ter em mente que em toda a média podem esconder-se muitas disparidades, portanto ela não é a medida mais transparente de todas.

Serão discutidos três períodos: de 1995 a 2001, o ano de 2003 e de 2005 a 2011.

O período de 1995 a 2001 foi marcado pela queda constante das proficiências em todos os anos escolares e em ambas as áreas. Esse período coincide com os dois governos Fernando Henrique Cardoso, e foi marcado por uma grande discussão envolvendo a queda da qualidade da educação pela significativa inclusão das camadas mais pobres da população na escola. Apesar de ser uma hipótese plausível, não se sabe de nenhum estudo que tenha sido feito para sustentá-la. No entanto, isso poderia explicar a queda da proficiência nos anos iniciais do ensino fundamental, mas não a queda nos anos finais, a menos que se imagine que o movimento de entrada na escola foi acompanhado por outro, de retorno a ela.

Em 2003, já no primeiro governo de Luiz Inácio Lula da Silva, o teste não sofre nenhuma alteração, e nos anos iniciais do ensino fundamental há uma melhora que, na época, o relatório técnico creditou às políticas implantadas pelo novo governo. Mas como explicar também as melhoras no 9º ano?

O ano de 2005, quando começa a Prova Brasil, marca o início da trajetória ascendente nas proficiências para os anos iniciais e finais do ensino fundamental, apesar de nos anos finais o crescimento ser de magnitude menor. Em 2007 entra em cena o Ideb, e o MEC divulga em seu *site* um simulado para ajudar os professores e as escolas a se prepararem para o teste. Nesse ano também, os gestores recebem, meses antes da realização do teste, o valor do Ideb de 2005 das redes e das escolas e as metas que deveriam ser atingidas no Ciclo de 2007. Será que o aumento das proficiências a partir de 2005 estaria melhorando a qualidade do ensino oferecido e isso se refletiria na elevação das proficiências medidas?

Sem pretender aqui ser exaustivo na discussão, algumas hipóteses podem ser levantadas. A primeira é que, encerrado o fluxo acelerado de entrada na escola, tenha havido uma acomodação, e, com isso, tenha sido possível melhorar o atendimento aos alunos. Outra é que o aumento no valor dos programas de transferência de renda, capitaneados pelo Programa Bolsa Família, e o aumento das famílias beneficiadas tenham contribuído para retirar da miséria muitas crianças e jovens, com reflexos na aprendizagem. Efeito semelhante poderia ser creditado ao aumento da renda da população, fruto do período de crescimento econômico vivido nos últimos anos. Até aqui, referimo-nos aos efeitos externos que impactam a aprendizagem, positivos e desejáveis. Outra hipótese é que o aumento da escolaridade para nove anos, a partir de 2006, colocando a criança mais cedo em contato com um ambiente de aprendizagem, possa também estar contribuindo para o aumento das proficiências, mas, nesse caso, por ser uma medida recente, a contribuição ainda deve ser mínima.

Pode ser também que houvesse um problema de gestão nas redes e nas escolas, e, com o impacto das pressões por melhores desempenhos, os gestores tenham desenvolvido estratégias para dar mais apoio às escolas, e esse apoio tenha chegado aos professores, estando eles, assim, mais preparados para melhorar a aprendizagem de seus alunos. Por fim, uma hipótese, que se lê nas entrelinhas das matérias jornalísticas e nas declarações de alguns gestores educacionais,

é a de que os gestores escolares e os professores estivessem fazendo "corpo mole", algo que o Ideb estaria denunciando. Com isso, eles começaram a se esforçar mais, alcançando melhores resultados. Até aqui todas as hipóteses levantadas refletiriam em benefício das crianças e dos jovens.

Mas existem também hipóteses que trariam prejuízos ao ensino. Pode ser que, por conta da pressão por melhorias dos indicadores das escolas, esteja havendo uma redução do currículo, que estaria moldando-se às Matrizes de Referências dos testes. Além disso, por conta dessa mesma pressão, as escolas podem centrar uma boa parte de suas atividades na preparação para o teste, principalmente pela aplicação de simulados, aproveitando os próprios modelos disponibilizados pelo MEC. Esses dois fatores também podem estar na raiz da explicação para o crescimento mais acelerado das proficiências nos anos iniciais do que nos finais. Como nos anos iniciais funciona a unidocência, é mais fácil para o professor organizar seu tempo e, dessa forma, preparar melhor seus alunos para o teste, algo mais complicado na etapa seguinte, com vários professores de diferentes áreas. Essas práticas, além de não colaborarem para a aprendizagem, estariam contribuindo para a homogeneização do ensino.

Todas essas hipóteses devem ser averiguadas, mas, no caso desse trabalho, a hipótese da preparação para o teste é investigada mais a frente.

PROVINHA BRASIL

A Provinha Brasil, aplicada pela primeira vez em 2008, testa os alunos no segundo ano de escolaridade nas áreas de Língua Portuguesa e Matemática no início e no final de cada ano, utilizando os chamados Testes 1 e 2, respectivamente, em cada um desses momentos. O teste é composto, em sua maioria, por itens de múltipla escolha, divididos em três tipos: aqueles em que o professor lê totalmente o item e os alunos marcam a alternativa correta; outros em que o professor lê uma parte do item, o aluno, a outra, e, então, assinala a resposta correta; por fim, itens que o aluno lê sozinho e marca a alternativa correta. Além desses, constam do teste

três outros itens, sendo que em dois deles o professor faz um ditado de palavras, e, no terceiro, faz um ditado de uma frase. Para os itens de múltipla escolha, são fornecidos os gabaritos, e, para os três itens do ditado, é fornecida uma grade de correção com seis possibilidades, em que três delas indicam aquelas consideradas corretas. Foram definidos cinco níveis de proficiência, utilizados para classificar os alunos segundo seu estágio de alfabetização, numerando-os de 1 a 5, e a cada um deles se associou um número de acertos, mas apenas para os itens de múltipla escolha. Para auxiliar a compreensão dos professores sobre os resultados do teste, foi apresentada uma interpretação pedagógica de cada um dos níveis.

Essa forma de apresentar os resultados levanta algumas preocupações. Primeiramente, o problema de uma classificação como essa é enquadrar um processo tão dinâmico e complexo como a alfabetização em uma determinada categoria, descrevendo-o a partir de itens construídos com base em uma Matriz de Referência que não necessariamente reflete o currículo praticado na escola e o definido pelo sistema de ensino em que a escola está inserida. Além disso, para o professor, o processo parece uma mágica: como associar um número de acertos de itens, independentemente do item que o aluno acertou, com uma etapa do processo de alfabetização que apresenta uma descrição tão fechada? Assim, apesar de, na Provinha Brasil, o próprio professor aplicar a prova, ficar com ela para compreendê-la melhor, corrigi-la e iniciar um processo de análise sobre o significado do desempenho dos seus alunos, tudo isso é feito a partir de um planejamento que não foi realizado por ele, a partir de itens que ele não criou e usando uma interpretação apresentada a ele de forma incompleta.

A forma como a Provinha Brasil é utilizada depende de cada rede. Em algumas, assim que os instrumentos são recebidos, eles são enviados para as escolas, que decidem quando devem aplicá-los e como utilizar seus resultados. Em outras redes, repete-se o que o governo federal faz com o Saeb: a Secretaria marca a data de aplicação, percorre as escolas aplicando os testes, coleta as folhas de resposta, faz a

correção, analisa os resultados e só depois os discute com as escolas, normalmente fixando metas que devem ser atingidas por cada uma. Assim, um instrumento que deveria ser autonomamente utilizado pelas escolas pode, em alguns casos, repetir o esquema do Saeb e se transformar em mais um teste com o objetivo de controlar a escola e seus professores.

Com o Pacto Nacional de Alfabetização na Idade Certa (Pnaic), que será discutido mais à frente, a Provinha Brasil passa a ser obrigatória em todos os municípios que a ele aderirem. Além disso, os resultados de cada rede devem ser inseridos em um *software* que será disponibilizado pelo Inep e que permitirá ter um panorama nacional, escola por escola, do desempenho dos alunos. Aumenta-se a pressão por resultados envolvendo crianças muito jovens, sem que se tenha estudado os impactos que isso pode trazer ao seu desenvolvimento.

AS AÇÕES DO GOVERNO FEDERAL QUE UTILIZAM OS RESULTADOS DAS AVALIAÇÕES

Nos dois próximos tópicos, discutem-se o Plano de Desenvolvimento da Educação (PDE), o principal programa educacional do governo federal e um de seus componentes, o Pnaic, buscando compreender como os resultados dos testes avaliativos os influenciaram.

PLANO DE DESENVOLVIMENTO DA EDUCAÇÃO (PDE)

O PDE constitui-se de um conjunto de ações e programas voltados para a educação básica e superior desenvolvidos pelo MEC. Lançado em 2007, o programa possui seis pilares (BRASIL, 2011), destacando-se dois deles: responsabilização, que o texto apresenta como uma tradução de *accountability*, afirmando que, se a educação é um direito de todos e dever do Estado e da família, deve-se responsabilizar os que não garantem esse direito, sobretudo a classe política; mobilização social, fundamental para acompanhar e fiscalizar as ações educacionais, sendo que para que isso aconteça deve haver transparência no trato das questões ligadas à educação e nas políticas destinadas a seu desenvolvimento.

O instrumento legal que normatizou o PDE foi o Plano de Metas Compromisso Todos pela Educação, que nada mais é do que a explicitação de como ocorrerão as transferências voluntárias da União para os estados e os municípios e em que áreas o MEC prestará assistência técnica e financeira. Sob a justificativa da necessidade de implementar o regime de colaboração entre os entes federados, é estabelecido um compromisso, de que tomam parte a União, os estados e os municípios, com base em 28 diretrizes definidas pelo governo federal que garantiriam a melhoria da educação básica (BRASIL; FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA, 2007).

Claramente, o documento impõe aos sistemas de ensino uma série de atribuições que lhes deveriam ser próprias, e sobre as quais a União não deveria ter ingerência, pois o que se deveria buscar seria um regime de colaboração. Além disso, associa o Ideb ao leme que deve guiar as ações desses sistemas, tornando-o, assim, o indicador da qualidade educacional. É importante destacar a limitação desse indicador, pois ele se refere apenas às medidas de desempenho em Língua Portuguesa e Matemática na Prova Brasil e do fluxo escolar, parcelas importantes, mas não determinantes, no conjunto de fatores que impactam uma educação de qualidade.

Com relação a aspectos educacionais propriamente ditos, das 28 diretrizes, apenas três se relacionam com o tema, com destaque àquela que estabelece uma meta para que a alfabetização esteja concluída até os oito anos de idade, equivalente ao final do 3º ano do ensino fundamental, algo que Soares (2010) considera possível, e que garantiria, em tese, a possibilidade de desenvolvimento de novas habilidades nos anos seguintes a essa etapa de ensino. No entanto, a diretriz impõe a realização de um exame para aferir se a alfabetização ocorreu ou não. Essa imposição é colocada parcialmente em prática com o lançamento do Pnaic.

O que chama a atenção no Decreto é que por trás de um compromisso, um simples acordo envolvendo os entes da federação, seja imposta a seus signatários uma série de obrigações fixadas pela União. Poder-se-ia argumentar que como o compromisso é de natureza voluntária, quem estiver

em desacordo com seus termos tem a opção de não o assinar. No entanto, o Art. 8º do Decreto abre a possibilidade para que a União, unilateralmente, se exima de colaborar com o ente que deixar de firmá-lo. O fato é que todos os estados e municípios, sem exceção, firmaram esse compromisso. Mesmo porque nenhum ente da federação tem condições de dispensar os recursos federais que podem vir a ser transferidos, nem tampouco deixar de assumir o compromisso público de melhorar a educação sob sua responsabilidade.

O PACTO NACIONAL DE ALFABETIZAÇÃO NA IDADE CERTA (PNAIC)

Lançado em julho de 2012, com investimentos previstos de R$ 3,3 bilhões, o Pnaic é um programa com o objetivo de garantir que todas as crianças estejam alfabetizadas até os oito anos de idade, ao final do 3º ano do ensino fundamental.

Ele possui dois componentes principais: um curso de formação em serviço para cerca de 360 mil professores em atuação no chamado ciclo de alfabetização, aquele que compreende do 1º ao 3º ano do ensino fundamental, e a aplicação de testes para verificar o estágio da alfabetização dos alunos. Os entes da federação que quiseram aderir ao pacto assinaram compromisso envolvendo as duas ações do programa.

O curso de formação em serviço, primeiro componente do Pnaic, ocorrerá em diversos momentos durante o período 2013-2014, com ênfase em Língua Portuguesa, no primeiro ano, e em Matemática, no segundo ano. O segundo componente são dois testes para acompanhar a evolução da aprendizagem. A primeira testagem ocorre no início e no final do 2º ano do ensino fundamental pela Provinha Brasil. Como já comentado antes, até 2012 esses testes eram utilizados apenas por aquelas redes que o quisessem, mas, a partir da adesão ao Pnaic, todas são obrigadas a realizá-lo, enviando seus resultados ao Inep. Além disso, a partir de 2014, tem sido utilizada a ANA para testar os alunos no final do 3º ano do ensino fundamental, último ano do ciclo de alfabetização.

Com as novas ações do Pnaic, os alunos dos anos iniciais do ensino fundamental, de duração de cinco anos, serão

testados quatro vezes pelo governo federal: no 2º ano, por duas vezes, com a Provinha Brasil, no 3º ano com a ANA e no 5º ano com a Prova Brasil, todos de forma censitária. Um enorme esforço que vai mobilizar toda a comunidade escolar.

Outra inovação do Pacto é a instituição da premiação por desempenho, estando previstos recursos da ordem de R$ 500 milhões para premiar as escolas que apresentarem os maiores progressos, podendo uma parte desse valor ser destinada à premiação da equipe escolar. Assim, o governo federal aproxima-se de outras iniciativas que também premiam a *performance* das escolas, provavelmente na expectativa de que, com esse bônus, os professores se esforcem mais para conseguir que seus alunos obtenham desempenhos melhores.

Alguns aspectos dos programas de formação continuada desenvolvidos pelo governo federal merecem ser discutidos. Primeiro é um programa pontual, que tem por objetivo alinhar os docentes com as "últimas decisões em matéria de política educacional dos governos" (LUCIO, 2010, p. 51).

O segundo aspecto, relacionado a uma possível razão para que programas federais de formação continuada sejam pouco efetivos, tem relação com a elevada rotatividade da carreira docente. Existem estudos (HORTA NETO, 2013) mostrando uma rotatividade média anual de 30% da equipe docente das escolas públicas de ensino fundamental, ou seja, de cada dez professores que estão cadastrados em um determinado ano em uma escola, somente sete serão novamente docentes nessa mesma escola no ano seguinte. Em alguns estados e em algumas redes, essa rotatividade supera 60%, e chega a ser maior na rede privada, mostrando tratar-se de um problema envolvendo a carreira docente como um todo, e não problema da rede pública. Quando se compara o intervalo de cinco anos, entre 2007 e 2012, a rotatividade nacional média dos docentes nas escolas públicas atinge patamares alarmantes: 43% para a rede estadual e 66% para a municipal.

Quando a rotatividade apresentada nesse estudo é comparada com estudos produzidos sobre a realidade americana, a rotatividade nas escolas brasileiras impressiona. As pesquisas nos Estados Unidos mostram que 30% dos novos professores abandonam a profissão depois de cinco anos (INGERSOLL, 2001), e

que em estados como o Texas, anualmente, entre 11% e 18% dos professores deixam a profissão, e entre 6% a 10% mudam de escola, sendo que a maior porcentagem refere-se a professores com até dois anos de formados (HANUSHECK; KAIN; RIVKIN, 1999).

Com uma rotatividade nos patamares em que se encontram, os programas de formação continuada de professores, complexos por serem gerenciados de Brasília, e de longa duração, não podem limitar-se a um único ciclo, pois sempre haverá um novo contingente de profissionais que não participaram desses programas. Dessa forma, os mesmos programas teriam que ser oferecidos por vários anos, ou então deveriam ser de tal forma estruturados que permitissem às secretarias continuar essa tarefa por conta própria. Nesse caso, esbarra-se em problemas de falta de pessoal especializado para levar a cabo essa tarefa.

Ao analisar o material desenvolvido pelo Pnaic, verifica-se que ele é essencialmente um curso sobre o processo de alfabetização, com suas técnicas e estratégias, algo que deveria fazer parte de qualquer curso de formação de professores. Como os professores em formação não têm tido a possibilidade de se aprofundar nesse tema durante a graduação, devido a problemas estruturais de seus cursos, fica a lacuna. Não seria mais produtivo o MEC atuar mais fortemente nos cursos superiores de formação de professores, já que, pela legislação, a União é que tem responsabilidade sobre a educação superior? Se isso fosse feito, a lacuna seria diminuída e permitiria a discussão com os outros entes da federação de programas mais simples, mas nem por isso de menor qualidade, voltados a temas específicos que pudessem ser desenvolvidos localmente. Da forma como os programas estão atualmente estruturados, a formação continuada está voltada para a complementação de falhas na formação inicial.

Pelos dados apresentados, fica claro que os resultados dos testes que vêm sendo aplicados nos últimos 25 anos não foram utilizados para o desenvolvimento dos dois programas analisados. Buscou-se, isso sim, instituir metas de desempenho.

Na seção a seguir, serão apresentadas evidências sobre ações que estão acontecendo nas escolas no sentido de buscar

aumentar o desempenho dos alunos nos testes, sem necessariamente melhorar suas aprendizagens.

A DISCUSSÃO DA QUALIDADE EDUCACIONAL NA MÍDIA ELETRÔNICA ESCRITA

Como uma das premissas do *accountability* é que, com a exposição pública dos resultados dos testes de desempenho, haja uma mobilização da sociedade em busca da melhoria da escola, surgiu o questionamento de como a mídia eletrônica estaria tratando os resultados desses testes. Para isso, entre 10 de janeiro de 2009 e 11 de outubro de 2012, foram coletadas matérias disponibilizadas na internet, como artigos de jornais, revistas e textos de *sites* e *blogs*. Isso foi feito com a utilização da ferramenta "Alerta", do Google, em uma pesquisa com as seguintes expressões-chave: avaliação educacional, avaliação da educação, *educational evaluation*, *evaluación educativa*, *evaluación de la educación*, Prova Brasil, Provinha Brasil, sistema nacional de avaliação da educação básica, Saeb. Além disso, foi utilizado também o *clipping* de notícias do Movimento Todos pela Educação e o do Inep.

No total, foram analisados 2.100 *links*, apresentados no Quadro 1, no qual estão classificados os textos coletados em quatro categorias: "avaliações", "julgamento do resultado obtido", "responsável pelo resultado obtido" e "preparação para os testes".

QUADRO 1 – Frequência das matérias coletadas no período entre 10 de janeiro de 2009 e 11 de outubro de 2012, utilizando a ferramenta "Alertas do Google", quanto à classificação em temas gerais e específicos

TEMA GERAL	TEMA ESPECÍFICO	FREQUÊNCIA
1. Avaliações	Saeb, Prova Brasil e Ideb	635
	Enem	493
	Avaliações estaduais	165
	Provinha Brasil	174
	Prova ABC	46
	Avaliações municipais	35
	Avaliações privadas	5
	Crítica à avaliação	0
	Subtotal	1.553
2. Julgamento do resultado obtido	Melhora da nota (*desempenho bom*)	175
	Queda da nota (*desempenho ruim*)	171
	Subtotal	346
3. Responsável pelo resultado obtido	Professora e equipe escolar	169
	Recursos	86
	Equipe da secretaria	49
	Curso (secretaria, particulares, *on-line*)	41
	Subtotal	345
4. Preparação para os testes	Secretarias	132
	Interna à escola	11
	Subtotal	143
Total		**2.387**

Fonte: Elaboração do autor.

Note-se que, como um mesmo texto poderia tratar de diferentes temas, a frequência total foi de 2.387, superior aos 2.100 *links* analisados. Por questão de espaço, apenas parte do Quadro 1 será tratado. Para uma discussão mais detalhada, conferir Horta Neto (2013).

A primeira observação é que dos 2.100 *links*, nenhum deles apresentava críticas aos testes avaliativos no Brasil. No máximo, o que se encontrou foram sugestões de aprimoramento ao sistema adotado, passando uma imagem de que existiria uma unanimidade com relação a eles.

O tema mais citado, com 1.553 registros, trata das avaliações em geral, e, dentro dele, os mais citados referem-se

ao Ideb, ao Saeb ou à Prova Brasil, recebendo 635 referências. A maior parte desses registros versa sobre o Ideb, informando o resultado alcançado, comparando-o com os anteriores e tecendo comentários sobre os desempenhos em Língua Portuguesa e Matemática. Quanto ao Saeb e à Prova Brasil, a maior parte dos textos discorre sobre a preparação para a aplicação dos testes e sobre suas realizações. No geral, a principal característica dos textos ligados a esse tema é ser de natureza informativa.

A respeito da Provinha Brasil, que recebeu 174 citações, apesar de ser um teste, à época da pesquisa, que permitia a cada escola aplicá-lo e corrigi-lo, as matérias produzidas pelas diversas mídias mostram que, gradativamente, as Secretarias de Educação assumem o controle da aplicação, da apuração e da divulgação dos resultados, utilizando para isso julgamentos de valor e o ranqueamento de escolas com base em níveis definidos pelo Inep para classificar os desempenhos dos alunos.

No caso das avaliações próprias, além dos testes elaborados pelos estados, as matérias apontaram o crescente número de municípios que utilizam essa estratégia a partir de 2011, indicando que 35 municípios passaram a utilizá-la, seja produzindo instrumentos localmente ou adquirindo-os de empresas e instituições especializadas. Juntamente com os testes, os municípios desenvolvem também indicadores, tomando como base a metodologia do Ideb. Apesar de as notícias tratarem de um número pequeno do total de 5.568 municípios brasileiros, o que se percebe é uma crescente utilização dessa estratégia. O tempo gasto nas escolas, na aplicação dos testes do município, do governo estadual e do federal e na preparação dos alunos para aumentar seu desempenho nesses testes, principalmente na forma de simulados, vem aumentando significativamente e deve estar tomando uma parcela importante do tempo que deveria ser dedicado à aprendizagem.

Com 345 referências, encontram-se as matérias que mostram quais seriam os responsáveis pelos resultados obtidos, fossem bons ou ruins. Atribui-se, em sua maioria, aos professores e à equipe escolar a responsabilidade pelos bons resultados, associando à função docente uma imagem messiânica,

relacionada a um dom para educar, e à escola como o templo sagrado do saber, não como um local de disputas e conflitos. As matérias ressaltam a dedicação e o empenho dos professores e o espírito de equipe dos profissionais da escola em torno do objetivo de garantir a aprendizagem dos estudantes como responsáveis pelo sucesso da escola. Esses fatores superariam todo tipo de obstáculos, principalmente a falta de recursos, os baixos salários e as condições sociais adversas das comunidades em que as escolas estão localizadas.

O tema "preparação para os testes" envolveu matérias que tratavam explicitamente de estratégias utilizadas pelas secretarias ou pelas escolas para melhorar o desempenho dos alunos na Prova Brasil, aumentando, assim, seu Ideb. O total de referências coletadas representaram redes de ensino que congregam, no ensino fundamental regular, 4.300 escolas de ensino fundamental onde estão matriculados quase 1,5 milhões de alunos, representando 6% do total de alunos matriculados nessa etapa de ensino e que estão localizadas nas regiões geográficas mais pobres do Brasil. Fica patente nos textos que a preocupação não é com o aprendizado das crianças e dos jovens, e sim em elevar o valor do Ideb para, dessa forma, demonstrar que o que vem sendo feito pelos governos e pelas escolas está dando resultado. Essa preocupação com o Ideb demonstra que, de certa forma, o governo federal está conseguindo mobilizar, pelo menos, a classe política para melhorar os indicadores educacionais.

É preciso deixar claro que é possível, em algumas situações específicas, utilizar simulados como uma das possíveis estratégias que o professor tem a seu dispor para trabalhar alguns aspectos de determinada disciplina. Mas essa estratégia, para ser efetiva no sentido de aprimorar determinados aspectos do que se quer desenvolver, tem de ter relação com a realidade de sala de aula: deve ser preparada pelo professor, dentro de determinado contexto, e tendo por base seu conhecimento sobre o conjunto de seus alunos. Caso contrário, o teste acaba apresentando o mesmo caráter dos simulados utilizados para preparar os alunos do ensino médio para o Enem, em seu papel de vestibular nacional. Nesse caso, os alunos buscam individualmente estratégias para obter uma

classificação mais alta que a de seus concorrentes, garantindo, assim, uma vaga no curso e na universidade escolhidos. Passada essa situação, e tendo sucesso em seu intento, pouco fica do que foi estudado, e o jovem, em pouco tempo, se esquece de grande parte do conteúdo utilizado para se preparar para o exame. Nesse caso, não houve aprendizagem significativa, pois tudo o que foi estudado só fazia sentido para garantir a vaga. No caso de testes como a Prova Brasil, o que seria de se esperar não é a busca de uma melhor classificação, mas sim a captação de informações que ajudassem a desenhar um quadro representativo da realidade educacional brasileira, para assim poder agir no sentido de melhorá-la. Da forma como parece que os simulados estão ocorrendo, os resultados que vêm sendo alcançados podem estar mascarando a realidade que se quer conhecer pelos testes.

Assim como no caso das secretarias de educação, algumas escolas também desenvolvem, por iniciativa própria, simulados para preparar seus alunos para o teste nacional. No caso das escolas, as matérias são mais escassas e foram colhidas na maior parte das vezes de *blogs* mantidos por elas ou por algum professor. Um dos casos descreve uma escola em um pequeno município do estado de Santa Catarina onde os professores afirmam que realizaram simulados a cada 15 dias. Principalmente nesse caso existe grande possibilidade de o ensino estar todo centrado na preparação para o teste.

Mesmo que se considere que seja necessário relativizar as informações obtidas a partir de matérias publicadas na internet, já que poderiam não apresentar dados fidedignos, a quantidade de alunos envolvidos no processo e a quantidade de depoimentos em favor da preparação para o teste não podem esconder que essa realidade é mais presente do que se poderia imaginar.

CONSIDERAÇÕES FINAIS

Desde os primeiros testes, tem-se produzido números com pouca relação com a prática pedagógica. Esses números isoladamente têm pouco significado para a escola, pois seu objetivo é buscar alternativas para garantir a aprendizagem de seus

alunos. Essa aprendizagem depende de outros fatores que não apenas um número que reflita o resultado de um teste, teste esse limitado àquilo que pode ser medido para algumas, não para todas, as áreas do conhecimento do currículo escolar e a um subconjunto de possibilidades para se elaborar um item, cada uma delas trazendo diferentes dificuldades.

Assim, como parece pouco provável que uma quantidade grande de escolas consiga apropriar-se dos resultados, apesar dos mais de 20 anos de tentativas para melhor expressá-los, provavelmente o problema esteja com o uso inadequado dos resultados dos testes avaliativos. Isso porque, inicialmente, esses instrumentos foram desenhados para avaliar políticas educacionais, evoluíram para avaliar sistemas educacionais e, sem que se alterassem seus instrumentos nem a forma como eram desenvolvidos, passou-se a querer que as escolas utilizassem seus resultados para orientar seu trabalho pedagógico. Talvez os testes avaliativos possam ser úteis para a comunidade escolar, caso ela participe ativamente no desenvolvimento dos instrumentos e na análise dos resultados obtidos, algo que se procurou nos anos iniciais do Saeb, mas foi abandonado depois. Com isso, não se quer afirmar que as medidas realizadas não tenham importância alguma, mas sim que têm, no geral, pouca capacidade de influenciar o trabalho pedagógico da escola.

Por outro lado, nos últimos anos, tem sido veiculada uma forte mensagem de que aumentar o número de testes é um bom caminho para aprimorar as aprendizagens. A prova disso é o aumento crescente do número de estados que estão testando seus alunos, repetindo os mesmos procedimentos dos testes federais sobre as mesmas escolas. A única justificativa plausível para ações como essa é a de aumentar a pressão em busca de desempenhos maiores, algo que não necessariamente significa uma melhoria na aprendizagem. Além disso, o próprio governo federal tem aumentado o número de testes e a abrangência dos existentes.

Esta pesquisa encontrou evidências da crescente utilização de estratégias de aplicar simulados às vésperas da aplicação dos testes, denotando uma clara preocupação com os resultados, e não com os processos educacionais. Dessa forma, pode es-

tar ocorrendo um estreitamento do currículo escolar àquilo que está sendo testado.

Por fim, apesar da grande ênfase dada pelo governo federal aos testes, o principal uso que se faz de seus dados é como auxiliar no processo de tomada de decisões relativas às transferências voluntárias da União. Com isso, pode-se estar estimulando apenas a busca por resultados, colocando os processos educacionais, que garantiriam uma aprendizagem significativa, em segundo plano.

REFERÊNCIAS

AFONSO, Almerindo J. Estado, mercado, comunidades e avaliação: esboço para uma rearticulação crítica. *Educação e Sociedade*, Campinas, v. 20, n. 69, p. 139-164, 1999.

_____. *Avaliação educacional*: regulação e emancipação. São Paulo: Cortez, 2005.

BONAMINO, Alícia C.; SOUZA, Sandra Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Resultados das edições anteriores*. Brasília, DF: Inep, 2011. Disponível em: <http://portal.inep.gov.br/web/saeb/resultados>. Acesso em: dez. 2014.

BRASIL. Ministério da Educação. *O Plano de Desenvolvimento da Educação*: razões, princípios e programas. Brasília: MEC, 2011.

_____. *Elementos conceituais e metodológicos para a definição dos direitos de aprendizagem e desenvolvimento do ciclo de alfabetização (1º, 2º e 3º anos) do ensino fundamental*. Brasília: MEC, 2012.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.; FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA. *Aprova Brasil*: o direito de aprender. Brasília: Unicef, 2007.

BROOKE, Nigel; CUNHA, Maria Amália A.; FALEIROS, Matheus. *A avaliação externa como instrumento da gestão educacional dos estados*. São Paulo: Game/FAE/UFMG, 2011.

FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA; BRASIL. Ministério da Educação; UNIÃO DOS DIRIGENTES MUNICIPAIS DE EDUCAÇÃO. *Redes de aprendizagem*: boas práticas de municípios que garantem o direito de aprender. Brasília: Unicef, 2008.

GATTI, Bernadete A. Avaliação educacional: pontuando uma história de ações. *EccoS revista científica*, São Paulo, v. 4, n. 1, p. 17-41, jun. 2002.

HANUSHECK, Eric A.; KAIN, Jonh F.; RIVKIN, Steven G. Do higher salaries buy better teachers? *National Bureau of Economic Research*, Washington, 1999.

HORTA NETO, João Luiz. Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema. *Revista Brasileira de Estudos Pedagógicos*, Brasília, v. 91, n. 227, p. 84-104, jan./abr. 2010.

_____. *As avaliações externas e seus efeitos sobre as políticas educacionais*: uma análise comparada entre a União e os Estados de Minas Gerais e São Paulo. 2013. Tese (Doutorado em Política Social) – Instituto de Ciências Humanas, Departamento de Serviço Social, Brasília. 2013.

INGERSOLL, Richard M. Teacher turnover and teacher shortages: an organization analysis. *American Educational Research Jounal*, Washington, DC, v. 38, n. 3, p. 499-534, 2001.

KORTE, Robert C.; NASH, Beverly E. *Review of the Sistema Nacional de Avaliação da Educação Básica* (Saeb). Brasilia: Inep, 1995.

LUCIO, Elizabeth O. *Tecendo os fios da rede*: o Programa Pró-Letramento e a tutoria na formação continuada de professores alfabetizadores da educação básica. 2010. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

OLIVEIRA, Romualdo P.; ARAÚJO, G. C. Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. *Revista Brasileira de Educação*, São Paulo, n. 28, p. 13-35, jan./abr. 2005.

RAVITCH, Diane. *The death and life of the graet american school system*: how testing and choice are undermining education. Nova York: Basic Books, 2010.

_____. *Reign of error*: the hoax of the privatization movement and the danger to america's public schools. Nova York: Alfred A. Knopf, 2013.

SAHLBERG, Pasi. *Finnish lessons*: what can the world learn from educactional change in Finland? Nova York: Teacher College Press, 2011.

SOARES, Magda B. Por uma alfabetização até os oito anos de idade. In: TODOS PELA EDUCAÇÃO. *De olho nas metas 2010*. São Paulo: Todos pela Educação, 2010. p. 14-17.

VIANNA, Heraldo M. *Avaliação do rendimento de alunos de escolas da rede pública*: um estudo em 15 capitais e 24 outras cidades. São Paulo: Fundação Carlos Chagas, 1990.

JOÃO LUIZ HORTA NETO
Doutor em Política Social e Mestre em Educação pela Universidade de Brasília (UnB). Pesquisador do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)
jlhorta@gmail.com

Recebido em: OUTUBRO 2013
Aprovado para publicação em: SETEMBRO 2014

NORMATIZAÇÕES FEDERAIS E A OFERTA DE MATRÍCULAS EM CRECHES NO BRASIL

MARIA DO CARMO MEIRELLES TOLEDO CRUZ
MARTA FERREIRA SANTOS FARAH
NATASHA BORGES SUGIYAMA

RESUMO

Este artigo analisa os efeitos dos atos normativos da União para apoiar municípios na oferta de vagas em creches. Mostra evoluções incrementais, impacto nas matrículas (2001-2013) e variação considerável entre os estados. Identifica que parte da regulamentação federal é uma forma de coordenação federativa, como o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb), e que parte são programas que contribuem na ampliação de vagas, como a Ação Brasil Carinhoso. Este artigo também aponta que a meta estabelecida pelo Plano Nacional de Educação, de atendimento de 50% das crianças em creches, está longe de ser cumprida.

PALAVRAS-CHAVE POLÍTICAS PÚBLICAS • FEDERALISMO • COOPERAÇÃO INTERFEDERATIVA • EDUCAÇÃO INFANTIL.

* Este artigo é a reformulação dos trabalhos apresentados por Cruz e Farah (2014a; 2014b) no *2014 Congress of IIAS Program*, realizado em Ifrane, Marrocos, de 13 a 17 de junho de 2014, e no *XXXVIII Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração – (EnANPAD) 2014*, ocorrido no Rio de Janeiro, de 14 a 17 de setembro de 2014. Contou com contribuições de Ciro Biderman, professor da Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas (FGV-EAESP); de Catherine R. Merchán, Fernanda Becker e Thomaz Barbosa, pesquisadores e alunos do programa de mestrado e doutorado em Administração Pública e Governo da FGV-EAESP; de Fátima Fernandes de Araújo, técnica do Centro de Estudos e Pesquisas de Administração Municipal (Cepam); de Dayane Fernandes, aluna de Ciências Sociais da Universidade de São Paulo (USP); e de pareceristas anônimos desta revista.

RESUMEN

Este artículo analiza los efectos de los actos normativos del gobierno federal para apoyar a los municipios en la oferta de plazas en guarderías. Muestra evoluciones incrementales, impacto en las matrículas (2001-2013) y considerable variación entre los estados. Identifica que parte de la reglamentación federal es una forma de coordinación federativa, como el Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação *(Fundeb), y que parte son programas que contribuyen para ampliar el número de plazas, como* Ação Brasil Carinhoso. *Este artículo también señala que la meta establecida por el* Plano Nacional de Educação, *de atender al 50% de los niños en guarderías, está lejos de llevarse a cabo.*

PALABRAS CLAVE POLÍTICAS PÚBLICAS • FEDERALISMO • COOPERACIÓN INTERFEDERATIVA • EDUCACIÓN INFANTIL.

ABSTRACT

This article analyzes the effects of the Union's normative acts to support municipalities in supplying places in day care centers. It shows incremental evolution, impact on enrollment (2001-2013) and a high disparity among the states. It identifies that part of the federal regulations consist of a form of federal coordination, such as the Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação *(Fundeb), and another part consists of programs that contribute to the expansion of openings, such as the* Ação Brasil Carinhoso. *The paper also points out that the goal to attend 50% of the children in day care centers, established by the P*lano Nacional de Educação, *is far from being reached.*

KEYWORDS PUBLIC POLICIES • FEDERALISM • INTER-FEDERAL COOPERATION • CHILDREN EDUCATION.

INTRODUÇÃO

A Constituição Federal (CF) de 1988 (BRASIL, 1988) definiu um padrão de relações intergovernamentais, novo na história da república brasileira, com forte diretriz pró-descentralização, e o estabelecimento de um rol significativo de competências compartilhadas entre as esferas de governo (FRANZESE, 2010). O Sistema Brasileiro de Proteção Social foi redesenhado (ARRETCHE, 1999) e estão sendo implementados programas de descentralização de boa parte das políticas sociais (assistência social, saúde, ensino fundamental, entre outras), transferindo, "paulatinamente, um conjunto significativo de atribuições de gestão para os níveis estadual e municipal de governo" (p. 111).

O processo de descentralização não foi igual para todas as políticas sociais e localidades. No caso da educação, em alguns estados já havia uma oferta subnacional, e, em outros locais, transferiram-se atribuições do ensino fundamental para os municípios com processos diferenciados de municipalização (ARRETCHE, 1999, 2012).

Este trabalho se propõe a analisar a política de educação infantil, mais especificamente para a etapa de creche, e as

medidas que a União vem adotando para apoiar e induzir os governos locais a ofertarem vagas em creches. O objetivo é identificar se os atos normativos[1] da União pós-1988 impactaram a oferta de vagas em creches.

O artigo destaca a importância das normatizações federais na garantia das funções reguladora e equalizadora da União, discutindo em que medida normas, diretrizes, instrumentos institucionais e programas federais garantem certa uniformidade e parâmetros gerais nos processos de implementação de políticas públicas, bem como asseguram que os objetivos das políticas sejam mantidos levando em conta as especificidades e capacidades institucionais de cada localidade (ALMEIDA, 2000; ABRUCIO, 2005; BICHIR, 2011).

Nos últimos anos tem havido uma ampliação do interesse teórico e político pela coordenação federativa e pela (re)centralização (ABRUCIO, 2005; ALMEIDA, 2005; ARRETCHE, 1996, 2012; BICHIR, 2011, 2014; FRANZESE, 2010; LÍCIO, 2012; FARAH, 2012, 2013). As transformações nas diversas políticas setoriais e seus deslocamentos (do governo central para os governos subnacionais, e destes para o governo central) vêm merecendo atenção especial nas literaturas nacional e internacional (RICH; GÓMEZ, 2012; EATON; DICKOVICK, 2004).

Para alguns teóricos, há um processo recente de recentralização das políticas sociais; para outros, trata-se de um processo de coordenação federativa cujo objetivo é reduzir as desigualdades inter-regionais e intermunicipais. Na educação infantil e, em especial, na política de creche, no entanto, o debate é recente. Este artigo pretende contribuir para essa reflexão ainda incipiente.

Este trabalho identifica as diferentes competências dos entes federados no setor de educação, apresenta os principais atos normativos federais com impacto na oferta de creches e mostra como mecanismos de equalização e outros instrumentos têm alterado de forma incremental (LINDBLOM, 1959) a política de creches, aumentando a oferta de vagas. Realiza também uma análise das matrículas de 2001 a 2013 e sua evolução no decorrer do período analisado, e finaliza mostrando que está em curso um processo incremental de construção de ações coordenadas visando à expansão de

1 Atos normativos compreendem a CF, Emendas Constitucionais, leis, resoluções, programas e outras ações da União para apoiar os governos locais na oferta de vagas em creches.

vagas em creches. A própria implementação da política indica avanços e dificuldades, o que, por sua vez, gera a promulgação de novos atos normativos que apontam para novos avanços.

A ESCOLHA DA POLÍTICA DE CRECHE E SUA EVOLUÇÃO RECENTE

A política de creche foi escolhida para análise pela importância que tem assumido recentemente na agenda societária e governamental. Sua implementação exige um arranjo complexo que envolve a cooperação interfederativa (União, estados e municípios) e a participação de atores não governamentais (família, entidades, etc.). Além disso, a política de creche envolve um serviço complexo que se deveria articular com outras políticas, como as de assistência social, de gênero, de emprego e, sobretudo, de atenção à criança (saúde, esportes, lazer, cultura).[2]

2 A integração das políticas para a infância, no Brasil, ainda é incipiente, pouco explorada e controversa entre profissionais e movimentos feministas. A discussão vem sendo ampliada nesta década, mas ainda há um longo caminho a ser percorrido para uma visão transversal da política.

Apesar de avanços nos últimos anos, o acesso ainda é restrito: 24% das crianças com até três anos estavam em creche em 2010 (ALVES; SILVA, 2013). A normatização e outros instrumentos de ação federal podem exercer papel importante na expansão, como ocorreu no ensino fundamental, contribuindo ao mesmo tempo para a ampliação do acesso e redução das desigualdades do serviço prestado.

Para entender a política de creches e sua evolução recente, é importante identificar sua trajetória anterior (legado de políticas prévias), incluindo as mudanças ocorridas nos instrumentos normativos e na engenharia operacional dessa política (ARRETCHE, 1999).

Deve-se relembrar que as primeiras unidades criadas no Brasil, no final do século XIX, estavam vinculadas a uma abordagem assistencialista, sendo atreladas a uma política de trabalho que visava a fornecer proteção aos filhos das mulheres de classes trabalhadoras (ROSEMBERG, 1984). As creches, portanto, não surgiram como política social de garantia de direitos de cidadania, mas como apoio às mulheres para lhes permitir o ingresso ao mercado de trabalho (MARCONDES, 2013).

Mesmo com a Consolidação das Leis do Trabalho (CLT), criada em 1943, a oferta do serviço de creche era entendida

como garantia da amamentação, não sendo concebida como educação infantil (ROSEMBERG, 1984), e havia pouca oferta de vagas.

Na década de 1970 e início da de 1980, movimentos políticos, sindicais e de profissionais da educação uniram suas pautas reivindicatórias, e a educação passa a fazer parte de demandas comuns a esses grupos. A articulação de diversos atores no processo de resistência contra a ditadura teve papel decisivo no processo constituinte de 1988 e, mais especificamente, nas políticas para a infância, na revisão de seu significado (ROSEMBERG, 1994), na formulação de uma nova concepção a respeito da criança (e não mais do "menor"), centrada na noção de direitos. Nesse contexto, a creche passa a ser vista como um dos elementos constitutivos do direito à educação (Art. 208, IV), e não apenas como uma ação assistencialista associada à política do trabalho.

Na CF (e regulamentação subsequente), a educação é indicada como direito de todos e dever do Estado e da família (Art. 205 e 208), devendo ser promovida com a colaboração da sociedade. É reconhecida como um direito social fundamental e também como um direito público subjetivo, e consagra-se a sua universalização.

A creche é então inserida na política educacional:

> A Constituição reconheceu como direito de educação à criança pequena – inédito na legislação brasileira – mediante a garantia de atendimento em creches e pré-escolas, das crianças de 0 a 6 anos de idade (Art. 208); bem como reconheceu como dever do Estado e do sistema público de educação, o atendimento em creches, incorporando-lhe, em decorrência, um caráter educativo, até então reconhecido como, basicamente, assistencial. (ARELARO, 2000, p. 51)

Ao ser definida pela Constituição como um direito da criança e dos trabalhadores urbanos e rurais (Art. 7º, XXV), a prestação do serviço de creche pode ser exigida do Estado.[3]

A CF estabeleceu as competências dos entes federativos em relação à provisão de serviços públicos. Em diversas áreas há sobreposição de atribuições e pouca clareza em relação à

3 A CF prevê meios à efetivação dos direitos com o Ministério Público e o Judiciário. Se o Poder Público não cumpre a sua obrigação, o interessado pode acionar o Judiciário e responsabilizá-lo (CURY; FERREIRA, 2010).

função da União, dos estados e municípios (ALMEIDA, 2005; ARRETCHE, 2012). Na política educacional, há competências comuns e compartilhamento de ações no ensino fundamental – entre estados e municípios (Art. 211, § 2º e §3º) – e no ensino médio – entre os estados e a União. A oferta de creches, por sua vez, é atribuição prioritária municipal, por determinação constitucional, devendo o município contar com cooperação técnica e financeira da União e dos estados.[4] Destaca-se que é dever do Estado a garantia de educação infantil, em creche e pré-escola (Art. 208, IV), mas facultada aos pais a inclusão das crianças em creche.

Os serviços de creche, por sua vez, podem ser oferecidos pela rede pública ou pela rede privada. A oferta pela rede privada se dá por meio de duas modalidades: a) lucrativa, em que a oferta é feita diretamente pelo setor privado, sendo regulada pelo mercado; e b) "sem fins lucrativos", em que os serviços são ofertados por instituições comunitárias, confessionais ou filantrópicas, por meio, em geral, de convênios[5] firmados com o Poder Público. Nesse caso, as creches conveniadas contribuem com a materialização da política municipal e ampliam a oferta de vagas.[6] A rede privada deve cumprir as normas gerais da educação nacional, e o Poder Público é responsável por sua autorização e avaliação de qualidade (Art. 209).

A incorporação da creche à área de educação a partir da CF, com base em uma concepção centrada na noção de direitos, colocou a creche na agenda governamental (KINGDON, 2006). Como apontam diversos autores (KUGELMAS; SOLA, 1999; RODDEN, 2005), a descentralização não significa que os níveis de governo mais abrangentes deixem de exercer um papel em relação às políticas descentralizadas. No caso das creches, especialmente, a União exerce um papel central no processo de constituição dessa política. Normas e ações federais passaram, assim, a apoiar o município na ampliação da oferta de vagas.

4 Os estados e a União não estão vedados pela CF de ofertarem vagas em creche.

5 O novo marco regulatório das organizações da sociedade civil, regulamentado pela Lei federal n. 13.019/2014 (BRASIL, 2014b), estabelece os mecanismos de parceria dos entes federados com o terceiro setor. A partir de sua implementação, prevista para julho de 2015, os municípios não mais firmarão convênios, mas sim termos de colaboração e de fomento com essas entidades. Estabelece regras para seleção das entidades e para etapas de execução, monitoramento e avaliação das parcerias. Há a exigência de chamamento público obrigatório, três anos de existência e experiência, entre outras mudanças. Os órgãos municipais deverão planejar e acompanhar a parceria buscando o controle dos resultados previstos.

6 Desde o final da década de 1970, há um debate na literatura sobre a relação do Estado com a sociedade civil organizada e a redefinição de esfera pública (BRESSER-PEREIRA, 1999; SUBIRATS et al., 2012; FARAH, 2001). A discussão sobre os limites e potencialidades da atuação dos municípios na oferta de vagas em creches, por meio de convênios, está inserida nesse contexto. Há atores que defendem a oferta exclusiva de vagas em creche pelo poder público. Na Conferência Nacional de Educação (Conae) de 2010, foi deliberada a "ampliação da oferta de educação infantil pelo poder público, extinguindo progressivamente o atendimento por meio de instituições conveniadas" (CONFERÊNCIA NACIONAL DE EDUCAÇAO, 2010, p. 68). Entretanto, outros avaliam que, neste momento, é necessário coexistirem a rede pública e conveniada, mesmo com as disparidades na oferta de serviços, para garantir a ampliação do acesso ao serviço. No presente artigo não será tratada a diferença de qualidade entre essas redes.

ATOS NORMATIVOS DA EDUCAÇÃO INFANTIL

A CF estabelece, como dever do Estado, a garantia do acesso à educação infantil, em creche e pré-escola. A creche é vista, portanto, como um direito da criança e também dos trabalhadores urbanos e rurais, e se torna um *locus* de educação, e não mais de assistência social.

À época da promulgação da Constituição, não estavam claras as atribuições do município em relação à creche; ele se responsabilizava apenas pelos programas de pré-escola, além da oferta do ensino fundamental (Art. 211, § 2º). No momento de proposição e homologação da CF, houve uma priorização da universalização do ensino fundamental, então ainda um grande desafio (FARAH, 1994). A educação infantil passou a receber maior atenção dos atores societários e governamentais mais recentemente.

Com a nova Constituição, ficou estabelecido que a União, os estados, o Distrito Federal e os municípios organizariam em regime de colaboração de seus sistemas de ensino (Art. 211), com suas respectivas redes, mas naquele momento não se criou um sistema único como ocorreu na política de saúde.

A CF definiu, ainda, que a União deveria empregar, no mínimo, 18% das receitas resultantes de impostos e transferências constitucionais na manutenção e no desenvolvimento do ensino. Os estados, o Distrito Federal e os municípios, por sua vez, precisam aplicar, no mínimo, 25% (Art. 212) da receita resultante de impostos, compreendida a proveniente de transferências, na manutenção e no desenvolvimento do ensino.[7]

7 Pelo menos 50% dos recursos dos mínimos estabelecidos para a educação devem ser utilizados para assegurar a universalização do atendimento e remuneração do magistério no ensino fundamental (Art. 60 do Ato das Disposições Constitucionais Transitórias).

A União passou a formular e implementar a política educacional nacional, a elaborar o Plano Nacional de Educação (PNE) e a articular os diferentes níveis e sistemas (municipais, estaduais e nacional), exercendo funções normativa, redistributiva e supletiva, em relação às demais instâncias educacionais. Os estados formulam e implementam políticas e planos educacionais próprios, em consonância com as diretrizes nacionais, e coordenam suas ações e as de seus municípios, entre outros atos.

O Estatuto da Criança e do Adolescente (ECA) (BRASIL, 1990) e, posteriormente, a Lei de Diretrizes e Bases da

Educação (LDB) (BRASIL, 1996b) tornam claro que compete aos municípios garantir a materialização do acesso à creche, à pré-escola e ao ensino fundamental. Tais atos normativos incorporaram, portanto, a mudança de concepção sobre as creches – agora entendidas como um serviço educacional.

O ECA e a LDB explicitam uma nova concepção da infância e da adolescência, centrada na noção de direitos, em oposição à que prevalecera por várias décadas no país, de caráter assistencialista-repressivo (FERRAREZI, 1994). A perspectiva de direitos que, a partir da Constituição, passa a orientar diversas políticas sociais (assistência social, saúde, educação, dentre outras) estende-se também às políticas para esse segmento da população (FARAH, 2001).

No ECA, é estabelecido o sistema de garantia dos direitos da criança e do adolescente envolvendo o poder judiciário, Ministério Público, Defensoria Pública, segurança pública, conselhos de direitos das crianças e dos adolescentes e tutelares, entre outros atores. No que diz respeito à creche, o ECA estabelece que é dever do Estado garantir à criança o atendimento em creche (Art. 54, IV), definindo ações de responsabilidade quando houver ação ou omissão de um ator do sistema (Art. 208, III). Assim, quando o direito à creche não é garantido, é possível aos interessados o questionamento judicial (CURY; FERREIRA, 2010). Entretanto, o ECA, na sua promulgação, não explicitou a responsabilidade do município na oferta de creche.

A Emenda Constitucional (EC) 14/1996 (BRASIL, 1996a), regulamentada pela Lei federal n. 9.424/1996 (BRASIL, 1996c), estabelece explicitamente as atribuições dos entes federados, definindo que os municípios atuarão prioritariamente no ensino fundamental e na educação infantil. Essa emenda modifica o financiamento da educação, criando um fundo de natureza contábil destinado à manutenção do ensino fundamental: o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef).

A inovação da EC 14/1996 consistiu na mudança da estrutura de financiamento do ensino fundamental público, que passou a vincular 15% dos recursos da educação a essa etapa de ensino, com distribuição automática de recursos,

de acordo com o número de alunos matriculados em cada rede de ensino fundamental, promovendo a partilha de responsabilidades entre os governos estaduais e respectivos governos municipais (BRASIL, 2004; CARVALHO, 2014).

O Fundef não financiava a educação infantil e o ensino médio. Já era prevista, na Constituição, a vinculação de 25% das receitas dos estados e municípios à educação. Com essa emenda, 60% desses recursos (o que representa 15% da arrecadação global de estados e municípios) ficavam reservados ao ensino fundamental. Os municípios poderiam utilizar os 10% restantes na educação infantil. O fundo deveria ter 60% de seus recursos reservados para a remuneração dos profissionais do magistério em exercício, e 40% para a manutenção e o desenvolvimento do ensino fundamental. O acompanhamento e o controle social sobre a distribuição, a transferência e a aplicação dos recursos eram exercidos por conselhos instituídos especificamente para esse fim no âmbito estadual, do Distrito Federal e municipal. Além do controle social, é previsto o controle externo por meio dos tribunais de contas e a fiscalização das câmaras municipais.

A proposta era equalizar a atuação no ensino municipal, muito desigual nos diversos estados e municípios. O Fundef promoveu essa equalização (LACZYNSKI, 2003), a expansão da matrícula no ensino fundamental (AFONSO, 2004) e o aumento dos salários dos professores; a destinação de recursos funcionou como um estímulo à expansão desse nível de ensino e criou a obrigatoriedade de destinação de parte dos recursos recebidos para a melhoria da remuneração dos professores. Entretanto, restringiu a autonomia de estados e de municípios, pois focalizou a maior parte dos recursos no ensino fundamental (ARRETCHE, 2012; CARVALHO, 2014). Havia uma clara indução federal por meio do Fundef, estimulando o investimento na educação fundamental que não contemplava a creche.

A LDB, aprovada no mesmo ano da EC 14/1996, por sua vez, estabelece o ensino infantil até seis anos como a primeira etapa da educação básica[8] (Art. 21, I). Inclui-se, de forma explícita, a creche como atribuição municipal, devendo a União e os estados apoiarem o município técnica e financeiramente.

8 Posteriormente, a faixa etária de abrangência da educação infantil foi modificada, envolvendo crianças até cinco anos.

Mostra, assim, que a educação na primeira infância passa a ser reconhecida. Entretanto, a LDB determina ainda que a prioridade de atuação municipal deve ser o ensino fundamental (Art. 11, V).

Fica definido que os sistemas municipais compreendem:

> [...] as instituições do ensino fundamental, médio e de *educação infantil* mantidas pelo Poder Público municipal; as instituições de educação infantil criadas e mantidas pela iniciativa privada; e os órgãos municipais de educação (LDB, Art. 18, itálicos nossos).

O sistema municipal envolve, portanto, todas as instituições públicas ou privadas que ofertam vagas na educação básica e os órgãos gestores da rede.

No caso das creches, pode ocorrer a participação da iniciativa privada, por meio de instituições sem fins lucrativos (entidades comunitárias, filantrópicas e confessionais), ou por meio da oferta direta no mercado. Neste caso de creches privadas lucrativas, o serviço é pago diretamente pelo cidadão.

A oferta indireta de vagas pelo Poder Público pode dar-se por meio de convênios com entidades filantrópicas, comunitárias e confessionais. Nas creches públicas ou nas conveniadas, o serviço é gratuito, assim como nos demais níveis de ensino, da educação infantil à superior.

A Lei de Responsabilidade Fiscal (LRF) (BRASIL, 2000) também tem impacto sobre a ação municipal em relação às creches. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal. Dentre as mudanças introduzidas por essa lei, destaca-se a limitação da despesa total com pessoal a 54% da receita corrente líquida no Poder Executivo municipal. As creches são intensivas em mão de obra, o que impõe um freio à sua expansão, dados os limites estabelecidos pela LRF.

Outro ato normativo que influenciou a atuação municipal em creche foi o Plano Nacional de Educação (PNE) de 2001 a 2010 (BRASIL, 2001), que determinava que os municípios atendessem em creche 50% das crianças com até três anos até 2010, estabelecendo padrões mínimos de infraestrutura para o funcionamento adequado das instituições públicas e privadas de educação infantil e um Programa Nacional de Formação dos Profissionais de educação infantil, determinando

formação mínima para os profissionais da educação. Preconizava a universalização da pré-escola até 2016.

O PNE trouxe ainda uma mudança importante para a política de creches: determinou a inclusão das creches (ou entidades equivalentes) no sistema nacional de estatísticas educacionais. Definia que os planos plurianuais dos entes federados fossem elaborados de forma a contemplar as metas do PNE e dos respectivos planos decenais (Art. 5º). O Plano estabelecia que os municípios deveriam priorizar a educação infantil na aplicação dos 10% dos recursos vinculados à manutenção e desenvolvimento do ensino não reservados para o ensino fundamental (Item 11.3.8), mas não previa um instrumento que permitisse a colaboração da União e dos estados no cofinanciamento.

Novas legislações também influenciaram a educação infantil. A Lei federal n. 11.114/2005 (BRASIL, 2005) tornou obrigatória a matrícula no ensino fundamental aos seis anos de idade, e a Lei federal n. 11.274/2006 (BRASIL, 2006b) instituiu o ensino fundamental de nove anos. Assim, a criança de seis anos migrou da educação infantil para o ensino fundamental, possibilitando a abertura de novas vagas na pré-escola e na creche.

A EC 53/2006 (BRASIL, 2006a) altera a CF, definindo que o dever do Estado com a educação será efetivado mediante a garantia de "educação infantil, em creche e pré-escola, às crianças até cinco anos de idade" (Art. 208, IV), incorporando a ideia de direito à creche e de dever do Estado preconizados pelo ECA. Portanto, a sua oferta torna-se obrigação dos municípios.[9]

Essa emenda cria o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb),[10] com regras similares ao Fundef. O financiamento para as creches foi então modificado, ocorrendo uma revisão da destinação do fundo – originalmente destinado exclusivamente ao ensino fundamental –, que passou a abranger a educação básica, incluindo, além do ensino fundamental, a educação infantil e o ensino médio. Há, portanto, um avanço na implementação progressiva do direito à educação, por meio do financiamento da educação infantil,

9 Há questionamentos referentes à obrigatoriedade da oferta de vagas em creche e esses têm gerado manifestações distintas por parte do Ministério Público e Judiciário. A maioria dos juízes paulistas (SILVEIRA, 2011) entende a educação infantil como uma obrigação constitucional que deve garantir o "efetivo acesso e atendimento em creches e unidades de pré-escola, sob pena de configurar-se inaceitável omissão governamental, apta a frustrar, injustamente, por inércia, o integral adimplemento, pelo poder público" (CURY; FERREIRA, 2010, p. 87). O Supremo Tribunal de Justiça (STJ) também se tem manifestado seguindo essa linha. Um exemplo é o voto da Ministra Ellen Grace, que aponta a educação infantil como prerrogativa constitucional indisponível, e que o Estado "deve criar condições objetivas que possibilitem o acesso à educação infantil" (BRASIL, 2009b). As autoras deste artigo entendem que não há um questionamento sobre o reconhecimento do direito à creche e sim a atribuição de responsabilidade pela materialização do mesmo. Portanto, defendem a obrigatoriedade de oferta de vagas pelo município, pois a creche está assegurada como um direito da criança e da mulher trabalhadora e um dever do Estado.

10 O Fundeb é regulamentado pela Lei federal n. 11.494/2007 (BRASIL, 2007a) e Decreto n. 6.253/2007 (BRASIL, 2007b).

incorporando de forma incremental as matrículas dessa etapa de ensino.

Outras mudanças trazidas pela EC 53/2006 (BRASIL, 2006a), que criou o Fundeb, tiveram impacto na política de creches: a subvinculação das receitas dos impostos e transferências dos estados, do Distrito Federal e dos municípios foi elevada de 15% para 20%; as obrigações de gasto do governo federal foram aumentadas (ARRETCHE, 2012); garantiram-se as cooperações técnica e financeira da União e dos estados nos programas de educação infantil e ensino fundamental, de responsabilidade dos municípios.

O Fundeb também busca equalizar nacionalmente a educação: a União apoia os estados e municípios, que não alcançam, com a própria arrecadação, o valor mínimo nacional estabelecido por aluno. Novas alterações nas relações interfederativas são introduzidas no setor da educação, com repercussões na oferta de creches. Segundo Carvalho (2014), o Fundeb propicia uma política indutora, ou seja, promove a especialização de cada ente federado em sua área de atuação prioritária, prevista na CF, reduzindo a autonomia de estados e municípios na oferta educacional, uma vez que a atuação em outras áreas acarreta custos sem repasses adicionais do fundo.

Em 2007, como parte das ações definidas pelo Plano de Desenvolvimento da Educação (PDE), então formulado, foi criado o Programa Nacional de Reestruturação e Aquisição de Equipamentos para a Rede Escolar Pública de Educação Infantil (Proinfância) que financia a construção de creches e pré-escolas, bem como a aquisição de equipamentos pelos municípios (FUNDO NACIONAL DE DESENVOLVIMENTO DA EDUCAÇÃO, s.d). O governo disponibiliza, assim, uma fonte de recursos para investimento em creches, complementando o incentivo dado pelo Fundeb, o qual financia apenas o custeio, por meio de matrículas, e não a expansão da rede (MARCONDES, 2013). Os municípios que aderirem ao programa devem optar pelos modelos construtivos propostos pelo programa e não podem fazer os seus próprios.

O Plano de Metas Compromisso Todos pela Educação, programa estratégico do PDE, por sua vez, estabeleceu a

obrigatoriedade de formulação de Planos de Ações Articuladas (PAR). Os municípios deveriam formular planos que contemplassem os níveis de educação sob sua responsabilidade: o ensino fundamental e a educação infantil.

Em 2008, por meio da Lei federal n. 11.738/2008 (BRASIL, 2008), foi instituído um piso salarial nacional para os profissionais do magistério público da educação básica, abrangendo os profissionais de creche. A União complementa os valores quando o ente federativo não tem disponibilidade orçamentária para cumprir o valor fixado. O objetivo é proporcionar melhoria da qualidade da educação por meio da valorização do professor, orientação já existente no Fundef e Fundeb.

Com a EC 59/2009 (BRASIL, 2009), torna-se obrigatório o ensino dos 4 aos 17 anos, a ser implementado até 2016, com apoios técnico e financeiro da União. Essa obrigatoriedade coexiste com a creche como dever do Estado (Art. 208, IV) e, portanto, um direito já previsto no ECA desde 1990. A emenda em questão coloca como obrigatória a frequência das crianças a partir da pré-escola (Art. 208, I), facultando aos pais a sua inclusão em creche.

Essa emenda também reduz a lacuna da coordenação interfederativa, estabelecendo que o plano nacional de educação, de duração decenal, deve ser estruturado com o "objetivo de articular o sistema nacional de educação em regime de colaboração e definir diretrizes, objetivos, metas e estratégias de implementação para assegurar a manutenção e desenvolvimento do ensino em seus diversos níveis, etapas e modalidades" (Art. 4º). Amplia, ainda, a abrangência dos programas suplementares (material didático escolar, transporte, alimentação e assistência à saúde), para todas as etapas da educação básica, o que inclui as creches.

Em 2010, o Proinfância é incluído no Programa de Aceleração ao Crescimento (PAC-2), no eixo Comunidade Cidadã, indicando a inclusão da creche na agenda prioritária no governo federal (MARCONDES, 2013).

A partir de 2011, iniciam-se novas ações da União para apoiar os municípios na ampliação da oferta de vagas em creches. A Lei federal n. 12.499/2011 (BRASIL, 2011b) prevê

apoio financeiro da União aos municípios que criarem novos estabelecimentos públicos de educação infantil construídos com recursos de programas federais (Art. 1º, I). Não está previsto apoio para ações próprias, municipais ou estaduais (construção, com recursos próprios, ou estaduais, de novos estabelecimentos; a reestruturação da rede para ampliar o número de vagas sem abertura de novas turmas; a melhoria da qualidade na prestação desse serviço etc.). Estabelece, ainda, que esses recursos devem ser aplicados exclusivamente em despesas correntes para a manutenção e o desenvolvimento da educação infantil pública.

Nova indução ocorre em 2012, por meio da Ação Brasil Carinhoso, do Plano Brasil Sem Miséria (BRASIL, [201-]), que objetiva retirar da extrema pobreza todas as famílias com crianças até seis anos e, ao mesmo tempo, aumentar o acesso dessas famílias a creches e serviços de saúde. Com essa finalidade, a União modifica a lei que cria o Programa Bolsa Família (PBF), ampliando o benefício de transferência de renda às famílias com crianças de até seis anos; garantindo apoio financeiro às localidades que ampliam a oferta de vagas em creche e aumentando o valor do repasse para a alimentação escolar[11] (BRASIL, [201-]).

11 Com o Brasil Carinhoso, o valor repassado pela União para a alimentação escolar passa de R$ 0,60, para cinco refeições, para R$ 1,00. O valor é baixo, se considerados os custos. O Município de São Paulo gasta, em média, R$ 6,00 por criança/dia (SÃO PAULO, 2013).

Com o Brasil Carinhoso, amplia-se o repasse do Fundeb em 50%, para aqueles que criam novas vagas em creches para crianças de até quatro anos beneficiárias do PBF em rede municipal própria ou conveniada com o Poder Público. Além de induzir a ação dos municípios na oferta de creches, o governo federal prioriza o público-alvo a ser atendido, focalizando as crianças mais pobres (BRASIL, [201-]). Esse programa foi estabelecido após o diagnóstico de que 14,6% das crianças com até seis anos estavam em situação de extrema pobreza (INSTITUTO DE PESQUISA ECONÔMICA APLICADA, 2013). O adicional (50%) é repassado pelo Ministério de Desenvolvimento Social e Combate à Fome (MDS) por ser uma ação ligada à política de enfrentamento à pobreza.

A Lei federal n. 12.695/2012 (BRASIL, 2012a), por sua vez, permite que, na distribuição dos recursos do Fundeb, sejam computadas todas as novas vagas oferecidas pela rede pública e pelas instituições comunitárias, confessionais ou

filantrópicas sem fins lucrativos, conveniadas com o Poder Público (Art. 13). Amplia-se, em decorrência, o financiamento a essas entidades, uma vez que, até então, os recursos do Fundeb deveriam ser destinados exclusivamente aos serviços oferecidos diretamente pela rede pública.

A Lei federal n. 12.722/2012 (BRASIL, 2012b) expande ainda mais o escopo do custeio do Fundeb, permitindo aplicação de recursos no cuidado integral e na segurança alimentar e nutricional das crianças, além das despesas já previstas com manutenção e desenvolvimento da creche (Art. 4º, § 4º).

A ampliação da abrangência do custeio é demanda dos municípios, que consideram alto o custo da alimentação escolar e do cuidado integral, se considerados os valores repassados para o atendimento das crianças em creche (CONFEDERAÇÃO NACIONAL DOS MUNICÍPIOS, 2013). Segundo a Confederação Nacional dos Municípios (CNM), os programas federais são subfinanciados, e o recurso repassado não cobre os gastos efetivos como é o caso, por exemplo, do Programa Nacional de Alimentação Escolar (Pnae). A suplementação do valor *per capita*, em 50%, e a permissão de utilização com alimentação escolar e com o cuidado às crianças, observa-se uma "escuta" do governo federal à demanda de municípios que reivindicavam mais apoio para custeio. Entretanto, essas modificações atendem apenas crianças do PBF, e não todas as matriculadas em creche.

Outra modificação criada para superar a falta de creches consiste na agilização do processo licitatório dos edifícios para sediar esses equipamentos. A Lei federal n. 12.722/2012 (BRASIL, 2012b) altera a Lei n. 12.462/2011 (BRASIL, 2011a), possibilitando que os sistemas públicos de ensino realizem licitações e contratos pelo Regime Diferenciado de Contratações Públicas (RDC).[12] Com essa forma de contratação, objetiva-se agilizar o processo de licitação e de construção desses equipamentos. Procurou-se, assim, contornar um dos obstáculos à expansão da rede de creches: a morosidade dos processos licitatórios do setor público.

A Lei federal n. 13.005/2014 (BRASIL, 2014a) aprova o novo PNE 2014-2024. Permanece a meta de ampliar a oferta de educação infantil em creches, de forma a atender no

12 O processo construção de creche é lento e complexo. Exige a definição de terreno; elaboração de projetos arquitetônico, de engenharia e executivos e licitação para construção. O RDC é criado para agilizar os processos licitatórios previstos na Lei federal n. 8.666/1993 (BRASIL, 1993), e pretende tornar mais eficientes e econômicas as contratações públicas da Copa das Confederações de 2013, Copa do Mundo de 2014 e Jogos Olímpicos e Paraolímpicos de 2016. Posteriormente, é ampliado para ações do PAC, incluindo obras e serviços de engenharia dos sistemas públicos de ensino, do Sistema Único de Saúde, e para estabelecimentos penais e unidades de atendimento socioeducativo.
Busca melhorar a relação entre custo e benefício para a Administração Pública. Para Fiuza (2012), as vantagens dessa legislação são a inversão das fases (julgamento das propostas precede a habilitação); a contratação integrada, que transfere para o contratado a responsabilidade pela elaboração dos projetos das obras; o cadastramento de fornecedores e a remuneração variável; entre outras. O contratado pode assumir a execução de todas as etapas da obra, mas fica responsável pelos riscos. Essas contratações devem respeitar normas especiais de proteção ao meio ambiente. Outras vantagens são: fase recursal única; orçamento estimado para a contratação é publicado após o encerramento da licitação, tendo caráter sigiloso, disponibilizado apenas para os órgãos de controle; pré-qualificação permanente; entre outras. Há a possibilidade da indicação de marca ou modelo, para a aquisição e bens, com a exigência de amostras do produto e certificação de qualidade desde que justificada. Ressalta-se que há uma ação do Ministério Público Federal questionando o Supremo Tribunal Federal sobre a inconstitucionalidade da lei.

13 O PNE 2001-2010 não atendeu às metas definidas para a creche.

mínimo 50% das crianças de até três anos até 2023.[13] A universalização da pré-escola está prevista para até 2016 para as crianças de quatro a cinco anos de idade. A lei inova ao estabelecer que será criada uma instância permanente de negociação e cooperação entre a União, os estados, o Distrito Federal e os municípios, bem como será fortalecido o regime de colaboração entre os estados e seus respectivos municípios, com a inclusão de instâncias permanentes de negociação, cooperação e pactuação em cada estado (Art. 7º, § 5º e 6º). Avança na discussão da qualidade, na orientação para a ampliação do acesso dos mais pobres às creches e no financiamento da educação pública, estabelecendo que, em até dez anos, o Brasil gastará 10% do Produto Interno Bruto (PIB) em educação (meta 20 do PNE), entre outros pontos.

No Quadro 1, são apresentadas as principais ações do governo federal na área de educação que exercem um impacto nas creches. Desde a CF de 1998, a União vem criando atos normativos, mecanismos de financiamento ou programas e ações para induzir os governos locais a desenvolverem ações que ampliem a oferta de vagas em creches e para apoiá-los nessas ações.

QUADRO 1 – Principais atos normativos que influenciaram a política de educação infantil por período

ANTES DE 1988	CONSTITUIÇÃO FEDERAL DE 1988	DÉCADAS DE 1990 E 2000	DE 2001 A 2010	APÓS 2010
– Creches fornecidas majoritariamente por entidades de assistência social.	– Educação infantil integra o sistema de ensino brasileiro. – A criança é considerada prioridade absoluta. – Creche é direito da criança. – Pré-escola e ensino fundamental prioritariamente de competência municipal. – Regras de financiamento. – Recursos vinculados: 18% de recursos orçamentários federais; 25% de estados; 25% de municípios e contribuição compulsória das empresas (salário-educação).	– Educação infantil (creche e pré-escola) e ensino fundamental prioritariamente de competência municipal. – Educação infantil envolve creche (até 3 anos) e pré-escola (até 6 anos). – Criação do Fundef, que não financia a educação infantil. – Despesa total com pessoal não pode exceder 60% da receita corrente líquida dos municípios.	– Educação infantil (creche e pré-escola) e ensino fundamental prioritariamente de competência municipal. – Educação infantil envolve creche (até 3 anos) e pré-escola (até 5 anos). – Educação básica obrigatória e gratuita dos 4 aos 17 anos. – Piso salarial nacional para os profissionais do magistério que trabalhem na educação básica pública. – Criação do Fundeb, que financia a educação infantil e os ensinos fundamental e médio. – Limitação dos gastos de custeio para manutenção e desenvolvimento da educação infantil. – Ampliação da complementação dos recursos da União de 1% para 10%. – Estabelecimento de metas de atendimento da educação infantil pelo PNE: 50% das crianças em creche até 2010. – Limites de gastos com pessoal e de empréstimos.	– Educação infantil (creche e pré-escola) e ensino fundamental prioritariamente de competência municipal. – Educação infantil envolve creche (até 3 anos) e pré-escola (até 5 anos) – Educação básica obrigatória e gratuita dos 4 aos 17 anos de idade. – Fundeb financia a educação infantil e os ensinos fundamental e médio. – Apoio financeiro da União aos novos estabelecimentos públicos de educação infantil. – Ampliação de 50% do Fundeb para novos estabelecimentos e novas vagas que atendam crianças do PBF. – Fundeb inclui matrículas das instituições comunitárias, confessionais ou filantrópicas sem fins lucrativos, conveniadas com o Poder Público. – Recursos de custeio podem ser aplicados com cuidado integral e segurança alimentar e nutricional. – Apoio financeiro da União para construção de creches e aquisição de equipamentos. – Creches podem ser licitadas pelo RDC. – Estabelecimento de metas de atendimento da educação infantil pelo PNE 2014-2024: 50% das crianças em creche até 2023.

Fonte: Elaborado pelas autoras com base nos trabalhos de Almeida (2000, 2005) e Arretche (2012).

Os atos normativos do governo federal sofreram transformações de 1988 a 2013 (Quadro 1). Foram utilizadas estratégias diferenciadas pelo governo federal para promover a coordenação de ações dos governos subnacionais. Promoveram-se normas que: a) restringem a liberdade de gasto dos governos subnacionais; b) definem responsabilidades ou competências dos entes da federação com relação à provisão e à gestão da política de creche; c) criam incentivos para que os governos subnacionais passem a assumir a responsabilidade de prover os serviços de creche (GOMES, 2009) e programas de apoio.

Foram redefinidas as próprias competências e obrigações dos governos municipais em relação à educação infantil, ampliando-se progressivamente a sua responsabilidade na política de creches. São exemplos desses atos o ECA, a LDB e a EC 14/1996. A participação das entidades conveniadas também foi estabelecida no processo, havendo um estímulo a sua participação por meio da permissão de utilização de recursos do Fundeb também para esse segmento da rede,[14] mesmo com a deliberação na Conferência Nacional de Educação (CONAE) de 2010 de priorização da rede pública própria. Esse movimento acompanha o lugar que a creche passa a assumir na política educacional e também no debate sobre os direitos das crianças.

14 Paralelamente a esse movimento de estimular a oferta por vagas em creches conveniadas, a Conae deliberou que as matrículas em creches conveniadas devem ser congeladas em 2014, e essa modalidade de parceria deve ser extinta até 2018, assegurando a demanda por rede própria. A análise do impacto dessa determinação ultrapassa o escopo deste artigo.

Normas, como a Lei de Responsabilidade Fiscal, o Fundef e o Fundeb, restringem a liberdade de gasto dos governos locais e induzem determinadas despesas. A União cria, ainda, mecanismos de equalização e apoia financeiramente (investimentos e custeio) a ação local, para que os municípios possam cumprir suas competências e atender beneficiários de programas federais prioritários. Com o Fundeb, o Brasil Carinhoso, o Proinfância, entre outros, o governo federal criou incentivos para que os governos municipais assumam a sua responsabilidade em relação à oferta de creches.

A seguir, procura-se verificar o impacto das normas federais, em especial do PNE, do Fundeb e do Brasil Carinhoso na expansão de vagas em creche. A análise considera as matrículas em creche de 2001 a 2013 nos 26 estados e no Distrito Federal, a taxa de atendimento de crianças em creche e os resultados do Brasil Carinhoso.

ESTUDO COMPARATIVO DO ACESSO À CRECHE

Os censos escolares, principal fonte de dados sobre a situação do ensino no Brasil, passaram a incluir dados sobre creches apenas em 1998, e os primeiros dados então divulgados sobre o setor são parciais, não incluindo todos os estabelecimentos (SISTEMA EDUCATIVO NACIONAL DO BRASIL, 2003). A partir de 2001, ano de aprovação da PNE 2001-2010, é possível consultar os dados sistematizados no Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), que diferencia creche e pré-escola, nas diversas redes de ensino.

O estudo do impacto da normatização federal nos municípios inicia contemplando as matrículas em creches de 2001 a 2013 por rede de ensino (municipal, estadual, federal e privada) nos estados.

Os dados de matrículas mostram crescimento de 147% entre 2001 e 2013; sendo de 31% entre 2001 e 2006, ano de aprovação do Fundeb, e de 71% entre 2007 e 2013 (Tabela 1).

TABELA 1 – Quantidade de matrículas em creche, por estado e rede de ensino (2001, 2006 e 2013)

ESTADO	2001					2006					2013				
	MUNICIPAL	ESTADUAL	FEDERAL	PRIVADA	TOTAL	MUNICIPAL	ESTADUAL	FEDERAL	PRIVADA	TOTAL	MUNICIPAL	ESTADUAL	FEDERAL	PRIVADA	TOTAL
AC	1.783	171	0	61	2.015	2.582	0	0	219	2.801	6.146	0	0	422	6.725
AL	5.654	142	0	3.007	8.803	8.217	0	0	2.804	11.021	20.435	0	0	6.301	26.736
AP	527	161		1.183	1.871	561	194		841	1.596	2.200	0	0	1.118	3.318
AM	4.927	0	0	2.731	7.658	10.062	18	0	2.582	12.662	16.155	0	0	5.692	21.847
BA	40.189	5.343	0	7.708	53.240	60.864	4.110	0	16.044	81.018	88.574	0	0	28.381	117.005
CE	58.230	254	16	26.386	84.886	67.531	86	16	22.637	90.270	97.445	17	17	37.080	134.678
DF	0	601	33	9.787	10.421	0	610	35	12.667	13.312	0	43	43	23.665	25.274
ES	22.872	0	0	3.672	26.544	30.344	0	81	4.518	34.943	59.812	87	87	6.653	66.552
GO	14.262	91	89	8.920	23.362	22.889	0	53	11.957	34.899	39.261	66	66	16.730	56.057
MA	19.395	0	0	7.495	26.890	29.473	0	0	16.153	45.626	62.809	0	0	22.013	84.822
MT	13.181	323	0	1.882	15.386	21.800	467	0	4.412	26.679	37.206	0	0	7.119	44.701
MS	10.116	1.336	0	3.458	14.910	15.449	1.583	0	5.225	22.257	36.327	0	0	9.651	46.199
MG	35.065	32	0	56.413	91.510	58.068	24	0	65.641	123.733	138.629	0	0	106.830	245.459
PA	26.493	608	118	2.650	29.869	27.516	622	0	3.317	31.455	37.316	0	0	7.585	45.019
PB	12.410	1.961	0	1.507	15.878	15.692	2.022	0	3.118	20.832	26.145	114	114	10.373	36.669
PR	59.318	135	104	31.381	90.938	71.937	6	88	35.014	107.045	122.331	66	66	43.170	165.567
PE	16.726	87	0	15.454	32.267	23.705	349	0	23.570	47.624	38.147	0	0	40.295	78.785
PI	20.151	1.893	0	4.037	26.081	17.816	0	42	4.476	22.334	31.125	0	0	4.501	35.626
RJ	34.089	632	133	43.965	78.819	55.933	331	130	52.662	109.056	115.718	322	322	88.789	204.913
RN	28.541	297	0	6.463	35.301	33.948	649	119	9.941	44.657	37.708	114	114	11.570	49.392
RS	37.756	163	95	22.931	60.945	45.784	173	48	29.213	75.218	81.582	132	132	60.476	142.384
RO	3.147	508	0	1.694	5.349	4.293	461	0	2.163	6.917	9.878	0	0	3.209	13.768
RR	725	1.753	0	462	2.940	1.639	105	0	160	1.904	2.249	0	0	1.874	4.123
SC	40.830	1.136	116	16.082	58.164	58.193	1.517	117	17.274	77.101	110.566	84	84	27.283	137.997
SP	147.389	567	182	128.178	276.316	202.306	4.107	204	160.369	366.986	465.140	341	341	408.691	874.973
SE	3.665	98	0	1.080	4.843	9.087	0	0	3.390	12.477	9.802	0	0	3.792	13.594
TO	6.067	0	0	2.074	8.141	7.809	57	0	2.273	10.139	14.811	0	0	3.058	17.869
Total	663.508	18.292	886	410.661	1.093.347	903.498	17.491	933	512.640	1.434.562	1.707.517	1.386	1.386	986.321	2.700.052

Fonte: Elaboração das autoras com base nos dados do Inep (BRASIL, 2001-2013).

As matrículas federais cresceram, e as estaduais decresceram, mas as duas redes são pouco representativas. O Distrito Federal e os estados do Ceará, Espírito Santo, Paraíba, Rio de Janeiro, Rio Grande do Norte, Rio Grande do Sul e São Paulo registraram incremento das matrículas federais no período. Em Mato Grosso, Pernambuco, Rio Grande do Sul, Rondônia e São Paulo, aumentaram as matrículas estaduais. O pequeno número de matrículas dessas redes sugere que se trata de atendimento a filhos de funcionários federais e estaduais ou de escolas de aplicação, mantidas pelas faculdades de educação de universidades estaduais e federais.

A rede de ensino que mais cresceu entre 2001 e 2013 foi a municipal (157%), seguida pela rede privada (140%).[15] As matrículas municipais representavam 61%, 63% e 63% em 2001, 2006 (ano de aprovação do Fundeb) e 2013, respectivamente. O governo local é o principal provedor de serviços públicos de creche, confirmando o papel que lhe é atribuído pela normatização federal posterior à Constituição. Constata-se, no entanto, que, em que pese o crescimento da rede municipal no período analisado, o percentual que representa não sofre alteração substantiva. A não discriminação na rede privada entre conveniadas e creches ligadas ao setor privado dificulta auferir o incremento da oferta total de responsabilidade do município.

15 Inclui escolas privadas e instituições comunitárias, confessionais ou filantrópicas sem fins lucrativos conveniadas com o Poder Público.

A rede municipal cresceu em todos os estados, exceto no Distrito Federal, que não possui rede municipal, mas prevalecem as desigualdades entre os estados. Há aqueles em que essas matrículas cresceram menos de 50%, como o Rio Grande do Norte (32%), e, em outros, mais de 260%, como Alagoas (261%) e Amapá (317%).

A rede privada também aumentou em todos os estados, exceto do Amapá. O aumento também é desigual. O Piauí registra, no período, crescimento de 11%, e a Paraíba, de 588%.

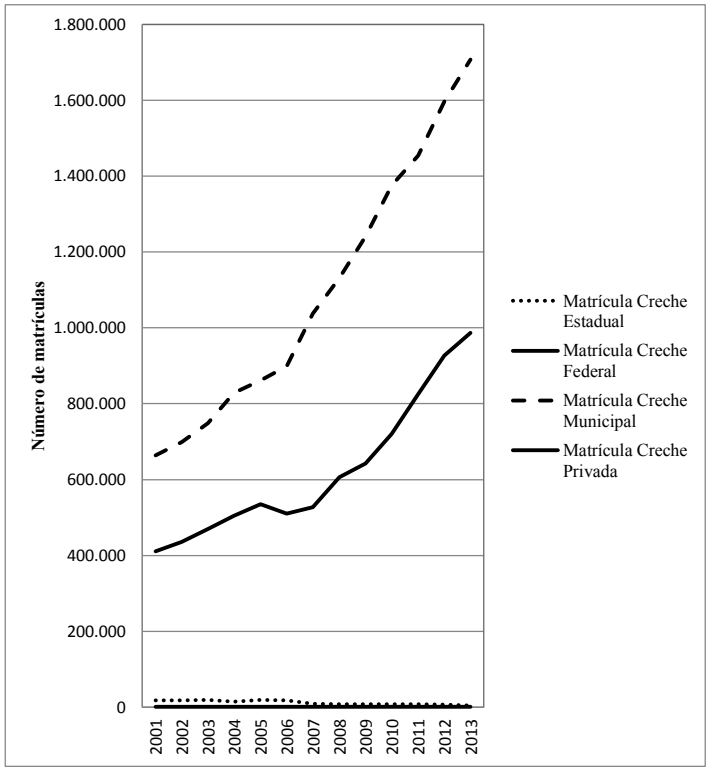
Na maioria dos estados, a rede municipal cresceu proporcionalmente mais do que a privada, exceto nos estados de Mato Grosso, Pará, Paraíba, Pernambuco, Rio Grande do Norte, Rio Grande do Sul e Sergipe, em que a rede privada teve um incremento proporcional maior. O único estado com oferta maior de vagas na rede privada do que na municipal é Pernambuco.

Também há transformações no próprio período. De 2001 a 2013, verifica-se que a rede municipal cresceu proporcionalmente mais do que a rede privada, até 2006, ano de aprovação do Fundeb. Depois de 2006, o crescimento proporcional da rede privada é maior. De 2006 a 2013, o crescimento das vagas em creches privadas foi de 92%, superando a taxa de expansão de creches municipais (89%). Essa informação remete à necessidade de um estudo mais aprofundado para identificar que parcela dessas vagas da rede privada é oferecida pela rede conveniada, incluída no financiamento, após a promulgação do Fundeb.

Algumas hipóteses preliminares podem ser aventadas para explicar esse aumento da rede privada: a ampliação de matrículas em instituições conveniadas teria ocorrido em função da LRF que limita a 54% de gastos com pessoal, o custo menor da criança nas instituições conveniadas, aumento da matrícula na rede privada com fins lucrativos em função do crescimento do poder aquisitivo da população de maior renda no período recente, a necessidade de ter a criança na escola por razões de alteração de mercado de trabalho ou novos padrões familiares.

No Gráfico 1, é possível verificar o incremento de vagas no período analisado com alguns atos normativos selecionados: PNE, Fundeb e Brasil Carinhoso. Verifica-se um incremento de matrículas após aprovação do PNE, em 2001. Entretanto, a tendência de crescimento é acentuada após a aprovação do Fundeb, em 2006, e de sua regulamentação, em 2008.

GRÁFICO 1 – Quantidade de matrículas em creche, por rede de ensino (2001-2013)



Fonte: Elaboração das autoras com base nos dados do Inep (BRASIL, 2001-2013).

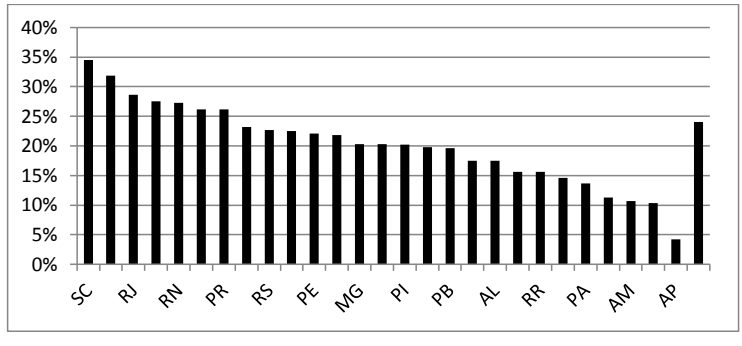
Foi observada, ainda, a partir de 2007, uma correlação entre a cobertura de atendimento e a existência do Fundeb, indicando que esse ato normativo teve impacto no atendimento. A correlação é maior a partir de 2008.

Ao analisar as matrículas de 2001 a 2013, observa-se que são ofertadas vagas em creche nas redes municipal e privada. Isso sugere que, se os instrumentos normativos federais foram capazes de induzir os municípios a incluir em sua agenda a oferta de creche, não foram capazes, por outro lado, de oferecer apoio suficiente para que a oferta direta pelos municípios se expandisse no ritmo definido pela política nacional. A expansão se deu com participação expressiva da rede privada, sendo necessários novos estudos para analisar o perfil dessa expansão e, no que se refere à rede conveniada, que tipo de apoio governamental a sustentou.

Nos limites deste artigo, não são analisadas as matrículas por município, para identificar as desigualdades intrarregionais, aspecto importante para compreender melhor as relações intergovernamentais e o desafio da coordenação federativa dessa política específica. Tampouco foi possível analisar a desigualdade intramunicipal entre as diferentes redes – a pública e a privada conveniada, também sob a responsabilidade dos municípios.

Outro aspecto pesquisado refere-se ao acesso à creche. Verifica-se, no Brasil, que, em 2010, 24% das crianças com até três anos estavam matriculadas (Gráfico 2). A desigualdade entre os estados é expressiva. Os estados que apresentam maior cobertura são Santa Catarina (35%), São Paulo (32%), Rio de Janeiro (29%) e Ceará (28%). Já os estados com menores coberturas municipais são Amapá (4%) e Roraima (10%).

GRÁFICO 2 – Taxa de atendimento de crianças com idades até três anos por estado (2010)



Fonte: Elaborado pelas autoras com base nos trabalhos de Alves e Silva (2013, p. 858).

Verifica-se uma "subpriorização" dessa etapa de ensino em relação ao ensino fundamental na oferta de vagas.[16] Muitas localidades investiram primeiramente no ensino fundamental, em função de sua obrigatoriedade determinada pela legislação federal, pela indução dos governos estaduais de processos de municipalização e pelo Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério. Posteriormente, os municípios iniciaram esforços para aumentar as matrículas de pré-escola e, mais recentemente, de creche.

16 Abrahão (2005) mostra que, do total de gastos, no Brasil, a educação infantil representava 11%, em 1995, e 9%, em 2002. Do gasto público em educação, o ensino superior fica com a maior parcela. No Estado de São Paulo, em 2012, dos gastos municipais com educação, 35% destinavam-se à educação infantil e 65% ao ensino fundamental (FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS; FUNDAÇÃO PREFEITO FARIA LIMA; FUNDAÇÃO DO DESENVOLVIMENTO ADMINISTRATIVO, 2013).

Não foi possível, nos limites deste artigo, analisar a variação entre os estados quanto à matrícula em creches, aspecto que requer novos estudos de caráter explicativo que considerem aspectos como perfil de urbanização do estado e porte de seus municípios; inserção da mulher no mercado de trabalho e características do próprio mercado de trabalho; relevância da educação infantil e das creches na agenda local; importância do tema "direitos das crianças" na agenda estadual e existência de ação indutora por parte dos governos estaduais, dentre outros fatores.

A importância da indução federal também pode ser vista com a iniciativa da Ação Brasil Carinhoso, que suplementa em 50% o Fundeb, para atender a crianças do PBF em creche. Com esse programa, há a inclusão de crianças em situação de pobreza. Em 2012, primeiro ano do incentivo, são atendidas 261.890 "novas" crianças e 453.569, em 2013, representando um crescimento de 73%, no período. As matrículas de crianças do Brasil Carinhoso representam 10% e 17% do total de matrículas em 2012 e 2013, respectivamente, indicando um aumento do acesso das crianças mais pobres.

Dados do Brasil Carinhoso mostram que a expansão do atendimento em creches nesse segmento – crianças em situação de pobreza – também é desigual entre os estados (Tabela 2).

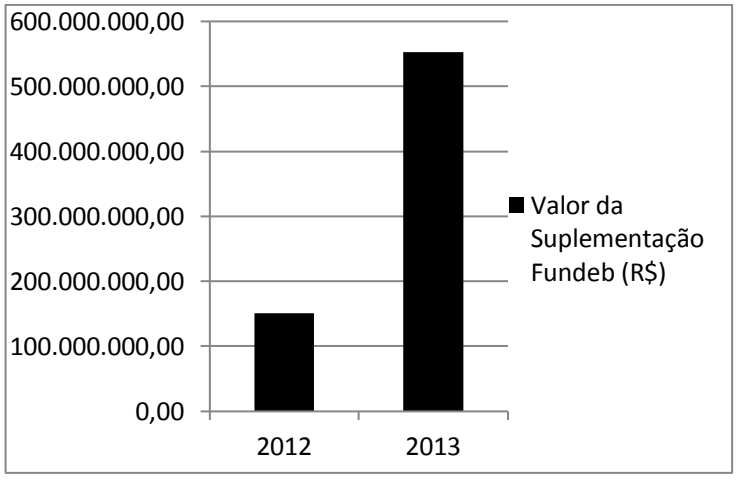
TABELA 2 – Quantidade de crianças do PBF atendidas no Brasil Carinhoso que receberam suplementação do Fundeb (2012 e 2013)

ESTADO	2012	2013	VARIAÇÃO
AC	610	1.987	226%
AL	3.424	8.055	135%
AP	272	632	132%
AM	1.403	3.837	173%
BA	17.694	33.640	90%
CE	21.216	43.107	103%
DF	0	1.976	-
ES	9.147	13.057	43%
GO	5.369	11.016	105%
MA	11.136	16.770	51%
MT	7.248	10.341	43%
MS	4.611	9.192	99%
MG	26.970	47.665	77%
PA	4.028	8.398	108%
PB	5.730	9.265	62%
PR	16.180	26.596	64%
PE	8.169	14.768	81%
PI	2.590	13.167	408%
RJ	11.241	28.057	150%
RN	10.502	16.376	56%
RS	13.263	19.198	45%
RO	1.420	1.760	24%
RR	216	498	131%
SC	7.444	11.549	55%
SP	65.165	92.814	42%
SE	2.126	3.069	44%
TO	4.716	6.779	44%
Brasil	**261.890**	**453.569**	**73%**

Fonte: Elaboração das autoras com base nos dados fornecidos pela Secretaria Extraordinária para Superação da Extrema Pobreza, por meio de pedido ao Sistema Eletrônico do Serviço de Informações ao Cidadão (e-SIC) (os de 2013 não são finais) (BRASIL, 2014).

Em 2012, inscreveram-se 2.246 municípios para receber suplementação do Fundeb por meio do Brasil Carinhoso. Em 2013, 3.450 localidades foram atendidas (62% dos municípios brasileiros). A suplementação financeira repassada aos municípios, por sua vez, passou de R$ 151.074.200,27, em 2012, para R$ 553.326.137,09 em 2013 (Gráfico 3).

GRÁFICO 3 – Valores repassados aos municípios de suplementação do Fundeb pelo Brasil Carinhoso (2012 e 2013)



Elaboração das autoras com base nos dados fornecidos pela Secretaria Extraordinária para Superação da Extrema Pobreza, por meio de pedido ao Sistema Eletrônico do Serviço de Informações ao Cidadão (e-SIC) (os de 2013 não são finais) (BRASIL, 2014).

Em alguns estados, como Mato Grosso, Rio Grande do Norte e Paraná, a suplementação do Fundeb pela associação ao Brasil Carinhoso ocorreu em mais de 80% dos municípios (82%, 82% e 84%, respectivamente). Entretanto, há estados, como o Amazonas, com 27% de seus municípios inseridos no programa. A apropriação desses recursos também é heterogênea: Piauí tem incremento de 998% no valor recebido, e Roraima registra 184%.

O Brasil Carinhoso prioriza as crianças mais pobres, e contribui para a expansão da oferta (e do acesso) a esse segmento. Antes da sua criação, a expansão da matrícula em creches não implicou redução da desigualdade de acesso à educação nessa faixa etária. Um exemplo é o Estado de São Paulo, onde, apesar do aumento de matrículas, de 2000 a 2010, a desigualdade de renda entre os que frequentam creche manteve-se praticamente inalterada. No período, a taxa de frequência entre as crianças de famílias mais ricas permaneceu muito maior do que a de crianças das famílias mais pobres (FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS; FUNDAÇÃO PREFEITO FARIA LIMA; FUNDAÇÃO DO DESENVOLVIMENTO ADMINISTRATIVO, 2013).

Os dados reunidos neste artigo mostram heterogeneidades regionais (ARRETCHE, 1999), apesar da existência de normas nacionais. Tais heterogeneidades exerceram um impacto significativamente variável na oferta de creches no território nacional. Faz-se necessário entender as complexas relações entre as diferentes esferas de governo, e entre os segmentos da sociedade civil, assim como as especificidades dos desafios da educação infantil e da oferta de creche em cada local. Assim se poderá contribuir para a proposição de medidas que contribuam para reduzir a desigualdade entre estados e municípios, envolvendo, de um lado, adequações à normatização existente, de outro, mecanismos de adaptação das políticas ao perfil diversificado dos municípios brasileiros.

CONSIDERAÇÕES FINAIS

Após a CF de 1988, atos normativos federais regulamentaram a educação infantil, orientados para a extensão do acesso e para redução das desigualdades entre as regiões brasileiras. Os dados sugerem que esses atos, fundos e programas federais contribuíram para o aumento de vagas em creches e para a expansão do acesso, havendo, no entanto, muito ainda a percorrer para se alcançar a meta da universalização e da redução das desigualdades regionais.

Parte da regulamentação federal criada é uma forma de coordenação federativa. Um exemplo é o Fundef e, posteriormente, o Fundeb, que trazem à discussão o uso dos recursos na manutenção e no desenvolvimento da educação. Esses fundos constituem instrumentos poderosos de influência federal sobre a municipalização e expansão da oferta de vagas (ARRETCHE, 2012). A transformação do Fundef em Fundeb constituiu uma inflexão importante na política educacional, ao incluir a educação infantil como foco do programa, ao lado do ensino fundamental e médio.

Os atos normativos promulgados, em especial pós-Fundeb, sofreram alterações importantes no que diz respeito à política de creches. Um exemplo é a inclusão de instituições comunitárias, confessionais ou filantrópicas sem fins lucrativos,

conveniadas com o Poder Público, no cômputo das matrículas para efeito da distribuição dos recursos previstos.

O marco legal, após 2011, aponta nova estratégia de indução do governo federal. A União, por meio do Brasil Carinhoso, cria um novo desenho de apoio para que as localidades ampliem a oferta de vagas em creche com o foco na erradicação da pobreza. O novo regramento amplia os recursos para novas vagas abertas para crianças do PBF e possibilita que os recursos do Fundeb sejam utilizados para cuidado e segurança alimentar e nutricional, no caso de crianças que integram o programa Brasil Carinhoso.

A União tem papel ativo na coordenação federativa (ARRETCHE, 2012) e implementa deliberações das Conferências Nacionais de Educação para aumentar o atendimento a crianças de baixa renda e o custeio dos serviços de creche. Entretanto, deve ser ressaltado que, pós-Brasil Carinhoso, são apoiadas, com os 50% adicionais de suplementação do Fundeb, apenas as novas vagas e turmas para crianças do PBF.

Outras pesquisas com os atores envolvidos podem identificar as motivações que levaram a União a permitir que somente beneficiários do PBF usem o suplemento de 50% do Fundeb com o cuidado integral e a segurança alimentar e nutricional das crianças. Será esse o novo movimento de indução? Essa estratégia de focalização pode ser entendida como antagônica aos objetivos de universalização da educação. Pode, por outro lado, ser percebida como medida que visa a criar a oportunidade para que crianças excluídas pela extrema pobreza tenham condições de acesso à creche e, mais do que isso, de permanência na creche.

É necessário ainda verificar se, no caso da política de creche, a autoridade decisória (*policy decison making*) está com a União e a de implementação (*policy-making*), com os municípios (ARRETCHE, 2012), e se a ação local é, a um só tempo, incentivada e limitada pela regulação de níveis superiores de governo. Assim, deve-se verificar se o desenho e o modo de operação são "fechados" pelo governo federal (FARAH, 2012, 2013) ou se há espaços de inovação e de adaptação ao local. A LDB dá aos municípios autonomia na gestão pedagógica, na estruturação de sua rede e no apoio ao educando.

Entretanto, fica uma questão: como os municípios se têm adaptado para implementar os programas federais e manter sua autonomia?

No apoio da União à ampliação da oferta de vagas, há indícios de que os municípios só podem aderir se cumprirem um conjunto de requisitos (FARAH, 2012) ou utilizar os recursos de determinada forma, como o Proinfância e Brasil Carinhoso. Coexistem com eles, no entanto, outros mecanismos, como o Fundeb, que têm um princípio equalizador, e que influenciam os municípios a implementarem essa etapa de ensino, sem, contudo, uma especificação estrita do "como fazer". A União tem utilizado, nesse setor, portanto, mecanismos diferenciados de coordenação federativa.

Corrobora-se, assim, a tese de que, para entender a implementação da política de creche e as relações interfederativas, é necessária uma análise mais aprofundada de sua evolução no decorrer do tempo, bem como de sua gestão fiscal (ALMEIDA, 2000). Mostra-se que a descentralização da educação cria uma forma de governo mais complexa e não é um processo finalizado de repasse de atribuições aos municípios (RODDEN, 2005). No caso específico da política de creche, a inclusão dessa etapa de educação no âmbito da política educacional nacional significou o desenho de um sistema federativo baseado na cooperação entre diferentes níveis de governo, no qual se atribuiu a responsabilidade pela oferta aos municípios.

Há evoluções incrementais no apoio da União para que os municípios possam cumprir suas atribuições em relação à oferta de creches. E está em curso uma transformação educacional, que modifica o processo de escolarização da primeira infância em nosso País. O debate deve ser aprofundado, sendo necessário conhecer as inter-relações que influenciaram a configuração dessa política, bem como verificados os condicionantes das diferentes ofertas de vagas no âmbito local e nos diversos estados.

Outro aspecto que sobressai no estudo consiste na variação entre os provedores de vagas em creche. As matrículas que mais cresceram, entre 2001 e 2013, são as municipais e privadas, parte das quais conveniadas com governos locais.

Após a regulamentação do Fundeb, as matrículas privadas cresceram mais do que as municipais.

Os mecanismos federais contribuíram para ampliar as matrículas nessas redes, e um estudo futuro deverá analisar as motivações dos gestores municipais para cada modalidade de oferta (pública ou conveniada). O aprofundamento dessa pesquisa revela-se importante para identificar, primeiramente, se o crescimento das matrículas da rede privada corresponde predominantemente à oferta de instituições conveniadas com o Poder Público ou à rede privada com fins lucrativos.

O estudo revela que, mesmo com normas, diretrizes, instrumentos e programas federais, a implementação é muito diferenciada entre os estados, indicando a necessidade de novas reflexões sobre as especificidades e capacidades institucionais de cada localidade (ALMEIDA, 2000; ABRUCIO, 2005; BICHIR, 2011, 2014) na oferta de creche.

Os dados apontam, ainda, que a meta estabelecida pelo PNE, de atendimento de 50% das crianças em creches, até 2023, está longe de ser atendida. Entretanto, a cobertura vem crescendo de forma gradual: em 1998, era de apenas 9%, e, em 2010, de 24% (ALVES; SILVA, 2013).

Entretanto, mesmo com os incentivos dados pela União para a ampliação da oferta em creches, os municípios continuam priorizando o ensino fundamental, etapa compartilhada com os estados. As vagas em creches, apesar da oferta crescente, ainda são insuficientes para atender à meta de atendimento de 50% das crianças com até três anos. Embora a educação infantil seja de competência exclusivamente local, os municípios vêm assumindo a oferta de creche como prioridade de forma lenta e gradual.

A implementação da política de creche – e a própria (re)formulação dessa política – são processos que vêm se consolidando de forma incremental (LINDBLOM, 1959). Envolvem mudança cultural em relação ao valor dado à educação infantil e em relação à noção do direito da criança à educação e dos direitos dos trabalhadores.

Algumas dificuldades e alguns desafios foram identificados, como o financiamento, a equalização entre municípios, a desigualdade de acesso entre segmentos da população,

a desigualdade de resposta de estados e municípios à indução federal e a "subpriorização" pelos municípios – ou simples atenção – à educação dessa faixa etária. A União, com seu papel de coordenação federativa, tem atuado para enfrentar esses problemas. Mas há ainda muito a percorrer nessas questões para que o Brasil avance na garantia do direito à creche.

REFERÊNCIAS

ABRAHÃO, Jorge. Financiamento e gasto público da Educação básica no Brasil e comparações com alguns países da OCDE e América Latina. *Educação e Sociedade*, Campinas, v. 26, n. 92, p. 841-858, out. 2005.

ABRUCIO, Fernando Luiz. A coordenação federativa no Brasil: a experiência do período FHC e os desafios do governo Lula. *Revista de Sociologia e Política*, n. 24, p. 41-67, jun. 2005.

AFONSO, José Roberto Rodrigues. Brasil, um caso a parte. In: REGIONAL SEMINAR OF FISCAL POLICY FROM CEPAM/ILPES, 16., Santiago do Chile, 2004. *Anais*… Santiago do Chile, 2004.

ALMEIDA, Maria Hermínia Tavares de. *Federalismo e proteção social*: a experiência brasileira em perspectiva comparada. In: SEMINÁRIO PACTO FEDERATIVO E GUERRA FISCAL NO BRASIL. São Paulo: FGV, 2000. Disponível em: <http://www.fflch.usp.br/dcp/assets/docs/MariaHerminia/federalismodef.pdf>. Acesso em: 12 jun. 2013.

_____. Recentralizando a federação. *Revista de Sociologia e Política*, n. 24, p. 29-40, jun. 2005.

ALVES, Thiago; SILVA, Rejane Moreira da. Estratificação das oportunidades educacionais no brasil: contextos e desafios para a oferta de ensino em condições de qualidade para todos. *Educação e Sociedade*, Campinas, v. 34, n. 124, p. 851-879, 2013. Disponível em: <http://www.scielo.br/pdf/es/v34n124/11.pdf>. Acesso em: 5 fev. 2014.

ARELARO, Lisete Regina Gomes. Para onde vai a educação infantil no Brasil? Algumas considerações face à nova LDB e à Emenda Constitucional 14/96. In: MACHADO, L. M. A. (Org.). *Educação infantil em tempos de LDB*. São Paulo: Fundação Carlos Chagas, 2000. p. 51-63.

ARRETCHE, Marta T. S. Mitos da descentralização: mais democracia e eficiência nas políticas públicas? *Revista Brasileira de Ciências Sociais*, São Paulo, n. 31, jun. 1996.

_____. Políticas sociais no Brasil: descentralização em um estado federativo. *Revista Brasileira de Ciências Sociais*, São Paulo, v. 14, n. 40, p. 111-141, jun. 1999.

_____. *Democracia, federalismo e centralização no Brasil*. Rio de Janeiro: FGV; Fiocruz, 2012.

BICHIR, Renata Mirandola. *Mecanismos federais de coordenação de políticas sociais e capacidades institucionais locais:* o caso do Programa Bolsa Família. 2011. Tese (Doutorado) – Programa de Pós-Graduação em Sociologia e Ciência Política, do Instituto de Estudos Sociais e Políticos, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2011. Disponível em: <http://www.fflch.usp.br/centrodametropole/antigo/static/uploads/Bichir_tese_rev.pdf>. Acesso em: 2 jun. 2014.

_____. *Os mecanismos de coordenação federal do Programa Bolsa Família*. Brasília: MDS, Sagi. 2014. Disponível em: <http://www.cienciapolitica.org.br/wp-content/uploads/2014/04/28_6_2012_22_3_15.pdf>. Acesso em: 30 mar. 2014.

BRASIL. Congresso Nacional. Constituição da república federativa do Brasil. *Portal da legislação* – governo federal. Brasília, DF, 1988. Disponível em: <http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm>. Acesso em: 10 out. 2014.

_____. Lei 8.069, de 13 de julho de 1990. Dispõe sobre o Estatuto da Criança e do Adolescente. *Portal da legislação* – governo federal. Brasília, DF, 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8069.htm. Acesso em: 3 jun. 2014.

_____. Lei 8.666, de 21 de junho de 1993. Regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. *Portal da legislação* – governo federal. Brasília, DF, 1993. Disponível em: < http://www.planalto.gov.br/ccivil_03/leis/l8666cons.htm>. Acesso em: 19 set. 2014.

_____. Emenda constitucional 14, de 12 de setembro de 1996. Modifica os arts. 34, 208, 211 e 212 da Constituição Federal e dá nova redação ao Art. 60 do Ato das Disposições Constitucionais Transitórias. *Portal da legislação* – governo federal. Brasília, DF, 1996a. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc14.htm>. Acesso em: 10 mai. 2014.

_____. Lei 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. *Portal da legislação* – governo federal. Brasília, DF, 1996b. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9394.htm>. Acesso em: 15 jan. 2014

_____. Lei 9.424, de 24 de dezembro de 1996. Dispõe sobre o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério. *Portal da legislação* – governo federal. Brasília, DF, 1996c. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9424.htm>. Acesso em: 10 jan. 2014.

_____. Lei complementar 101, de 4 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal. *Portal da legislação* – governo federal. Brasília, DF, 2000. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm>. Acesso em: 3 abr. 2014.

_____. Lei 10.172, de 9 de janeiro de 2001. Aprova o Plano Nacional de Educação. *Portal da legislação* – governo federal. Brasília, DF, 2001. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10172.htm>. Acesso em: 28 mar. 2014.

_____. Lei 11.114, de 16 de maio de 2005. Torna obrigatório o início do ensino fundamental aos seis anos de idade. *Portal da legislação* – governo federal. Brasília, DF, 2005. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Lei/L11114.htm>. Acesso em: 30 abr. 2014

_____. Emenda constitucional 53, de 19 de dezembro de 2006. Dá nova redação aos arts. 7º, 23, 30, 206, 208, 211 e 212 da Constituição Federal e ao art. 60 do Ato das Disposições Constitucionais Transitórias. *Portal da legislação* – governo federal. Brasília, DF, 2006a. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc53.htm>. Acesso em: 14 abr. 2014.

_____. Lei 11.274, de 6 de fevereiro de 2006. Dispõe sobre a duração de nove anos para o ensino fundamental, com matrícula obrigatória a partir dos seis anos de idade. *Portal da legislação* – governo federal. Brasília, DF, 2006b. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Lei/l11274.htm>. Acesso em: 7 mai. 2014.

_____. Lei 11.494, de 20 de junho 2007. Regulamenta o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação – Fundeb. *Portal da legislação* – governo federal. Brasília, DF, 2007a. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/l11494.htm>. Acesso em: 3 set. 2014.

_____. Decreto 6.253, de 13 de novembro de 2007. Dispõe sobre o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação - FUNDEB, regulamenta a Lei no 11.494, de 20 de junho de 2007, e dá outras providências. *Portal da legislação* – governo federal. Brasília, DF, 2007b. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/Decreto/D6253.htm>. Acesso em 5 mai. 2014.

_____. Lei 11.738, de 16 de julho de 2008. Institui o piso salarial profissional nacional para os profissionais do magistério público da educação básica. *Portal da legislação* – governo federal. Brasília, DF, 2008. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11738.htm>. Acesso em: 14 out. 2014.

_____. Emenda constitucional 59, de 11 de novembro de 2009. Acrescenta § 3º ao art. 76 do Ato das Disposições Constitucionais Transitórias para reduzir, anualmente, a partir do exercício de 2009, o percentual da Desvinculação das Receitas da União incidente sobre os recursos destinados à manutenção e desenvolvimento do ensino de que trata o art. 212 da Constituição Federal, dá nova redação aos incisos I e VII do art. 208, de forma a prever a obrigatoriedade do ensino de quatro a dezessete anos e ampliar a abrangência dos programas suplementares para todas as etapas da educação básica, e dá nova redação ao § 4º do art. 211 e ao § 3º do art. 212 e ao caput do art. 214, com a inserção neste dispositivo de inciso VI. *Portal da legislação* – governo federal. Brasília, DF, 2009. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc59.htm>. Acesso em: 2 set. 2014.

_____. Fundo Nacional de Desenvolvimento da Educação. *Programas/Proinfância*. Brasília, DF: FNDE , [200-]. Disponível em: <http://www.fnde.gov.br/programas/proinfancia/proinfancia-apresentacao>. Acesso em: 15 jun. 2013.

_____. Lei 12.462, de 4 de agosto de 2011. Institui o Regime Diferenciado de Contratações Públicas - RDC; altera a Lei no 10.683, de 28 de maio de 2003, que dispõe sobre a organização da Presidência da República e dos Ministérios, a legislação da Agência Nacional de Aviação Civil (Anac) e a legislação da Empresa Brasileira de Infraestrutura Aeroportuária (Infraero); cria a Secretaria de Aviação Civil, cargos de Ministro de Estado, cargos em comissão e cargos de Controlador de Tráfego Aéreo; autoriza a contratação de controladores de tráfego aéreo temporários; altera as Leis nos 11.182, de 27 de setembro de 2005, 5.862, de 12 de dezembro de 1972, 8.399, de 7 de janeiro de 1992, 11.526, de 4 de outubro de 2007, 11.458, de 19 de março de 2007, e 12.350, de 20 de dezembro de 2010, e a Medida Provisória no 2.185-35, de 24 de agosto de 2001; e revoga dispositivos da Lei no 9.649, de 27 de maio de 1998. *Portal da legislação* – governo federal. Brasília, DF, 2011a. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/Lei/L12462.htm>. Acesso em: 14 mai. 2014.

_____. Lei 12.499, de 29 de setembro de 2011. Autoriza a União a transferir recursos financeiros aos municípios e ao Distrito Federal, com a finalidade de prestar apoio financeiro à manutenção de novos estabelecimentos públicos de educação infantil. *Portal da legislação* – governo federal. Brasília, DF, 2011b. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12499.htm>. Acesso em: 14 set. 2014.

_____. Lei 12.695, de 25 de julho de 2012. Dispõe sobre o apoio técnico ou financeiro da União no âmbito do Plano de Ações Articuladas; altera a Lei no 11.947, de 16 de junho de 2009, para incluir os polos presenciais do sistema Universidade Aberta do Brasil na assistência financeira do Programa Dinheiro Direto na Escola; altera a Lei no 11.494, de 20 de junho de 2007, para contemplar com recursos do FUNDEB as instituições comunitárias que atuam na educação do campo; altera a Lei no 10.880, de 9 de junho de 2004, para dispor sobre a assistência financeira da União no âmbito do Programa de Apoio aos Sistemas de Ensino para Atendimento à Educação de Jovens e Adultos; altera a Lei no 8.405, de 9 de janeiro de 1992; e dá outras providências. Portal da legislação – governo federal. Brasília, DF, 2012a. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12695.htm>. Acesso em: 6 jan. 2014.

_____. Lei 12.722, de 3 de outubro de 2012. Altera as Leis nos 10.836, de 9 de janeiro de 2004, 12.462, de 4 de agosto de 2011, e 11.977, de 7 de julho de 2009; dispõe sobre o apoio financeiro da União aos Municípios e ao Distrito Federal para ampliação da oferta da educação infantil; e dá outras providências. Portal da legislação – governo federal. Brasília, DF, 2012b. Disponível em: < http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/Lei/L12722.htm>. Acesso em: 7 mai. 2014.

_____. Lei 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *Portal da legislação* – governo federal. Brasília, DF, 2014a. Disponível em: <http://www.planalto.gov.br/CCIVIL_03/_Ato2011-2014/2014/Lei/L13005.htm>. Acesso em: 6 jul. 2014.

_____. Lei 13.019, de 31 de julho de 2014. Estabelece o regime jurídico das parcerias voluntárias, envolvendo ou não transferências de recursos financeiros, entre a administração pública e as organizações da sociedade

civil, em regime de mútua cooperação, para a consecução de finalidades de interesse público; define diretrizes para a política de fomento e de colaboração com organizações da sociedade civil; institui o termo de colaboração e o termo de fomento; e altera as Leis 8.429, de 2 de junho de 1992, e 9.790, de 23 de março de 1999. *Portal da legislação* – governo federal. Brasília, DF, 2014b. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2014/Lei/L13019.htm>. Acesso em: 10 set. 2014.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sistema de consulta a matrícula do Censo Escolar*. Brasília, DF: MEC, 2001-2013. Disponível em: <http://portal.inep.gov.br/basica-censo-escolar-matricula>. Acesso em: 30 maio 2013.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. *Fundef* – Fundo de manutenção e desenvolvimento do ensino fundamental e de valorização do magistério: manual de orientação. Brasília, DF: MEC, maio 2004. Disponível em: <http://www.oei.es/quipu/brasil/Manual_FUNDEF.pdf>. Acesso em: 21 abr. 2014.

BRASIL. Ministério de Desenvolvimento Social e Combate à Fome. *Brasil carinhoso*. Brasília, DF: [201-]. Disponível em: <http://www.mds.gov.br/brasilsemmiseria/brasil-carinhoso>. Acesso em: 13 abr. 2014.

BRASIL. Supremo Tribunal Federal. *Ag. Reg. no Recurso Extraordinário 464.143*. Agravante: Município de Santo André. Agravado: Gabriel Duarte Galbero. Relatoria: Min. Ellen Grace São Paulo, 2009. Disponível em: < http://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=608283>. Acesso em: 8 nov. 2014.

_____. *Plano Brasil Sem Miséria*. [201-]. Disponível em: <http://www.brasilsemmiseria.gov.br/>. Acesso em: 03 abr. 2014.

_____. Secretaria Extraordinária para Superação da Extrema Pobreza. [Mensagem pessoal]. Mensagem recebida por <carminhameirelles@gmail.com> em 26 mar. 2014.

BRESSER-PEREIRA, Luiz Carlos; GRAU, Nuria Cunill. Entre o estado e o mercado: o público não estatal. In: BRESSER-PEREIRA, Luiz Carlos; GRAU, Nuria Cunill (Org.). *O público não-estatal na reforma do estado*. Rio de Janeiro: FGV, 1999. p. 15-48.

CARVALHO, Cristina Helena Almeida de. Financiamento da educação brasileira – perspectiva histórica e desafios futuros. In: FUNDAÇÃO JOÃO MANGABEIRA *Avaliação de políticas públicas federativas*. São Paulo: Quanta, 2014.

CONFERÊNCIA NACIONAL DE EDUCAÇÃO. *Documento final*. Brasília, DF: MEC, 2010. Disponível em: <http://conae.mec.gov.br/images/stories/pdf/pdf/documetos/documento_final_sl.pdf>. Acesso em: 2 set. 2014.

CONFEDERAÇÃO NACIONAL DOS MUNICÍPIOS. *Pauta municipalista e agenda legislativa*. Brasília: CNM, 2013.

CRUZ, Maria do Carmo Meirelles Toledo; FARAH, Marta Ferreira Santos. As normatizações federais e os impactos nas creches em âmbito subnacional. In: ENANPAD, 38, 2014a, Rio de Janeiro. *Anais*… Rio de Janeiro, 14 a 17 de setembro de 2014a.

_____. The decentralization of the early childhood education in Brazil: challenges to the local government. In: CONGRESS OF IIAS_PROGRAMA, IFRANE, 13 a 17 de junho de 2014, Marrocos. *Anais...* Marrocos: IIAS 2014b.

CURY, Carlos Roberto Jamil; FERREIRA, Luiz Antonio Miguel. Justiciabilidade no campo da educação. *Revista de Política e Administração da Educação*, Porto Alegre, v. 26, n. 1, p. 75-103, jan./abr. 2010.

EATON, Kent; DICKOVICK, J. Tyler. The politics of re-centralization in Argentina and Brazil. *Latin American Research Review*, n. 39, v. 1, p. 90-122, 2004.

FARAH, Marta Ferreira Santos. Reconstruindo o Estado: gestão do setor público e reforma da Educação. *Planejamento e Políticas Públicas*, Brasília, v. 1, n. 11, p. 189-237, 1994.

_____. Parcerias, novos arranjos institucionais e políticas públicas no nível local de governo. *Revista de Administração Pública*, Rio de Janeiro, n. 35, v. 1, p. 119-144, jan./fev. 2001.

_____. *Políticas sociais e governos locais*: indução ou coordenação federativa? In: COLÓQUIO INTERNACIONAL DE PODER LOCAL, 12., 10 a 12 de dezembro de 2012, Salvador. *Trabalho apresentado...* Salvador: UFBA, 2012.

_____. Políticas públicas e municípios: inovação ou adesão? In: LUKIC, M. R.; TOMAZINI, C. (Org.). *As ideias também importam*: abordagem cognitiva e políticas públicas no Brasil. Curitiba: Juruá, 2013. v. 1. p. 171-193.

FERRAREZI, Elisabete Roseli. *Evolução das políticas dirigidas à infância e à adolescência*: a Secretaria do Menor de São Paulo e a introdução de um novo paradigma. 1994. Dissertação (Mestrado) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 1994.

FIUZA, Eduardo P. S. O regime diferenciado de contratações públicas e a agenda perdida das compras públicas. *Radar*: tecnologia, produção e comércio exterior, Brasília, n. 19, p. 7-20, maio 2012. Disponível em: <http://www.ipea.gov.br/portal/images/stories/PDFs/radar/120522_radar19_cap1>. Acesso em: 18 set. 2014.

FRANZESE, Cibele. *Federalismo cooperativo no Brasil*: da Constituição de 1988 aos sistemas de políticas públicas. 2010. Tese (Doutorado) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2010. Disponível em: <http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/8219/72060100752.pdf?sequence=1>. Acesso em: 10 abr. 2014.

FUNDO NACIONAL DE DESENVOLVIMENTO DA EDUCAÇÃO. *Programas/Proinfância*. Brasília: Fundo Nacional de Desenvolvimento da Educação , s.d. Disponível em: <http://www.fnde.gov.br/programas/proinfancia/proinfancia-apresentacao>. Acesso em: 15 jun. 2013.

GOMES. Sandra. Políticas Nacionais e Implementação Subnacional: Uma Revisão da Descentralização Pós-Fundef, *Revista de Ciências Sociais*, Rio de Janeiro, v. 52, n. 3, p. 659-690, 2009.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. *Políticas sociais*: acompanhamento e análise. Brasília, DF: Ipea, 2013. v. 21. Disponível em: <http://www.ipea.gov.br/agencia/images/stories/PDFs/politicas_sociais/bps_21_completo.pdf>. Acesso em: 3 nov. 2014.

KINGDON, John W. Como chega a hora de uma ideia? In: SARAVIA, E.; FERRAREZI, E. (Org.). *Políticas públicas*: coletânea. Brasília: Enap, 2006. v. 1. p. 219-224.

KUGELMAS, Eduardo; SOLA, Lourdes. Recentralização/descentralização: dinâmica do regime federativo no Brasil dos anos 90. Tempo Social. *Revista de Sociologia da USP*, São Paulo, v. 11, n. 2, p. 63-83, 1999.

LACZYNSKI, Patrícia. *Fundef*: seus primeiros resultados no estado de São Paulo. 2003. Dissertação (Mestrado) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2003. Disponível em: <http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/2228/74588.pdf?sequence=2>. Acesso em: 21 abr. 2014.

LÍCIO, Elaine Cristina. *Para além da recentralização*: os caminhos da coordenação do Programa Bolsa Família (2003-2010). 2012. Tese (Doutorado) – Universidade de Brasília, Brasília, 2012. Disponível em: <http://repositorio.unb.br/bitstream/10482/10759/1/2012_ElaineCristinaLicio.pdf>. Acesso em: 2 jun. 2014.

LINDBLOM, Charles E. The Science of "Muddling Through". *Public Administration Review*, v. 19, n. 2, p. 79-88, 1959.

MARCONDES, Mariana Mazzini. *A política de creches do PAC-2 e o cuidado*: análise na perspectiva da indivisibilidade e interdependência de direitos. In: SEMINÁRIO INTERNACIONAL FAZENDO GÊNERO 10, 2013, Florianópolis: UFSC; UDESC. *Anais....* Florianópolis, 2013.

RICH, Jessica A. J.; GÓMEZ, Eduardo J. Centralizing Decentralized Governance in Brazil. *Publius*, v. 42, n. 4, p. 636-661, 2012.

RODDEN, Jonathan. Federalismo e descentralização em perspectiva comparada: sobre significados e medidas. *Revista Sociologia e Política*, Curitiba, n. 24, p. 9-27, jun. 2005.

ROSEMBERG, Fúlvia. O movimento de mulheres e a abertura política no Brasil: O caso da creche. *Cadernos de Pesquisa*, São Paulo, n. 51, p. 7-79, nov. 1984. Disponível em: <http://educa.fcc.org.br/pdf/cp/n51/n51a07.pdf>. Acesso em: 22 maio 2014.

SÃO PAULO. Fundação Sistema Estadual de Análise de Dados; Fundação Prefeito Faria Lima; Fundação do Desenvolvimento Administrativo. *Elementos para a melhoria das condições de oferta de educação infantil no estado de São Paulo* – relatório final, São Paulo: Seade/Cepam/Fundap, 2013.

SÃO PAULO. Secretaria Municipal de Educação. Departamento de Suprimento Escolar. *Valores médios de alimentação escolar*. São Paulo: Secretaria Municipal de Educação de São Paulo, 2013.

SILVEIRA, Adriana Aparecida Dragone. Judicialização da educação para a efetivação do direito à educação básica. *Jornal de Políticas Educacionais*, v. 5, p. 30-40, 2011.

SISTEMA EDUCATIVO NACIONAL DO BRASIL. Educação Infantil. In: ORGANIZAÇÃO DOS ESTADOS IBERO-AMERICANOS PARA EDUCAÇÃO, CIÊNCIA E CULTURA. *Informe OEI* – Ministério 2002. Brasília, DF, 2003. p. 66-83. Disponível em: <http://www.oei.es/quipu/brasil/educ_inicial.pdf>. Acesso em: 03 abr. 2013.

SUBIRATS, Joan; KNOWEPFEL, Peter; LARRUE, Corinne; VARONE, Frédéric. *Análisis y gestión de políticas públicas*. 2. ed. Barcelona: Planeta, 2012.

MARIA DO CARMO MEIRELLES TOLEDO CRUZ
Mestre em Administração Pública e doutoranda em Administração Pública e Governo pela Fundação Getulio Vargas – Escola de Administração de Empresas de São Paulo (FGV-EAESP). Técnica do Centro de Estudos e Pesquisas de Administração Municipal (Cepam). São Paulo, São Paulo, Brasil
mcmcruz@sp.gov.br; carminhameirelles@gmail.com

MARTA FERREIRA SANTOS FARAH
Doutora em Sociologia pela Universidade de São Paulo (USP). Coordenadora do Curso de Mestrado e de Doutorado em Administração Pública e Governo da Fundação Getulio Vargas – Escola de Administração de Empresas de São Paulo (FGV-EAESP) e professora dos cursos de Graduação e de Pós-Graduação *stricto sensu* em Administração Pública e Governo da FGV-EAESP. Integra o Departamento de Gestão Pública e participa do Centro de Estudos de Administração Pública e Governo. São Paulo, São Paulo, Brasil
marta.farah@fgv.br

NATASHA BORGES SUGIYAMA
Doutora em Governo pela Universidade de Texas, Austin. Professora dos cursos de Graduação e de Pós-Graduação de Ciência Política da Universidade de Wisconsin-Milwaukee, Milwaukee, Wisconsin, Estados Unidos
sugiyamn@uwm.edu

Recebido em: NOVEMBRO 2014
Aprovado para publicação em: DEZEMBRO 2014

EVIDÊNCIAS DO EFEITO DA REPETÊNCIA NOS PRIMEIROS ANOS ESCOLARES

ERISSON VIANA CORREA
ALICIA BONAMINO
TUFI MACHADO SOARES

RESUMO

No contexto do debate sobre as vantagens e desvantagens da repetência escolar como medida pedagógica, este trabalho pretende verificar os efeitos dessa prática sobre o desempenho de estudantes dos anos iniciais do ensino fundamental. O objetivo é investigar se alunos que repetiram, mesmo ao custo de um ano extra de escolarização, apresentam ganhos reais de proficiência que justifiquem a utilização de tal medida. Para tanto, desenvolvemos um estudo observacional que compara dois grupos de alunos, repetentes e promovidos, organizados por pareamento assistido, analisando o impacto da repetência nas medidas de proficiência em Língua Portuguesa e Matemática desses grupos em diferentes momentos da sua trajetória escolar. Os resultados mostram que, ao longo do tempo, os alunos repetentes aprendem menos que os promovidos. Quando comparados em uma mesma série, independente do ano letivo em que a cursaram, os repetentes se saem melhor, embora o efeito em Língua Portuguesa se perca ao longo das séries.

PALAVRAS-CHAVE REPETÊNCIA • AVALIAÇÃO DO RENDIMENTO ESCOLAR • PESQUISA LONGITUDINAL • ENSINO FUNDAMENTAL.

* O presente trabalho foi realizado com apoio do Programa Observatório da Educação, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes/Brasil).

RESUMEN

En el marco del debate sobre las ventajas y desventajas de la repitencia escolar como medida pedagógica, este trabajo pretende verificar los efectos de tal práctica sobre el desempeño de estudiantes de los años iniciales de la educación fundamental. El objetivo es investigar si los alumnos que repitieron, incluso con el costo de un año más de escolarización, presentan aumentos reales de proficiencia que justifiquen la utilización de tal medida. Para ello desarrollamos un estudio observacional que compara dos grupos de alumnos, repitentes y promovidos, organizados por pares asistidos, analizando el impacto de la repitencia en las medidas de proficiencia en Idioma Portugués y Matemática de dichos grupos en diferentes momentos de su trayectoria escolar. Los resultados muestran que, con el tiempo, los alumnos repitentes aprenden menos que los promovidos. Cuando comparados en un mismo curso, independientemente del año lectivo en que se encuentran, los repitentes obtienen mejores resultados, aunque el efecto en Idioma Portugués se pierda a lo largo de los distintos cursos.

PALABRAS CLAVE REPITENCIA • EVALUACIÓN DEL RENDIMIENTO ESCOLAR • INVESTIGACIÓN LONGITUDINAL • EDUCACIÓN FUNDAMENTAL.

ABSTRACT

In the context of the debate about the advantages and disadvantages of school retention as a pedagogic method, this study intends to verify the effects of this practice on the performance of students in the first years of schooling. The objective is to investigate whether students who repeat, even at the cost of an extra year of schooling, show real gains in proficiency that justify the use of such a measure. To accomplish this, we developed an observational study that compares groups of students, retained and promoted, organized by assisted pairing, analyzing the impact of retention on the proficiency measures on these groups in Portuguese language and Mathematics, at different moments along their school path. The results show that, over time, the retained students learn less than the promoted students. When compared within the same grade, independent of the school year of the student, the retainees were more successful, although the effect on the Portuguese language gets lost over the grades.

KEYWORDS RETENTION • ASSESSMENT OF SCHOOL PERFORMANCE • LONGITUDINAL RESEARCH • ELEMENTARY EDUCATION.

INTRODUÇÃO

Os altos níveis de repetência estão entre os mais graves e crônicos problemas educacionais do país. Sua existência está ligada à baixa qualidade do ensino, e as altas taxas de defasagem idade/série e de evasão escolar são suas consequências diretas. A extensão e a intensidade do uso desse recurso por parte das unidades escolares têm sido um fator de enorme pressão sobre os indicadores de fluxo e rendimento nos sistemas de ensino brasileiros.

As preocupações com a disseminação da repetência escolar no Brasil vêm de longa data. Nas décadas de 1930 e 1940, com as primeiras tentativas de sistematizar os levantamentos estatísticos sobre a situação educacional do país, Teixeira de Freitas e Lourenço Filho protagonizaram intensos debates acerca do melhor tratamento e da interpretação dos dados estatísticos relativos ao ensino básico (TEIXEIRA DE FREITAS, 1937, 1941; LOURENÇO FILHO, 1941, 2002; GIL, 2007; GADOTTI; ROMÃO, 2000). Decisões equivocadas tomadas àquela época incorporaram aos dados oficiais erros graves que distorceriam as estatísticas, superestimando as taxas

de evasão e subestimando as de repetência. Como resultado, inúmeras análises baseadas nesses dados chegaram a diagnósticos equivocados e imprecisos sobre esse problema educacional. A política educacional decorrente privilegiou por muito tempo a expansão quantitativa do atendimento escolar em detrimento da melhoria da qualidade.

Zaia Brandão, Anna Maria Baeta e Any Dutra Rocha (1983), em clássico estudo sobre o estado do conhecimento acerca da evasão e da repetência, verificaram que as investigações evidenciavam um estrangulamento na base do sistema escolar, concentrado na passagem da primeira para a segunda série, além da inadequação do tratamento pedagógico dado à questão da repetência.

Concomitantemente, um grupo de pesquisadores apontou as falhas das estatísticas oficiais em uma série de estudos desenvolvidos desde a década de 1980 e considerados fundamentais para a compreensão do problema da repetência no Brasil (FLETCHER, 1985; FLETCHER; CASTRO, 1986; FLETCHER; RIBEIRO, 1987; RIBEIRO, 1991; KLEIN; RIBEIRO, 1991).

Com base em um modelo denominado Profluxo, que utilizava dados coletados na Pesquisa Nacional de Amostra por Domicílio (Pnad) do Instituto Brasileiro de Geografia e Estatística (IBGE), constatou-se que as altíssimas taxas de repetência na primeira série do então ensino de primeiro grau interrompiam de tal forma o fluxo, que produziam, já ao final do primeiro ano de escolaridade, um gargalo que bloqueava a entrada de novas coortes de alunos no sistema escolar. Essa retenção elevava substancialmente a distorção idade/série e depreciava o rendimento dos alunos, em um processo acumulado que se arrastava desde a entrada até a saída daquela coorte do sistema, configurando, assim, talvez o nosso maior problema educacional.

Diversas medidas de correção de fluxo foram adotadas com maior ênfase a partir desses achados. A repetência, que estava em um patamar de 60% nos anos 1980, caiu para próximo de 45% na década de 1990 e recuou mais um pouco nos anos 2000, estabilizando-se em um patamar ainda alto, em torno dos 25% (KLEIN, 2006).

Grande parte da queda observada na primeira série se deve à introdução das Classes de Alfabetização em alguns

estados, com a instituição de uma política muito comum na época: a "subseriação".[1] Além disso, a introdução do Ciclo Básico de Alfabetização (CBA) acabou com a possibilidade de reprovação na 1ª série, diminuindo as taxas de repetência (KLEIN; RIBEIRO, 1998). A queda significativa nas taxas de repetência, ao longo das décadas de 1980 e 1990, principalmente na primeira série, também foi reflexo direto da adoção de medidas como as Classes de Aceleração, a organização escolar em Ciclos e a Promoção Automática. No entanto, a despeito da importante melhoria nas taxas de fluxo alcançada pelo Brasil na década de 1990, ainda estamos entre os países com as maiores taxas de repetência no mundo, perdendo apenas para Angola (SOARES, 2007).[2]

1 Subseriação refere-se à prática de repartir a primeira série em duas etapas de escolarização. Essa estratégia era corrente até a instituição do ensino fundamental de nove anos.

2 Em recente relatório do Pisa (ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO, 2011), o Brasil aparece situado entre os três países que apresentam a maior quantidade de alunos com experiência prévia de repetência, atrás apenas de Tunísia e Macau.

LITERATURA SOBRE A REPETÊNCIA

A literatura sobre repetência e as políticas educacionais vigentes no Brasil, a partir da década de 1990, sofreram influência significativa das descobertas da década anterior, motivadas pelas conclusões do modelo Profluxo.

O artigo "A pedagogia da repetência" (RIBEIRO, 1991) é representativo do momento histórico de mudança na perspectiva das análises sobre o problema da repetência. Sérgio da Costa Ribeiro colocou em xeque os dados estatísticos fornecidos pelo Ministério da Educação (MEC), mostrando que o maior entrave para o aumento da escolaridade da população brasileira era a repetência e não as altas taxas de evasão.

Durante esse período, o debate sobre os problemas decorrentes da repetência escolar no Brasil esteve circunscrito à mensuração e análise do fluxo escolar e à compreensão dos aspectos sociais e culturais envolvidos nesse fenômeno.

Mais recentemente, outros estudos buscaram compreender os efeitos de políticas de não retenção. Naércio Menezes-Filho *et al.* (2008) avaliaram o impacto da implantação de programas de progressão continuada nos indicadores de rendimento das escolas e no desempenho dos alunos, com base nos dados do Censo Escolar 2006 e da Prova Brasil 2005. Os estudos identificaram o aumento da aprovação acompanhado da diminuição do abandono escolar em todas as séries, entretanto, com perda significativa para o desempenho dos alunos

da 8ª série. Carvalho (2009) avaliou o impacto das políticas de não retenção no desempenho acadêmico de alunos do ensino fundamental de escolas públicas brasileiras. Utilizando dados do Censo Escolar 2001 e 2005, e resultados do Saeb[3] 2001 e da Prova Brasil 2005, o autor constatou que a prática de não retenção não teve impacto algum sobre o desempenho dos alunos, nem mesmo entre aqueles com pior desempenho, grupo que se julgava ser beneficiado por essa ação.

3 Saeb é o Sistema de Avaliação da Educação Básica composto por dois processos: a Avaliação Nacional da Educação Básica (Aneb), amostral; e a Avaliação Nacional do Rendimento Escolar (Anresc) também conhecida como Prova Brasil, censitária. Ambas utilizam os mesmos testes e instrumentos, e são realizadas nacionalmente a cada dois anos.

Neves e Pazello (2012) investigaram o impacto das mudanças nas políticas de promoção das escolas a partir de resultados do Saeb, no período 1999-2003. Os achados dessa pesquisa mostram que as escolas que aboliram a repetência tiveram um acréscimo no desempenho em Matemática, mas não em Língua Portuguesa, entre os alunos da 4ª série quando comparados às escolas que seguiram com o regime seriado. Não foram encontrados, no entanto, efeitos significativos para a 8ª série em nenhuma das duas disciplinas. Entre as escolas que abandonaram os sistemas de progressão e passaram a utilizar o regime de repetência, a pesquisa indicou um resultado superior em Matemática na 4ª série apenas para a regressão realizada no âmbito do aluno. Na 8ª série, os alunos apresentaram um desempenho inferior em Matemática, tanto no âmbito do aluno quanto no âmbito da escola, e desempenho inferior em Língua Portuguesa, significativo apenas para o âmbito da escola.

Mais recentemente, com a emergência dos dados longitudinais, já começam a despontar no Brasil trabalhos sobre o efeito da repetência propriamente dita, e não apenas das políticas antirrepetência. Embora incipiente no Brasil, esse tipo de estudo tem uma longa tradição no debate internacional.

Parte importante desse debate foi introduzida no Brasil por Marcel Crahay (2006) em um artigo que revisa uma série de estudos sobre o efeito da repetência. Nele, o autor sintetiza quase um século do percurso da pesquisa sobre o tema e suas principais evoluções metodológicas. Para o autor, a trajetória recente dessas pesquisas pode ser considerada paradigmática, pelas múltiplas tentativas de aproximação da análise científica à realidade da prática escolar, motivada pelos esforços de diversos pesquisadores que, ao longo desse período,

procuraram tanto contornar obstáculos metodológicos como responder às objeções lançadas a esse tipo de pesquisa. O estudo conclui que o fato de repetir uma série e recomeçar todo o programa de curso não ajuda alunos com dificuldades a superarem os obstáculos e obterem sucesso escolar.

Luciana Soares Luz (2008) é autora de um dos primeiros estudos no contexto brasileiro a utilizar dados longitudinais para analisar os efeitos da repetência. Fazendo uso dos resultados da base Fatores Associados ao Desempenho Escolar – Centro de Desenvolvimento de Planejamento Regional da Universidade Federal de Minas Gerais (Cedeplar)/Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), a autora realizou um pareamento entre alunos repetentes e promovidos da 7ª série, em 2002, a partir do método de escores de propensão. Na comparação dos escores desses alunos entre a 7ª série em 2002 e a 8ª série em 2008, os resultados obtidos mostram que os alunos repetentes obtêm ganhos na proficiência inferiores aos dos alunos nas mesmas condições e que foram promovidos. Ao final do ano repetido, o desempenho dos repetentes é muito parecido com o de seus novos colegas de turma e bastante inferior ao apresentado pelos alunos que progrediram.

Mais recentemente, Juliana Riani, Vania Silva e Tufi Soares (2010) conduziram novo estudo sobre efeitos da repetência no contexto brasileiro, a partir dos dados do Proalfa – Avaliação da Alfabetização da Rede Pública do Estado de Minas Gerais – com o objetivo de compreender se a retenção dos alunos que cursavam o 3º ano de escolaridade em 2009 foi ou não benéfica em termos de aprendizado. Os autores utilizaram modelos hierárquicos e, por meio deles, mostraram que dado dois alunos com uma mesma proficiência em 2008, um deles sendo promovido e o outro retido no 3º ano, o promovido tende a apresentar maior proficiência em 2009. Os resultados obtidos indicam, também, que o resultado geral da escola tem grande influência sobre os alunos de baixo desempenho, o que leva à consideração de que expor alunos repetentes ao mesmo conteúdo, com base no pressuposto de que existiria uma deficiência de aprendizagem decorrente unicamente da sua inabilidade para apreender, não resolverá

o problema, uma vez que fatores intraescolares possuem impacto significativo sobre a aprendizagem dos alunos.

METODOLOGIA

Antes dos primeiros esforços de estudos baseados em metodologias quase experimentais, diversos pesquisadores já se ocupavam em avaliar os efeitos da repetência em uma perspectiva comparada por meio da observação dos alunos retidos e promovidos. Entretanto, essas investigações tendiam a favorecer os alunos promovidos, uma vez que não se preocupavam em isolar as características individuais e familiares que poderiam interferir na formação e na comparação dos grupos de alunos retidos e não retidos (JACKSON, 1975).

As primeiras investigações que se preocuparam em captar os efeitos da repetência no desempenho dos alunos se baseavam em desenhos bastante elementares, denominados esquemas pré e pós-teste. Essas pesquisas selecionavam amostras formadas exclusivamente por alunos com dificuldade escolar. Seus conhecimentos eram testados a partir de exames, em que o primeiro teste era aplicado antes da decisão da repetência e o outro depois. Para garantir igualdade nas medidas, os pesquisadores utilizavam os mesmos testes cognitivos e isso impunha uma forte limitação ao estudo (JACKSON, 1975). Todos os estudos que se utilizaram desse delineamento naturalmente concluíram a favor de uma clara evolução cognitiva dos alunos repetentes.

Constatada essa limitação metodológica, os pesquisadores procuraram alternativas que permitissem mensurar os efeitos da repetência de uma forma mais apurada. A pesquisa experimental foi o modelo adotado para a compreensão dos efeitos da repetência, o que exigiria dos pesquisadores das ciências sociais e humanas posturas pouco convencionais diante de impedimentos de diversas ordens. Em primeiro lugar, o rigor exigido pela pesquisa experimental demandava que se procedesse a uma escolha aleatória dos indivíduos que comporiam os grupos. Depois, para garantir que não se introduzisse nenhum viés na comparação, seria necessário que os grupos formados apresentassem certa equivalência

inicial que permitisse a comparabilidade entre eles. Por último, um estudo para ser considerado experimental de fato exigiria que o pesquisador tivesse pleno domínio sobre o tratamento a ser administrado.

As dificuldades encontradas para se estabelecer com segurança dois grupos distintos de alunos em que cada indivíduo de um grupo tenha as mesmas características socioeconômicas e de aprendizagem de um indivíduo do outro grupo, aliadas às restrições de escolha dos alunos por meio de sorteio, no qual todos tenham as mesmas probabilidades de receber o tratamento, colocaram os pesquisadores diante de um problema ético. A primeira proposta no sentido de superar essas dificuldades foram os estudos quase experimentais ou *passive observacional studies* (estudos observacionais),[4] assim denominados em referência à obra de Cook e Campbell (1979). Esses estudos buscaram acompanhar, durante determinado intervalo de anos, a evolução cognitiva de um conjunto de alunos com capacidades cognitivas iniciais semelhantes, sendo uns submetidos à retenção na mesma série e outros promovidos à série seguinte.

4 Assim denominados em referência à obra de Thomas Cook e Donald Campbell (1979).

Nesta pesquisa aproveitaremos o potencial empírico desse tipo de delineamento aliado à possibilidade de exploração dos dados da Pesquisa Geres (Estudo Longitudinal da Geração Escolar). Com isso, pretendemos oferecer uma contribuição ao debate sobre as vantagens e desvantagens da repetência no Brasil.

DELINEAMENTO DA PESQUISA

Este estudo pretende responder a duas questões básicas: 1) Os alunos que passam pela experiência da repetência apresentam ganhos reais de proficiência que justifiquem a utilização de tal medida mesmo ao custo de um ano extra de escolarização? 2) No caso de haver ganhos reais na aprendizagem do aluno repetente, esses ganhos se mantêm ao longo dos anos iniciais do ensino fundamental?

Para responder a elas, comparam-se dois grupos de alunos, um formado por repetentes e outro por promovidos. Chamaremos o primeiro grupo, formado por alunos repetentes na 1ª série/2º ano do ensino fundamental, de "grupo

experimental" ou "grupo de tratamento". O segundo grupo, formado por alunos que não sofreram repetência, será denominado "grupo controle". O comportamento dos alunos do grupo controle representará o comportamento do grupo experimental caso os alunos não tivessem sido submetidos à repetência. As análises comparativas foram realizadas tanto para Língua Portuguesa quanto para Matemática.

Os dados utilizados nesta investigação fazem parte da base de dados global da Pesquisa Geres 2005 (BROOKE; BONAMINO, 2011). Realizado entre 2005 e 2008, este estudo objetivou investigar a evolução das aprendizagens em Língua Portuguesa e Matemática de cerca de 21.500 alunos dos anos iniciais do ensino fundamental, matriculados em mais de 300 escolas de cinco cidades brasileiras: Belo Horizonte, Campinas, Campo Grande, Rio de Janeiro e Salvador. A pesquisa foi realizada em regime de colaboração interinstitucional entre seis universidades.[5]

O desenho do Geres foi construído a partir de uma amostra estratificada, com a participação de aproximadamente 300 escolas de diversas dependências administrativas: federais, estaduais, municipais e privadas. Durante os quatro anos do projeto, foram aplicadas cinco ondas de avaliação, duas no primeiro ano, em março e novembro de 2005, e as demais ao final de cada ano, até 2008. Além dos testes de proficiência em Língua Portuguesa e Matemática, foram aplicados também questionários contextuais sobre a escola, os professores, os diretores, os pais e os próprios alunos, com o objetivo de coletar dados que revelassem aspectos escolares, sociais e familiares que pudessem influenciar no desempenho destes últimos.

A construção dos testes cognitivos contou com a participação de especialistas da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) e do Centro de Alfabetização, Leitura e Escrita da Universidade Federal de Minas Gerais (Ceale/UFMG) para a elaboração da matriz de referência e dos testes de Língua Portuguesa, e da Universidade Federal de Juiz de Fora (UFJF) para a matriz de referência e dos testes de Matemática, bem como para a elaboração das escalas de proficiência da pesquisa.

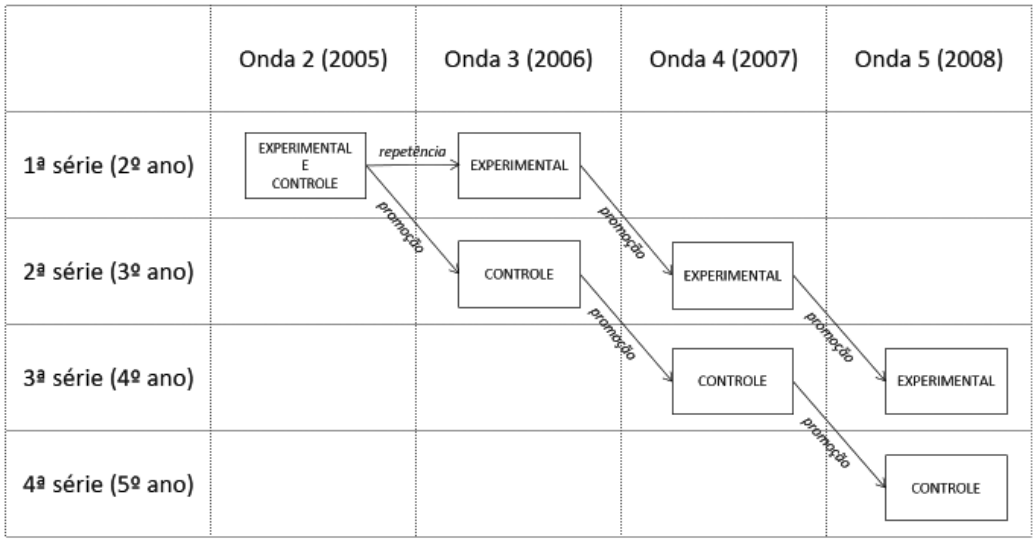
5 A realização do projeto ficou a cargo da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Universidade Federal de Minas Gerais (UFMG), Universidade Federal de Juiz de Fora (UFJF), Universidade Federal da Bahia (UFBA), Universidade Estadual do Mato Grosso do Sul (UEMS) e Universidade Estadual de Campinas (Unicamp).

As proficiências do Geres foram estimadas por meio do modelo proposto pela Teoria da Resposta ao Item (TRI) e sua escala foi equalizada com a do Saeb. Suas vantagens incluem uma melhor comparabilidade dos resultados alcançados pelos alunos nos testes construídos a partir de um conjunto de diferentes itens e aplicados em diferentes momentos (BROOKE; BONAMINO, 2011). Isso permite que esses alunos possam ter suas proficiências avaliadas de acordo com uma mesma escala, apesar de terem respondido testes diferentes. Os questionários contextuais, por sua vez, foram elaborados no sentido de permitir aos pesquisadores testar hipóteses relativas às características escolares, familiares, sociais, dos professores, gestores e alunos que poderiam influenciar nos resultados escolares. Além disso, há outros blocos de itens que foram construídos com base no interesse de testar a existência de certos construtos teóricos, como a motivação do aluno, por exemplo.

A principal vantagem do Geres é ser longitudinal, oferecendo dados cognitivos comparáveis durante os quatro anos iniciais do ensino fundamental. Isso permite um desenho de estudo observacional capaz de captar possíveis efeitos pedagógicos da repetência ao longo do tempo.

As observações correspondem às ondas de avaliação do estudo longitudinal. Isso quer dizer que nossa primeira observação será realizada com os alunos antes da decisão da repetência e as observações seguintes após a repetência. A Figura 1 apresenta a configuração da pesquisa, a partir do delineamento adotado.

FIGURA 1 – Esquema conceitual da pesquisa sobre o efeito da repetência



Fonte: Elaboração dos autores.

Na Figura 1, estão representadas, num esquema de fluxo, as trajetórias escolares dos dois grupos a serem pesquisados. No eixo horizontal, verificam-se os anos de aplicação do estudo longitudinal, divididos em quatro avaliações, que vão de 2005 a 2008. No eixo vertical, observam-se as séries cursadas pelos mesmos grupos de alunos. À medida que os grupos avançam nas colunas, há uma mudança no ano calendário, da mesma forma, quando mudam de linha, representa um avanço na série cursada.

Esse delineamento permite abordagens de dois tipos. No plano vertical, de acordo com a evolução das proficiências dos alunos no decorrer dos anos, que chamaremos "comparação por ondas". No plano horizontal, a partir das séries cursadas, o que denominaremos "comparação por série". A partir desse desenho pretendemos compreender se a repetência, a despeito de um ano a mais de escolaridade, leva os alunos do grupo experimental a apresentar maiores ganhos de proficiência do que os alunos do grupo controle que tiveram o mesmo ponto de partida.

Para a seleção dos indivíduos que comporão os grupos recorreremos a uma técnica denominada pareamento assistido[6] (ou *matching*). Essa técnica consiste em escolher para cada

6 O pareamento assistido consiste em tentar encontrar visualmente registros iguais na base de dados para os grupos de tratamento e controle. Ele permite alcançar o melhor pareamento possível desde que não haja muitos casos sendo tratados, nem muitas variáveis de controle. É particularmente indicado no caso de haver muitas opções para compor o grupo de controle, o que ocorre no presente estudo.

aluno do grupo experimental, outro indivíduo que não recebeu o tratamento e tem características idênticas a ele. O processo é descrito em detalhes na próxima seção.

O PROCESSO DE PAREAMENTO

Para proceder ao pareamento escolheram-se, como variáveis de controle, a proficiência em Língua Portuguesa e a proficiência em Matemática da 2ª Onda de avaliação do estudo longitudinal, o Indicador de Nível Socioeconômico (NSE), a dependência administrativa e o município, tendo como objetivo controlar possíveis distorções na comparação entre alunos provenientes de diferentes configurações sociais, redes ou tipos de escolas.

As unidades não pareadas, ou seja, aquelas que não se encaixam nos perfis definidos para esta pesquisa por não estarem de acordo com os critérios estabelecidos foram descartadas e, portanto, não farão parte dos resultados. Foram excluídos, também, os alunos que sofreram repetências a *posteriori*, pois o período de escolaridade visado pelo estudo longitudinal (os quatro anos iniciais do ensino fundamental) impõe uma limitação temporal que não permite a análise da trajetória de alunos retidos mais de uma vez.

Dentre os 19.757 alunos que iniciaram o estudo longitudinal, foram identificados 297 casos elegíveis para o grupo experimental, o que equivale aos alunos que estariam cursando pela segunda vez a 1ª série/2º ano do ensino fundamental, no momento da aplicação da Onda 3, e que não sofreram mais nenhuma repetência até o fim do estudo longitudinal.

Identificaram-se 9.558 casos de alunos elegíveis para o grupo controle, abrangendo os alunos com trajetórias escolares regulares, sem repetência, e com participação do estudo longitudinal em todos os anos. Esses alunos estavam cursando a 1ª série/2ª ano quando se aplicou a Onda 2, cursavam a 2ª série/3º ano quando se aplicou a 3ª Onda de avaliação e assim por diante.

O número total de repetentes segundo os critérios escolhidos foi de 297. Com o procedimento realizado conseguimos parear com grande aproximação 229 casos em

Matemática e 243 em Língua Portuguesa. Dado o número de variáveis utilizadas (rede, estrato, proficiência e NSE) e a grande restrição que impusemos ao pareamento, ao admitirmos um distanciamento muito pequeno entre os casos controle e experimental, concluímos que o processo de *matching* foi bem-sucedido, com uma taxa de aproveitamento próxima a 80% para Matemática e Língua Portuguesa. Com isso, a distribuição de casos pelas redes ficou mais equânime, como pode ser observado no Quadro 1, a seguir.

QUADRO 1 – Distribuição do número de pares por rede e município – Língua Portuguesa e Matemática

	MUNICÍPIO	REDE				TOTAL
		ESPECIAL[7]	ESTADUAL	MUNICIPAL	PRIVADA	
L. PORTUGUESA	Belo Horizonte[8]	0	21	4	0	25
	Campinas[9]	0	3	43	3	49
	Campo Grande[10]	0	38	98	0	136
	Rio de Janeiro[11]	19	0	7	7	33
	Total	19	62	152	10	243
MATEMÁTICA	Belo Horizonte	0	20	3	0	23
	Campinas	0	2	38	3	43
	Campo Grande	0	37	94	0	131
	Rio de Janeiro	19	0	6	7	32
	Total	19	59	141	10	229

Fonte: Elaboração dos autores.

As distribuições dos repetentes e promovidos são idênticas e, como se pode ver no Quadro 2, bastante parecidas também entre as amostras de Matemática e Língua Portuguesa para outras variáveis de contexto. Para os valores de nível socioeconômico utilizamos como critério que alunos pareados,

7 Na coluna Rede, o estrato Especial abrange os Colégios de Aplicação vinculados a Universidades nas cidades do Rio de Janeiro e Belo Horizonte, bem como as escolas federais de educação básica no município do Rio de Janeiro.

8 Em Belo Horizonte, a rede municipal oferecia ensino fundamental de nove anos, organizado em três ciclos com duração de três anos cada. A rede estadual também oferecia ensino fundamental de nove anos com regime de ciclos, porém com configuração diferente. O primeiro ciclo tinha duração de três anos, atendendo crianças a partir de 6 anos de idade. Os ciclos seguintes tinham duração de dois anos cada.

9 Em Campinas, as redes municipal e estadual ofereciam ensino fundamental com duração de oito anos, porém ambas com Classes de Alfabetização (CA) na pré-escola. A rede municipal se organizava em regime seriado, enquanto a rede estadual mantinha dois ciclos formados por quatro anos de escolaridade cada um.

10 Em Campo Grande, a rede municipal oferecia ensino fundamental de nove anos, com os dois primeiros anos de escolaridade compondo um único ciclo e atendendo crianças de 6 e 7 anos de idade. Os demais anos de escolaridade seguiam em regime seriado. A rede estadual também oferecia ensino fundamental de nove anos, porém em regime seriado.

11 A rede municipal do Rio de Janeiro oferecia ensino fundamental de nove anos, cujos três primeiros anos compunham um ciclo (atendendo crianças de 6, 7 e 8 anos de idade) seguido do regime de seriação.

seja no grupo experimental ou controle, guardassem entre si a distância de no máximo 0,6 pontos na escala de NSE. A diferença dos alunos repetentes com relação aos alunos promovidos foi de -0,05 pontos para o pareamento em Língua Portuguesa e de 0,04 pontos para o pareamento em Matemática.

Isso permitiu amostras mais homogêneas não só do ponto de vista da rede e da cidade, como também das características dos alunos, como podemos observar no Quadro 2, a seguir.

QUADRO 2 – Distribuição percentual dos grupos pareados quanto à cor/raça, sexo e frequência à pré-escola

CARACTERÍSTICAS		L. PORTUGUESA		MATEMÁTICA	
		REPETENTES	PROMOVIDOS	REPETENTES	PROMOVIDOS
Cor / raça	Branco	35,0	29,6	32,8	29,3
	Pardo	30,9	46,1	31,4	45,0
	Preto	17,3	14,8	17,5	14,0
	Amarelo	2,1	2,9	3,1	3,5
	Indígena	5,3	4,1	5,2	6,1
	N/I	9,5	2,5	10,0	2,2
Sexo	Feminino	35,8	40,3	35,8	47,2
	Masculino	57,6	57,6	57,2	51,1
	N/I	6,6	2,1	7,0	1,7
Frequência à Pré-escola	Não	23,0	19,3	24,5	20,5
	Sim	58,0	69,1	55,5	65,5
	N/I	18,9	11,5	20,1	14,0

Fonte: Elaboração dos autores.

Como se pode observar, o número de meninos é maior tanto entre os repetentes quanto entre os promovidos, seja em Língua Portuguesa ou em Matemática. E as meninas estão mais presentes entre os promovidos. No que diz respeito à cor declarada, as distribuições são bem parecidas. Há um número maior de alunos pardos entre os promovidos para ambas as disciplinas. No que se refere à frequência à pré-escola, o número de alunos promovidos que frequentaram a educação infantil é cerca de 10 pontos percentuais maior que o dos alunos repetentes, tanto para o pareamento em Língua Portuguesa quanto em Matemática.

No que diz respeito à medida de desempenho, em Língua Portuguesa a proficiência média ficou em torno de 106 pontos. Em Matemática, a média é de 112 pontos. As variações no tempo e nas séries observadas para as proficiências serão analisadas na próxima seção.

RESULTADOS

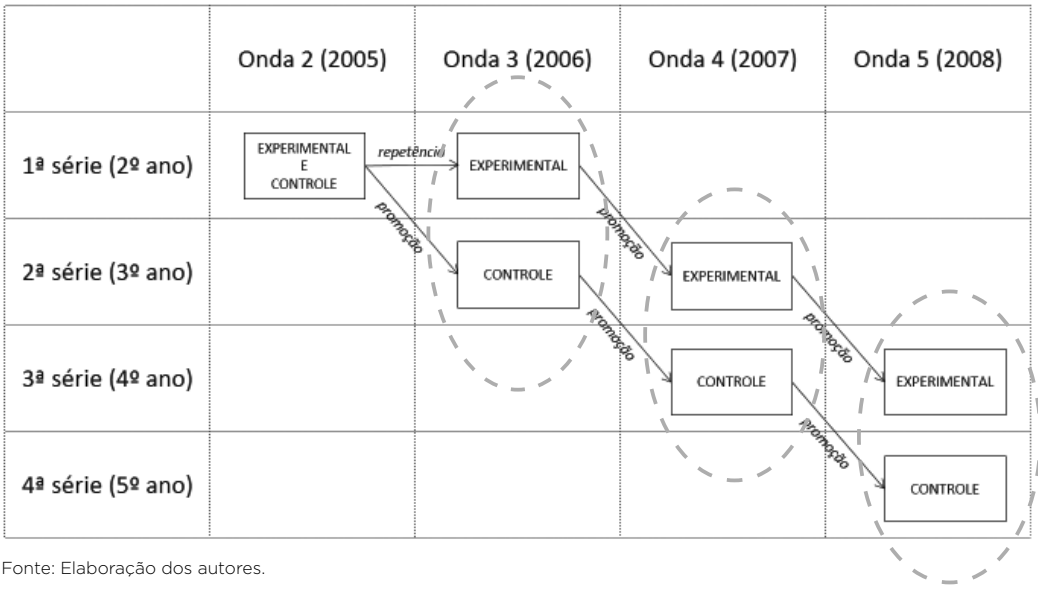
O EFEITO DA REPETÊNCIA SOBRE A PROFICIÊNCIA

Tradicionalmente, os autores que têm se voltado para a compreensão do efeito da repetência sobre o desenvolvimento cognitivo dos alunos se utilizam de duas abordagens básicas para a análise dos resultados (HOLMES, 1989). A primeira é uma estratégia de comparação denominada *same-age comparison*, pela qual os autores comparam os resultados de alunos repetentes com seus pares promovidos, levando em conta apenas a idade do aluno ou o ano em que a avaliação foi realizada, e ignorando a série cursada. A segunda estratégia adotada, denominada *same-grade comparison*, consiste em comparar o desempenho obtido pelos alunos repetentes em uma determinada série com o resultado obtido pelo aluno promovido para aquela mesma série no ano anterior. Adotaremos essas duas abordagens na pesquisa, adaptando-as ao estudo longitudinal.

COMPARAÇÃO POR ANO OU ONDA (*SAME-AGE COMPARISON*)

Por meio desta primeira abordagem buscamos comparar as proficiências dos alunos em um mesmo ano letivo ou em uma mesma onda de avaliação, independente da série que tenham cursado naquele ano, como pode ser observado na Figura 2.

FIGURA 2 – Estratégia de comparação por ano ou por onda de avaliação

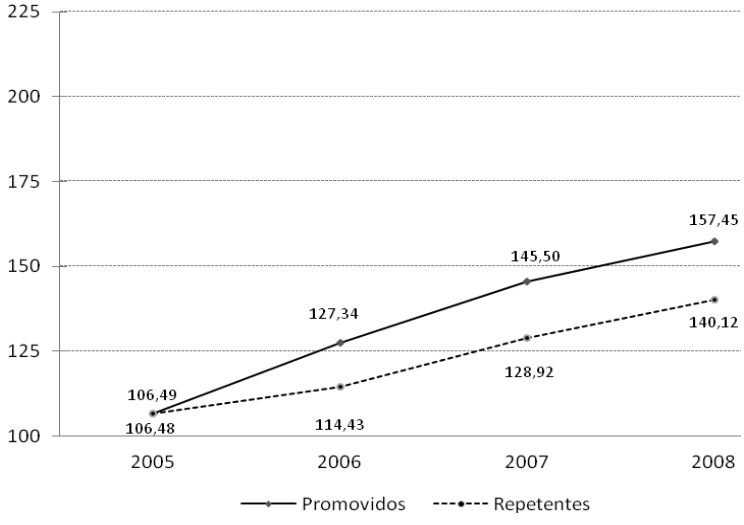


Fonte: Elaboração dos autores.

Incluímos no esquema de delineamento três círculos tracejados que indicam três comparações possíveis por meio dessa abordagem.

Os gráficos 1 e 2, a seguir, comparam as médias de proficiência em Língua Portuguesa e Matemática dos alunos repetentes e promovidos no decorrer das cinco ondas de avaliação. É importante relembrar que esses alunos foram submetidos ao pareamento a partir da 2ª Onda, o que ocorreu no final de 2005, quando todos os alunos ainda cursavam a 1ª série/2º ano e nenhum deles ainda havia passado pela experiência da repetência.

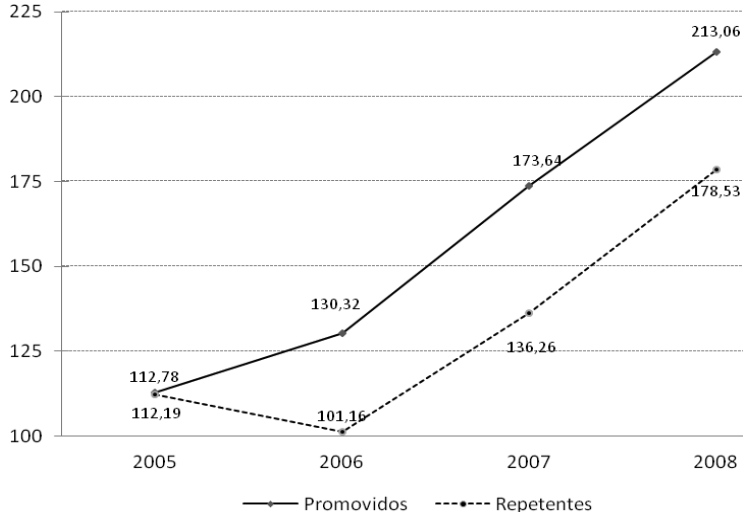
GRÁFICO 1 – Proficiência média em Língua Portuguesa dos alunos repetentes e promovidos: comparação por Onda (2005-2008)



Fonte: Elaboração dos autores.

Como se pode ver na comparação por ano, as médias dos alunos repetentes e promovidos ao final de 2005 estão em patamares idênticos, próximos dos 106 pontos na escala de proficiência em Língua Portuguesa. A partir daí, ambos os grupos continuam progredindo, embora apenas os promovidos o façam em um ritmo constante, com crescimento médio em torno de 17 pontos ao ano, entre 2005 e 2007. No grupo de alunos repetentes, em 2006, quando estes cursaram pela segunda vez a 1ª série/2º ano, há um crescimento médio de apenas 8 pontos, menos que a metade do crescimento atingido pelos alunos promovidos. A explicação para esse resultado pode estar no fato de os alunos promovidos terem sido expostos a conteúdos novos, constantes do programa da 2ª série/3º ano, enquanto os alunos repetentes estariam revendo conteúdos da 1ª série/2º ano. A retomada do ritmo do crescimento entre 2006 e 2007 em 14 pontos na escala de proficiência corrobora com esta possibilidade: apenas 4 pontos a menos que os alunos promovidos no mesmo período e 6 pontos a menos que no período anterior. Entre 2007 e 2008, os dois grupos praticamente empatam no crescimento médio, acrescentando aproximadamente 11 pontos.

GRÁFICO 2 – Proficiência média em Matemática dos alunos repetentes e promovidos: comparação por Onda (2005-2008)



Fonte: Elaboração dos autores.

Ambos os grupos partem de 112 pontos na escala de proficiência em Matemática, na Onda 2. Entretanto, no ano seguinte, promovidos e repetentes apresentam desempenhos muito díspares. Enquanto o grupo de alunos promovidos cresce 17 pontos em 2006, o grupo de repetentes apresenta queda de 11 pontos na escala entre 2005 e 2006. A partir daí, o crescimento dos dois grupos se acelera significativamente, acumulando 22 pontos entre os repetentes e 33 pontos entre os promovidos ao final de 2008.

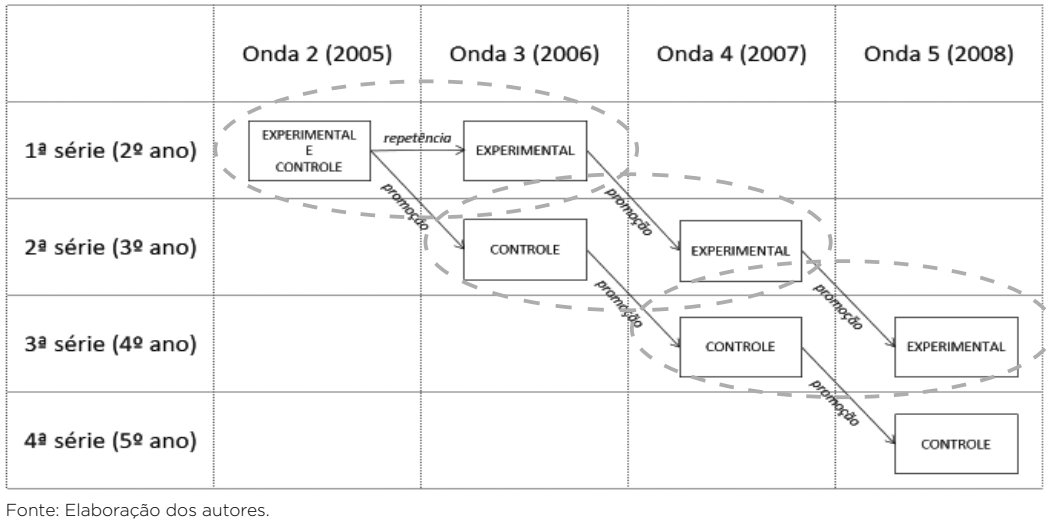
Embora os repetentes apresentem, em relação a 2005, uma queda na média da proficiência em Matemática em 2006, a proficiência dos alunos promovidos também apresenta um crescimento menor em 2006 do que nos demais anos. Esse fenômeno já havia sido identificado pelo estudo longitudinal no universo dos alunos participantes e hipóteses sobre essa desaceleração verificada no ritmo de aprendizagem de habilidades de Matemática, entre o 2º e o 3º ano escolar, estão sendo analisadas (BROOKE; BONAMINO, 2011).

COMPARAÇÃO POR SÉRIE (*SAME-GRADE COMPARISON*)

Utilizando os mesmos grupos da abordagem anterior, realizamos uma comparação baseada na série (*same-grade comparison*). Esta abordagem consiste em comparar a proficiência desses dois grupos de alunos a partir da série, independente do ano em que esta série foi cursada. Isso equivale a dizer que, apesar de as medidas serem comparadas a partir de uma mesma série, esses dois grupos de alunos mantém entre si um ano letivo de diferença.

A Figura 3 acompanha o esquema apresentado na seção anterior relativo às ondas de avaliação. A diferença é que as comparações demarcadas pelo pontilhado se dão em um plano horizontal, isto é, em uma abordagem que privilegia a série. Dessa forma, a comparação possibilita destacar qual a diferença entre as proficiências dos alunos repetentes e promovidos em uma mesma série.

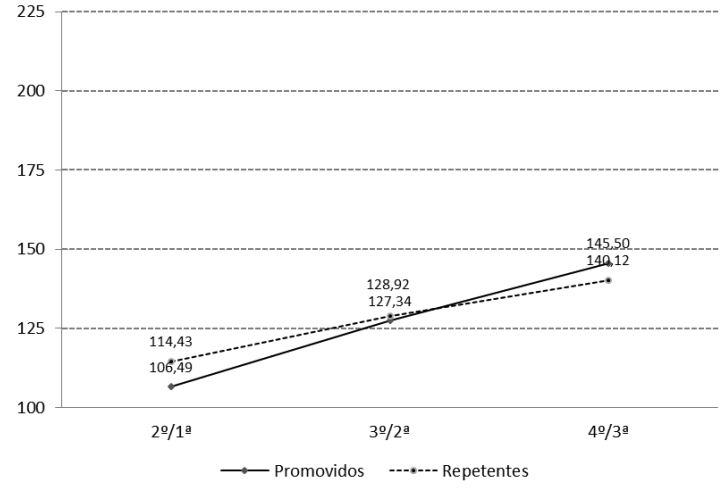
FIGURA 3 – Estratégia de comparação por série



Fonte: Elaboração dos autores.

Os gráficos 3 e 4, a seguir, apresentam a comparação das proficiências de ambos os grupos em Língua Portuguesa e Matemática, respectivamente, ao final da 1ª, 2ª e 3ª séries. Note-se que o grupo experimental estará sempre cursando as séries em questão com um ano de atraso em relação ao grupo controle.

GRÁFICO 3 – Proficiência média em Língua Portuguesa dos alunos repetentes e promovidos – comparação por série (2º ao 4º ano/1ª à 3ª série)

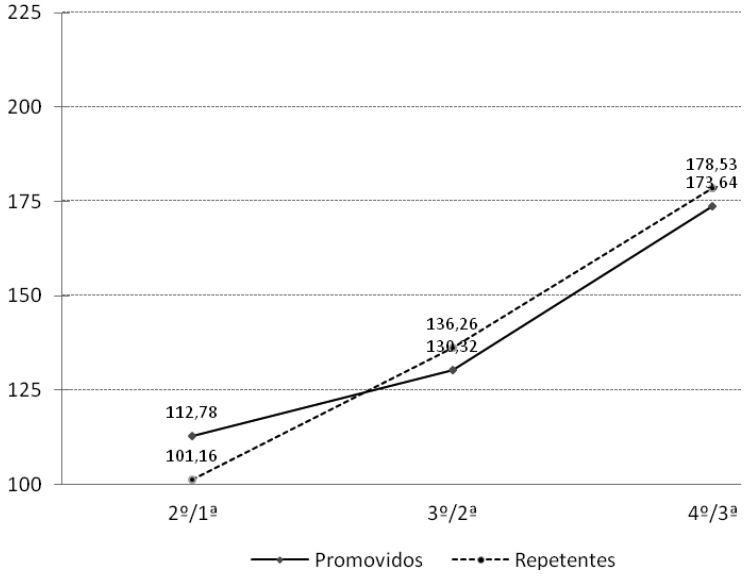


Fonte: Elaboração dos autores.

O Gráfico 3 mostra que, passado um ano após a repetência e ignorando-se o fato de que um ano letivo separa os resultados de ambos os grupos, os alunos repetentes apresentam um resultado mais satisfatório que os promovidos. Entretanto, essa constatação não subsiste a uma observação longitudinal mais extensa. Observa-se que a primeira medida dos alunos repetentes é de 114 pontos. Isso representa a média de proficiência desse grupo no ano de 2006, ou seja, o desempenho desses alunos no estudo longitudinal quando estavam cursando a 1ª série/2º ano pela segunda vez. Nota-se, também, que, após um ano cursando a mesma série, a diferença que esses alunos obtiveram em relação à proficiência de seus pares na mesma série no ano anterior é de apenas 8 pontos. Isso equivale a dizer que ao final de um ano de repetição de um mesmo currículo, a repetência como medida de recuperação das aprendizagens levou à agregação de apenas 8 pontos na média de proficiência em relação não só aos seus pares promovidos, mas também em relação ao próprio grupo, uma vez que as médias destes estavam pareadas no ano anterior. Essa diferença não se mantém no decorrer da

trajetória escolar. Quando comparados os resultados da 2ª série/3º ano, percebe-se que o suposto "efeito repetência", que já não era grande, reduz-se para apenas 1 ponto, invertendo-se os resultados já na 3ª série/4º ano, quando há uma diferença de 5 pontos, em média, a favor dos não repetentes.

GRÁFICO 4 – Proficiência média em Matemática dos alunos repetentes e promovidos – comparação por série (2º ao 4º ano/1ª à 3ª série)



Fonte: Elaboração dos autores.

Matemática apresenta um comportamento diferente quando se comparam as proficiências dos grupos por série. Os alunos repetentes apresentam uma média de 112 pontos na 1ª série/2º ano, que é menor que a dos promovidos em cerca de 8 pontos. Essa diferença a favor dos promovidos na 1ª série/2º ano se deve à queda verificada na proficiência dos alunos repetentes em 2006. A partir da 2ª série/3º ano nota-se uma diferença de 6 pontos a favor dos repetentes. Essa diferença cai cerca de 5 pontos na 3ª série/4º ano.

O período escolar que o estudo longitudinal nos permite acompanhar não é suficiente para saber o que acontece nas séries posteriores à 4ª série/5º ano do ensino fundamental.

O Quadro 3 apresenta, de forma sintética, as diferenças entre as médias dos repetentes e promovidos, na comparação por onda e por série.

QUADRO 3 – Diferenças entre a média das proficiências dos alunos repetentes e dos promovidos

	COMPARAÇÃO POR ONDA			COMPARAÇÃO POR SÉRIE		
	2006	2007	2008	1ª/2º	2ª/3º	3ª/4º
Português	-12,9	-16,5	-17,3	7,9	1,5	-5,3
Matemática	-29,1	-37,3	-34,5	-11,6	5,9	4,8

Fonte: Elaboração dos autores.

Na comparação por Onda, em 2006, um ano após a decisão de reter o aluno, os repetentes têm, em média, 12 pontos de proficiência a menos em Língua Portuguesa que seus pares promovidos. Em 2007, essa diferença aumenta para aproximadamente 15 pontos, estabilizando-se nesse patamar em 2008. Em Matemática, a diferença entre repetentes e promovidos é mais acentuada, chegando a 29 pontos ao final de 2006, 37 pontos ao final de 2007 e 28 pontos em 2008.

Na comparação por série, ao final de um ano repetido, os alunos repetentes apresentam 8 pontos a mais em Língua Portuguesa em relação aos seus colegas promovidos e ao próprio resultado anterior. Essa diferença se reduz para 3 pontos ao final da 2ª série, invertendo-se a favor dos promovidos ao final da 3ª serie, em cerca de 3 pontos. Em Matemática, os resultados se apresentam favoráveis à repetência. Alunos repetentes apresentam queda de cerca de 7 pontos na proficiência em relação aos seus pares e a si mesmos ao final da 1ª série do ensino fundamental. Há uma inversão dessa tendência na 2ª e na 3ª séries, que os coloca acima dos promovidos em 13 pontos. Esses resultados estão sintonizados com pesquisas internacionais, que comparam os alunos com base na idade (ou no ano letivo), o que tende a favorecer os alunos promovidos, enquanto as comparações baseadas na mesma série tendem a favorecer os estudantes retidos (HOLMES, 1989). As significâncias estatísticas das diferenças obtidas pelo teste t, para amostras independentes, são apresentadas no Quadro 4.

QUADRO 4 – Teste de comparação entre as médias dos alunos repetentes e promovidos

		LÍNGUA PORTUGUESA				MATEMÁTICA			
		MÉDIA	ERRO PADRÃO	ESTAT. t	P(T<=t)	MÉDIA	ERRO PADRÃO	ESTAT. t	P(T<=t)
2005(1)	Repetentes	90,85	1,71	-2,89	0,00	89,15	2	-2,84	0,00
	Promovidos	97,52	1,56			96,53	1,66		
2005(2)	Repetentes	106,48	1,31	-0,01	0,50	112,19	1,88	-0,22	0,41
	Promovidos	106,49	1,29			112,78	1,84		
2006	Repetentes	114,43	1,34	-6,5	0,00	101,16	2,88	-6,85	0,00
	Promovidos	127,34	1,47			130,32	3,13		
2007	Repetentes	128,92	1,42	-8,08	0,00	136,26	3,11	-8,5	0,00
	Promovidos	145,5	1,48			173,64	3,11		
2008	Repetentes	140,12	1,54	-7,99	0,00	178,53	3,76	-6,45	0,00
	Promovidos	157,45	1,53			213,06	3,81		
1ª /2º	Repetentes	114,43	1,34	4,28	0,00	101,16	2,88	-3,4	0,00
	Promovidos	106,49	1,29			112,78	1,84		
2ª/3º	Repetentes	128,92	1,42	0,77	0,22	136,26	3,11	1,34	0,09
	Promovidos	127,34	1,47			130,32	3,13		
3ª/4º	Repetentes	140,12	1,54	-2,52	0,01	178,53	3,76	1	0,16
	Promovidos	145,5	1,48			173,64	3,11		

Fonte: Elaboração dos autores.

CONCLUSÃO

Quando um estudante falha em demonstrar as competências necessárias para ser promovido à série seguinte, uma das decisões que cabe ao corpo docente no final do ano é em relação à retenção do aluno na série, sob a justificativa/expectativa de que, no próximo ano, a maturidade e a reexposição aos conteúdos da série irão prepará-lo melhor para as demandas escolares que lhe serão exigidas no decorrer das séries seguintes. Entretanto, esse tipo de decisão docente e os argumentos a favor da retenção na série como intervenção pedagógica para alunos de baixo desempenho acadêmico não têm encontrado respaldo nas evidências empíricas obtidas a partir de uma variedade de metodologias, em mais de meio século de pesquisas sobre o efeito da retenção. Os resultados desses estudos sugerem que reter alunos em uma

mesma série não acrescenta a eles maiores benefícios em relação aos seus pares promovidos, principalmente nos anos iniciais do ensino fundamental. Pode-se, ainda, baseando-se em outros estudos e outros autores, conjecturar que a repetência pode ter um impacto negativo no desenvolvimento acadêmico, social ou emocional do aluno, com consequências sérias em termos de abandono e evasão escolar.

No trabalho aqui realizado, procurou-se compreender as relações que se estabelecem entre a repetência e o desenvolvimento cognitivo dos alunos repetentes e não repetentes, que foram pareados por meio do pareamento assistido, utilizando-se de variáveis como o NSE e a proficiência prévia em Língua Portuguesa (leitura) e o NSE e a proficiência prévia em Matemática, como controle. A comparação foi realizada por série e por ano letivo ou onda de avaliação.

A comparação do desenvolvimento acadêmico de alunos repetentes e promovidos a partir de uma mesma série mostra que, apesar de ocorrerem ganhos na proficiência em Língua Portuguesa entre os alunos retidos, esses ganhos não se mantêm ao longo do tempo. Em Matemática, embora um pequeno efeito de 5 pontos na escala do Saeb se mantenha até a 4ª série/5º ano, não é suficiente a ponto de colocar esses alunos em níveis de aprendizagem diferentes dos promovidos, de maneira a justificar a repetência. Isso, associado aos resultados da comparação em relação às ondas – que apontam para uma diferença crescente entre promovidos e retidos a favor dos primeiros – permite-nos concluir que, do ponto de vista pedagógico, a repetência não garante a esses alunos melhores condições de aprendizagem que seus colegas promovidos.

Esses resultados se colocam a favor das políticas que visam à superação da retenção nos anos iniciais do ensino fundamental, uma vez que demonstram que reter alunos nesse segmento escolar pode comprometer o aprendizado num momento crucial de desenvolvimento das habilidades de leitura e escrita. Os achados contradizem os discursos que advogam a favor da repetência, mostrando que reter o aluno não é tão vantajoso como se pensa em termos cognitivos, pois não garante vantagens compensatórias futuras aos repe-

tentes. Mesmo para Matemática, que apresentou resultados mais favoráveis à repetência, essa diferença, embora estatisticamente significativa, não se mostrou relevante em termos pedagógicos, dado que os níveis de habilidades consolidadas não diferem entre promovidos e repetentes para justificar a adoção dessa medida.

REFERÊNCIAS

BRANDÃO, Zaia; BAETA, Anna Maria Bianchini; ROCHA, Any Dutra Coelho. *Evasão e repetência no Brasil*: a escola em questão. 2. ed. Rio de Janeiro: Dois Pontos, 1983.

BROOKE, Nigel; BONAMINO, Alícia (Org.). *Geres 2005*: razões e resultados de uma pesquisa longitudinal sobre a eficácia escolar. Rio de Janeiro: Wallprint, 2011.

CARVALHO, Sandro Sacchet de. *Um estudo do impacto das políticas de não retenção sobre o desempenho acadêmico dos alunos nas escolas públicas brasileiras*. 2009. Tese (Doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

COOK, Thomas D.; CAMPBELL, Donald Thomas. *Quasi-experimentation*: design and analysis for field settings. Chicago: Rand McNally, 1979.

CRAHAY, Marcel. É possível tirar conclusões sobre os efeitos da repetência?. *Cadernos de* Pesquisa, São Paulo, v. 36, n. 127, p. 223-246, jan./abr. 2006.

FLETCHER, Philip. A repetência no ensino de 1º grau: um problema negligenciado da educação brasileira. Uma análise preliminar e sugestão de avaliação adicional. *Revista Brasileira de Administração da Educação*, Porto Alegre, v. 3, n.1, p. 10-41, jan./jun. 1985.

FLETCHER, Philip; CASTRO, Cláudio de Moura. Os mitos, as estratégias e as prioridades para o ensino de 1º grau. *Educação e Realidade*, Porto Alegre, v. 11, n. 1, p. 35-42, jan./jun. 1986.

FLETCHER, Philip; RIBEIRO, Sérgio da Costa. O ensino de primeiro grau no Brasil de hoje. *Em Aberto*, Brasília, ano 6, n .33, jan./mar. 1987.

GADOTTI, Moacir; ROMÃO, José Eustáquio Romão. *Evolução do ensino fundamental no Brasil*: análise das estatísticas e indicadores educacionais. Instituto Paulo Freire, 2000. Disponível em: <http://acervo.paulofreire.org/xmlui/handle/7891/3389>. Acesso em: 07 jan. 2013.

GIL, Natália. Interpretação das estatísticas de educação: um espaço de disputas simbólicas. *Revista Brasileira de História da Educação*, n. 13, p. 121-151, jan./abr. 2007.

HOLMES, C. Thomas. Grade-level retention effects: a meta-analysis of research studies. In: SHEPARD, L.A.; SMITH, M.L. (Ed.), *Flunking grades:* research and policies on retention. London, Bristol: The Falmer, 1989. p. 16-33.

JACKSON, Gregg. The research evidence on the effects of grade retention. *Review of Educational Research*, v. 45, n. 4, p. 613-636, 1975.

KLEIN, Ruben; RIBEIRO, Sérgio da Costa. *O censo educacional e o modelo de fluxo*: o problema da repetência. São Paulo: USP/NUPES, nov. 1991.

_____. A pedagogia da repetência ao longo das décadas. *Ensaio: Avaliações e Políticas Públicas em Educação*, Rio de Janeiro, v. 3, n. 20, p. 55-61, jul./set. 1998.

LOURENÇO FILHO, Manoel Bergström. A evasão escolar no ensino primário. *Revista Brasileira de Estatística*, Rio de Janeiro, n. 7, p. 539-552, jul./set. 1941.

_____. *Tendências da educação brasileira*. 2. ed. Brasília, DF: MEC/Inep, 2002.

LUZ, Luciana Soares. *O impacto da repetência na proficiência escolar*: Uma análise longitudinal do desempenho de repetentes em 2002-2003. 2008. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Cedeplar, Belo Horizonte, 2008.

MENEZES-FILHO, Naércio et al. *Avaliando o impacto da progressão continuada nas taxas de rendimento e desempenho escolar do Brasil*. São Paulo: bmec-SP; Fea-USP, 2008. Disponível em: <http://www.sebh.ecn.br/seminario_6/sebh_artigo_Roberta.pdf>. Acesso em: 27 dez. 2012.

NEVES, Rafael Correia das; PAZELLO, Elaine Toldo. O efeito de políticas de não repetência sobre o desempenho dos estudantes do ensino fundamental. In: ENCONTRO NACIONAL DE ECONOMIA, 40., 2012, Porto de Galinhas. *Anais...* Porto de Galinhas: Anpec, 2012. Disponível em: < http://www.anpec.org.br/encontro/2012/inscricao/files_I/i11-8f8c89f27396496a52604c0a6e6309dd.pdf>. Acesso em: 30 jan. 2013.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. Quando os estudantes repetem um ano ou são transferidos da escola: o que isso significa para os sistemas de educação?. *Pisa em Foco*, n. 6, jul. 2011. Disponível em: <http://www.pisa.oecd.org>. Acesso em: 30 dez. 2012.

RIANI, Juliana de Lucena Ruas; SILVA, Vânia Candida da; SOARES, Tufi Machado. *Repetir ou progredir?* Uma análise nas escolas públicas de Minas Gerais. Belo Horizonte: Cenpec, 2010. Disponível em: <http://www.cenpec.org.br>. Acesso em: 30 jan. 2013.

RIBEIRO, Sérgio da Costa. A pedagogia da repetência. *Estudos Avançados*, São Paulo, v. 5, n. 12, maio/ago. 1991.

SOARES, Sergei Suarez Dillon. A repetência no contexto internacional: o que dizem os dados de avaliações das quais o Brasil não participa. Brasília: Instituto de Pesquisa Econômica Aplicada, 2007. (Texto para Discussão, n. 1300).

TEIXEIRA DE FREITAS, Mario Augusto. O que dizem os números sobre o ensino primário. In: LOURENÇO FILHO, Manuel Bergströn (Org.). *Bibliotheca da Educação*. v. 27. São Paulo: Melhoramentos, 1937.

_____. Ainda a evasão escolar no ensino primário brasileiro. *Revista Brasileira de estatística*, Rio de Janeiro, v. 2, n. 7, p. 553-642, jul./set. 1941.

ERISSON VIANA CORREA
Doutorando do Programa de Pós-Graduação em Educação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Integrante do Laboratório de Avaliação da Educação (LAEd) da Puc-Rio. Orientador Pedagógico do Colégio Pedro II, Rio de Janeiro, Rio de Janeiro, Brasil
erisson@cp2.g12.br

ALICIA BONAMINO
Professora Doutora do Departamento de Educação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Coordenadora do Laboratório de Avaliação da Educação (LAEd) da PUC-Rio, Rio de Janeiro, Rio de Janeiro, Brasil
alicia@puc-rio.br

TUFI MACHADO SOARES
Professor Doutor do Departamento de Estatística e do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora (UFJF). Coordenador de Pesquisa do Centro de Políticas Públicas e Avaliação da Educação (CAEd) da UFJF, Juiz de Fora, Minas Gerais, Brasil
tufi@caed.ufjf.br

Recebido em: JULHO 2014
Aprovado para publicação em: DEZEMBRO 2014

OS QUESTIONÁRIOS DO SAEB: O QUE ELES REALMENTE MEDEM?

CAMILA AKEMI KARINO
LUÍS GUSTAVO DO AMARAL VINHA
JACOB ARIE LAROS

RESUMO
O presente estudo investigou as evidências de validade de construto dos questionários do Saeb de 2009. Depois de um processo de tratamento de dados e análise exploratória, foram analisadas as respostas de 1.906.616 alunos do 9º ano do ensino fundamental, de 77.905 professores e de 27.292 diretores, além das informações relacionadas às 27.045 escolas de todo Brasil. Análise Fatorial foi utilizada para identificar os construtos mensurados pelos questionários. Os resultados indicaram a existência de quatro fatores no questionário de escola, seis fatores no questionário de diretor, seis fatores no questionário de professor e quatro fatores no questionário de aluno. Uma reflexão crítica é necessária para avaliar se os construtos mensurados pelos questionários do Saeb são aqueles que realmente importam em relação ao desempenho escolar.

PALAVRAS-CHAVE QUESTIONÁRIOS CONTEXTUAIS • AVALIAÇÃO DA EDUCAÇÃO • SAEB • VALIDADE DE CONSTRUTO.

RESUMEN

El presente estudio investigó las evidencias de validez de constructo de los cuestionarios del Saeb del año 2009. Después del proceso de análisis exploratorio, fueron investigadas las respuestas de 1.906.616 alumnos del 9º año de enseñanza fundamental, de 77.905 profesores y de 27.292 directores, fuera las informaciones relativas a las 27.045 escuelas de todo el Brasil. La técnica de análisis factorial fue utilizada para identificar los constructos medidos por los cuestionarios. Los resultados revelaron a existencia de cuatro factores en el cuestionario de la escuela, seis factores en el cuestionario del director, seis factores en el cuestionario del profesor y cuatro factores en el cuestionario del alumno. Una reflexión crítica se hace necesaria para evaluar si los constructos medidos por los cuestionarios del Saeb son los que realmente interesan en relación al desempeño escolar.

PALABRAS CLAVE CUESTIONARIOS CONTEXTUALES • EVALUACIÓN EDUCACIONAL • SAEB • VALIDEZ DE CONSTRUCTO.

ABSTRACT

The present study investigated the evidence of construct validity of the Saeb questionnaires of 2009. After a process of data treatment and exploratory analysis, further analysis of the responses was carried out. The responses analyzed were from 1,906,616 students in the ninth grade, as well as of 77,905 teachers and 27,292 school directors, in addition to information related to 27,045 schools around Brazil. Factor analysis was used to identify the constructs measured by the questionnaires. The results indicated the existence of four factors in the school questionnaire, six factors in the director questionnaire, six factors in the teacher questionnaire and four factors in student questionnaire. Critical reflection is needed to assess whether the constructs measured by the Saeb questionnaires are those that really matter to school performance.

KEYWORDS CONTEXTUAL QUESTIONNAIRES • EDUCATION EVALUATION • SAEB • CONSTRUCT VALIDITY.

INTRODUÇÃO

A promoção de educação a todos e de qualidade é objetivo de um país que busca ascensão. Nesse contexto, tornam-se essenciais o conhecimento das escolas que precisam de intervenção, a identificação de boas práticas e políticas públicas e a geração de estudos acerca dos fatores intervenientes na promoção de uma educação de qualidade. E o cerne da obtenção dessas informações está na realização periódica de boas avaliações do sistema educacional.

A avaliação tem como objetivo precípuo a geração de informação, no sentido tanto de apresentar um retrato da realidade quanto de subsidiar ações. Foi com o intuito de gerar informações sobre o desempenho e os resultados dos sistemas educativos que surgiram as avaliações externas no Brasil (PESTANA, 1998). Assim, a primeira avaliação externa instituída no Brasil, o Sistema de Avaliação da Educação Básica (Saeb), teve como principal objetivo oferecer subsídios para formulação, reformulação e monitoramento de políticas públicas que buscam a universalização do acesso e a ampliação da qualidade, da equidade e da eficiência da educação brasileira (BRASIL, 2007).

Atualmente, esse sistema é constituído por três avaliações diferentes: Avaliação Nacional da Educação Básica (Aneb – amostral); Avaliação Nacional do Rendimento Escolar (Anresc); e Avaliação Nacional da Alfabetização (ANA) (BRASIL, 2013). A partir dessa nova constituição, o Saeb passa a monitorar a educação desde o 3º ano do ensino fundamental até o 3º ano do ensino médio.

As avaliações externas são constituídas, essencialmente, por provas de avaliação do desempenho escolar e questionários que buscam avaliar fatores contextuais. Os questionários devem subsidiar a obtenção de informações acerca dos fatores contextuais que interferem na qualidade da educação e no desempenho escolar. Entretanto, que informações são essas? Quais são as informações coletadas? O que se busca medir por meio dos questionários? Diferentemente da delimitação e do aporte teórico que dão suporte à constituição das provas, não há tanta clareza sobre o que se pretende medir e o que se está medindo por meio dos questionários.

Considerando a necessidade de se discutirem os questionários do Saeb, instrumentos essenciais para a realização de uma boa avaliação, este estudo tem como objetivo analisar as evidências de validade de construto dos questionários aplicados na edição de 2009.

A VALIDADE DE CONSTRUTO DOS QUESTIONÁRIOS DO SAEB

A evidência da validade de construto é uma das principais características psicométricas de um instrumento de medida, ou seja, a evidência de que o instrumento está medindo realmente aquilo que se pretende medir. Essa característica é essencial, entre outras razões, por dois aspectos interligados à validade: utilidade e consequência do uso de um teste (URBINA, 2007).

Segundo Urbina (2007), a utilidade de um instrumento de medida está relacionada aos benefícios que ele traz para a tomada de decisões, sendo contingente na proporção em que o uso do instrumento promove uma maior precisão das inferências e decisões. Assim, a carência de evidências de validade do construto que está sendo medido tem implicações

diretas acerca da utilidade de um instrumento de medida e, na ausência completa de evidências de validade, o teste torna-se inútil. Em conjunto à utilidade de um instrumento, deve-se pensar nas consequências individuais e sociais que o teste proporciona. A avaliação está relacionada a um julgamento de valor e se o termômetro desse julgamento não estiver medindo a essência do que se pretende medir, consequências negativas podem emergir.

Há pouca literatura acerca do que se pretende medir por meio dos questionários do Saeb e o que realmente eles estão medindo. Na maior parte das vezes, há informações gerais sobre o que os questionários buscam identificar: fatores contextuais; nível socioeconômico; aspectos culturais; etc. Todavia, verifica-se carência de uma definição clara e objetiva de cada um desses fatores, uma vez que eles podem representar uma série de informações. Há uma carência ainda maior de estudos que apresentem evidências de validade desses questionários.

Pestana (1998) menciona que os instrumentos do Saeb buscam avaliar quatro eixos: eficiência no ensino (medida por meio de provas de avaliação do desempenho); contexto (que engloba nível socioeconômico, perfil e autonomia das escolas); processo (que envolve planejamento e projeto pedagógico); e insumos (que incluem infraestrutura, instalações e equipamentos). Essa definição parece apontar algumas diretrizes a serem seguidas na construção dos instrumentos, mas ainda necessita de concepções teóricas mais claras e atualizadas.

O melhor aporte teórico acerca do que se pretende medir com os questionários do Saeb foi encontrado no documento *Saeb 2001: Novas perspectivas* (BRASIL, 2001) e no artigo de Franco *et al.* (2003). Nesse artigo, relata-se que a definição dos construtos a serem avaliados nos questionários teve como base teorias e resultados de pesquisas educacionais, especificamente: Lee, Bryk e Smith (1993); Willms (1992); Sammons, Hillman e Mortimore (1995); Mello (1994) e Brasil (1998). Com base nesses estudos, buscou-se priorizar na construção dos questionários do Saeb 2001:

- seis construtos relacionados ao aluno – características sociodemográficas, capital social, capital cultural, motivação e autoestima, práticas de estudo e trajetória escolar;
- sete construtos relacionados à sala de aula – características sociodemográficas do professor, formação do professor, experiência profissional, condições de trabalho, estilo pedagógico, expectativa e miscelânea (que englobaria presença e rotatividade de professor e turno);
- treze construtos relacionados à escola – características sociodemográficas do diretor, formação do diretor, experiência do diretor, liderança, condições de trabalho do diretor e da equipe, trabalho colaborativo, organização do ensino e políticas de promoção, clima acadêmico, clima disciplinar, recursos pedagógicos, situação das instalações e equipamentos, atividades extracurriculares e miscelânea (que englobaria itens como admissão de alunos, critérios de formação de turmas, acesso aos resultados do Saeb, etc.).

Os autores esclarecem que esse referencial conceitual foi o melhor que pôde ser feito naquele momento e indicam a necessidade de investigações futuras que possibilitem aprimorar os instrumentos tanto em termos de linguagem quanto na definição de alguns construtos. Apesar de fazer referência aos instrumentos de 2001, o artigo ainda é atual considerando-se que pouco foi alterado nos questionários e que não há publicação desde então acerca da concepção teórica dos questionários.

Diante da relevância do Saeb no contexto brasileiro, algumas pesquisas já foram realizadas a fim de investigar a validade de construto de seus questionários. Em uma das pesquisas, por exemplo, foi feita uma análise fatorial utilizando os 163 itens relacionados à escola dos diferentes questionários do Saeb 2001 – Escola, Diretor, Turma, Professor e Aluno (RAUTER, 2003; RAUTER; LAROS; JESUS, 2007). Nessa pesquisa, foram encontrados 19 fatores: instalações físicas; liderança do diretor; recursos materiais e didáticos;

comprometimento dos professores (avaliado pelo professor); comprometimento dos professores (avaliado pelo diretor); atitude positiva do alunado; clima disciplinar; formação continuada do diretor; experiência do diretor; TV e vídeo; segurança; vídeos didáticos; reforço escolar; recursos financeiros e pedagógicos; colaboração entre professores; equipe pedagógica; políticas contra repetências e abandono; estabilidade e assiduidade dos professores; e não ocorrência de roubos e depredações.

Todavia, é importante destacar que os autores mencionam as fragilidades psicométricas existentes em vários desses fatores. Além disso, Rauter (2003) aponta características dos questionários que dificultam a Análise Fatorial e que indicam problemas na construção dos questionários. Assim, a autora identificou como principais problemas dos questionários do Saeb 2001: a falta de simplicidade na formulação das perguntas (a questão deve ser formulada da forma mais objetiva possível – frases longas e detalhamentos desnecessários cansam a leitura), e a falta de precisão na formulação das perguntas (a questão deve medir apenas um único atributo). A autora observou também que os questionários do Saeb são muito extensos, o que pode comprometer a disposição para responder ao instrumento.

Outros estudos não tinham como foco principal a validade de construto dos questionários, mas buscavam encontrar evidências para fatores específicos, que, em geral, são utilizados em análises multinível para verificar o efeito escola. Andrade (2005), a fim de realizar uma análise dos fatores associados ao desempenho, procurou reduzir as variáveis dos questionários do Saeb 2001 a 11 fatores: nível socioeconômico; capital social; recursos culturais disponíveis em casa; experiência profissional dos professores; liderança do diretor; condições de trabalho do diretor e da equipe; trabalho colaborativo; clima acadêmico; clima disciplinar; recursos pedagógicos; e situações das instalações e equipamentos da escola.

Como aporte teórico, é importante também buscar referência nos questionários utilizados em avaliações externas estrangeiras e internacionais, tal como no Naep (Avaliação

Nacional de Progresso Educacional) e no Pisa (Programa Internacional de Avaliação de Estudantes). No primeiro – um sistema de avaliação de abrangência nacional nos Estados Unidos –, são utilizados questionários aplicados a alunos, professores e escolas. Nos questionários dos alunos, são coletadas informações demográficas, nível socioeconômico, experiências em sala de aula e suporte educacional. Os professores são avaliados quanto a formação, treinamento e práticas pedagógicas. Por meio dos questionários das escolas, são avaliadas algumas características, tais como corpo discente, absenteísmo e evasão, corpo docente, práticas pedagógicas, recursos e envolvimento de pais e da comunidade (<http://nces.ed.gov/nationsreportcard/bgquest.aspx>).

Já o Pisa é uma iniciativa de avaliação internacional comparada que envolve estudantes de dezenas de países e economias. Na edição de 2009, o questionário respondido pelos estudantes buscava informações sobre características individuais, contexto familiar e recursos das casas, clima escolar, duração das aulas, acesso e uso de bibliotecas, gosto por leitura e estratégias de leitura e compreensão (essa edição tinha como domínio principal a leitura). O questionário das escolas visava avaliar a estrutura e organização, corpo docente e discente, recursos, currículos e avaliações, clima escolar, práticas e políticas educacionais, características do diretor ou responsável (ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, 2012).

MÉTODO

PARTICIPANTES

Este estudo utilizou dados secundários coletados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) relativos à edição do Saeb realizada em 2009. Foram analisadas as respostas de estudantes, professores e diretores do 9º ano do ensino fundamental de todo o país, além daquelas relacionadas às escolas. Foram consideradas escolas das diferentes dependências administrativas (federais, estaduais e municipais) e localizadas tanto nas cidades quanto no meio rural.

A base de dados contém informações de 2.816.856 alunos, variando de 13 a 1.405 alunos por escola. Das 32.343 escolas envolvidas no levantamento, 38,4% são estaduais e 62,5% municipais; 83,4% são urbanas. Na distribuição entre as regiões, 9,5% das escolas localizam-se no Norte, 34,3% no Nordeste, 33,1% no Sudeste, 16% no Sul e 7,1% no Centro-Oeste.

Também são disponibilizadas informações sobre professores e diretores. Neste estudo, foram utilizados os dados de 127.291 professores, dos quais 71,6% são mulheres e 67,9% possuem idade entre 30 e 49 anos. Entre os 32.344 diretores, 75,5% são mulheres e 67,1% têm idade entre 30 e 49 anos.

INSTRUMENTOS

Os instrumentos utilizados foram os questionários de contexto do Saeb. O questionário respondido pelos alunos é composto por 47 questões relacionadas, de modo geral, com o perfil do aluno, estrutura e itens presentes nas casas, escolaridade dos pais ou responsáveis, hábitos de leitura e incentivo dos pais aos estudos, trajetória e práticas escolares.

Os questionários de professores e diretores são compostos por 123 e 134 questões, respectivamente, e avaliam as condições de trabalho, práticas pedagógicas e de gestão, formação e experiência profissional, além de características sociodemográficas. As informações relacionadas com a infraestrutura das escolas são obtidas por meio do questionário de escola, composto por 66 itens, em que são avaliadas características como infraestrutura, estado de conservação e segurança (BRASIL, 2009).

PROCEDIMENTOS

Os questionários foram respondidos pelos participantes no mesmo dia em que foram aplicados os testes de conhecimento aos alunos. Os professores das disciplinas avaliadas, Língua Portuguesa e Matemática, e os diretores das escolas receberam os questionários antes do início das provas, devolvendo-os preenchidos ao final destas. Os alunos responderam ao questionário logo após a realização dos testes. Os questionários das escolas foram preenchidos por profissionais contratados pelo Inep.

ANÁLISE DE DADOS

Os seguintes procedimentos de análise foram utilizados:

- análise exploratória dos dados – para verificação da integridade das bases de dados e da qualidade das respostas, foi realizada análise de dados *missing* e de normalidade univariada. A opção *pairwise* foi utilizada na Análise Fatorial para tratamento dos dados *missing*;
- Análise Fatorial – teve como base a Análise Fatorial dos eixos principais (*Principal Axis Factoring, PAF*), utilizando rotação *Promax*. Na Análise Fatorial foram excluídos itens que apresentaram cargas elevadas em mais de um fator e itens com carga fatorial no fator principal menor que 0,30 (LAROS, 2012; TABACHNICK; FIDELL, 2007);
- análise de fidedignidade – foi utilizado o coeficiente lambda 2 (λ_2) de Guttman para estimar a fidedignidade dos fatores. Também considerou-se a correlação item-resto. Itens com correlação baixa entre o item e a escala total foram excluídos. Coeficientes de fidedignidade acima de 0,70 são considerados satisfatórios para a maioria das pesquisas básicas (KAPLAN; SACCUZZO, 2001; GEORGE; MALLERY, 2003).

As análises foram realizadas utilizando o *software* estatístico *SPSS (Statistical Package for the Social Sciences)*, versão 18.

RESULTADOS

ANÁLISE EXPLORATÓRIA DOS DADOS

Durante as análises de consistência da base de dados, foram verificados percentuais altos de dados omissos. Dois tratamentos foram aplicados a todos os questionários: exclusão das variáveis com mais de 10% de dados ausentes, e exclusão de casos com mais de 70% de dados ausentes. A Tabela 1 apresenta um resumo das características gerais dos questionários.

TABELA 1 – Características dos questionários do Saeb 2009 – 9º ano

QUESTIONÁRIO	TAMANHO DA BASE DE DADOS ORIGINAL	TAMANHO DA BASE DE DADOS FINAL	TOTAL DE VARIÁVEIS	Nº DE VARIÁVEIS COM MAIS DE 10% DE MISSING	Nº DE VARIÁVEIS CONSIDERADAS NA ANÁLISE FATORIAL
Escola	32.002	27.045	67	15	51*
Professor	127.291	77.905	121	68	53
Diretor	32.344	27.292	134	48	86
Aluno	2.817.851	1.906.616	47	9	38

Fonte: Elaboração dos autores.

Nota (*): Uma variável foi retirada da Análise Fatorial, pois consistia de uma percepção geral do avaliador acerca das condições da escola.

Nota-se que os questionários de escola e de aluno foram mais bem respondidos, uma vez que apresentam percentuais menores de variáveis com mais de 10% de dados ausentes (22% e 19%, respectivamente). Por outro lado, sobretudo no questionário de professor, destaca-se a quantidade de variáveis com mais de 10% de dados ausentes. A análise qualitativa das variáveis com percentual elevado de dados ausentes indicou problemas estruturais na formulação das perguntas (falta de simplicidade e de clareza, problemas nas opções de respostas, etc.), bem como a existência de perguntas comprometedoras ou passíveis de desejabilidade social. Além disso, esses resultados confirmam a existência de dificuldades na aplicação dos questionários de professor e diretor, seja pela logística adotada na coleta dos dados, seja pela falta de motivação dos agentes em responder aos questionários.

É preciso mencionar também que outras variáveis foram retiradas, pois tinham respostas qualitativas que não poderiam ser usadas numa análise quantitativa. Por exemplo, no questionário de diretor é perguntado como ele assumiu a direção da escola, se foi por processo seletivo, eleição ou indicação. Foram excluídas também variáveis como sexo, idade e etnia autodeclarada. Na análise do questionário de aluno, especificamente, não foram consideradas questões como gosto por Matemática e Língua Portuguesa, pretensões após o término do ensino fundamental, tipo de escola em que estudou, reprovação e abandono.

Considerando a relevância de uma boa qualidade da resposta nas análises subsequentes, julgou-se que seria melhor eliminar as variáveis com percentual acima de 10% de dados

ausentes. Dessa forma, optou-se por manter nos questionários somente as variáveis que foram adequadamente respondidas para, a partir delas, buscar definir o que parece medir os questionários do Saeb 2009.

Depois da exclusão das variáveis com mais de 10% de dados ausentes, a maioria das variáveis apresentava percentuais em torno de 2% de dados ausentes. No caso do questionário de escola, duas variáveis apresentaram excepcionalmente percentuais em torno de 5% e 8% de dados ausentes: *quantidade-computadores* e *estrutura-biblioteca*. Verificou-se que, na primeira variável, o avaliador externo precisava indicar o número total de computadores na escola, mas não havia opção a ser marcada caso a escola não tivesse computador (n=0). Logo, provavelmente o percentual de dados ausentes agregava também as escolas que não possuíam computador.

Para tratar essa variável, utilizou-se outra variável do questionário de escola que avaliava o estado de conservação dos computadores (*conservação equipamentos-computador*), na qual havia a opção "inexistente". Assim, nas escolas em que opção "inexistente" estava preenchida na variável *conservação equipamentos-computador*, os dados omissos da variável *quantidade-computadores* foram tratados como ausência de computador em vez de dado ausente. Após esse tratamento, restaram na base de dados somente 2,6% de respostas ausentes na variável *quantidade-computadores*.

Já a variável *estrutura-biblioteca* foi constituída a partir da inferência de uma variável do questionário de escola, pois não havia uma pergunta que avaliava diretamente a existência de uma biblioteca. Assim, a partir da pergunta: "Qual o percentual de alunos que utilizam a biblioteca durante o mês?", desdobraram-se duas variáveis: frequência de uso da biblioteca pelos alunos e existência de biblioteca na escola, considerando que uma das opções de resposta era "não há biblioteca na escola". A primeira variável foi eliminada das análises por apresentar mais de 10% de dados ausentes, enquanto a segunda, denominada *estrutura-biblioteca*, permaneceu nas análises mesmo com 8% de dados ausentes. Optou-se por manter a variável devido à relevância do quesito biblioteca para a infraestrutura escolar.

Foram ainda realizados outros tratamentos a fim de se manter o maior número de variáveis possível. No questionário de diretor, um dos aspectos abordados está relacionado à ocorrência de agressões, questionando o tipo de agressão (física ou verbal), quem foi agredido (professor, funcionário ou aluno) e quem foi o agressor (professor, funcionário ou aluno), o que possibilitava um total de 18 respostas. De modo a diminuir o percentual de dados ausentes, foram considerados apenas os dois tipos de agressões para as três vítimas, independentemente do agressor. Já no questionário de aluno, uma variável única foi criada para relatar a existência de *freezer* em casa. Essa variável assumiu o valor 1 se o aluno respondeu que na sua casa tem *freezer* junto com a geladeira ou *freezer* separado da geladeira.

Após o tratamento das variáveis com quantitativo alto de dados ausentes, foi verificado se havia na base de dados casos (escolas, diretores, professores e alunos, a depender do questionário) com menos de 70% das questões respondidas. O intuito dessa análise era retirar os casos com muitas respostas faltantes. O questionário que apresentou maior redução no número de casos com base nesse critério foi o de aluno. De todo modo, considerando-se o tamanho da base de dados inicial, a exclusão dos casos com menos de 70% de variáveis respondidas não compromete as análises posteriores e permite manter uma base com qualidade melhor das respostas.

ANÁLISE FATORIAL E ANÁLISE DE CONSISTÊNCIA INTERNA

O número inicial de fatores foi determinado pela análise do *screeplot*, do percentual de variância explicada por cada fator (acima de 2%) e pelos autovalores (acima de 1) a partir de uma análise de componentes principais. Em seguida, executou-se uma sequência de análises fatoriais, no sentido de buscar uma solução considerando os fatores, as cargas fatoriais dos itens e a coerência de conteúdo dos itens e fatores.

As tabelas 2 a 5 apresentam as cargas fatoriais, as correlações item-resto e a fidedignidade por fator das análises dos questionários de escola, professor, diretor e aluno, respectivamente.

TABELA 2 – Cargas fatoriais depois da rotação Promax (CF), correlações item-resto (r_{it}), fidedignidade dos fatores e a correlação entre os fatores do questionário de escola

FATOR 1 – CONDIÇÃO DA ESTRUTURA		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Sala	0,80	0,72
Parede	0,74	0,67
Piso	0,74	0,65
Estrutura elétrica	0,68	0,65
Telhado	0,67	0,57
Hidráulica	0,67	0,68
Entrada da escola	0,66	0,64
Portas	0,66	0,66
Corredor	0,65	0,62
Banheiro	0,64	0,67
Pátio	0,61	0,59
Cozinha	0,60	0,60
Janelas	0,59	0,61
Sala de aula arejada	0,40	0,40
Iluminação da sala de aula	0,35	0,36
Média (15 itens)	0,63	0,61
Fidedignidade	0,91	

FATOR 2 – CONSERVAÇÃO DOS EQUIPAMENTOS		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Telefone	0,68	0,61
Fax	0,63	0,56
Quantidade de computadores	0,62	0,52
Impressora	0,58	0,49
Retroprojetor	0,56	0,49
Computador	0,52	0,45
Som	0,49	0,44
Máquina copiadora	0,47	0,42
Projetor de *slides*	0,43	0,40
Equipamentos contra incêndio	0,42	0,43
Videocassete	0,40	0,38
Televisão	0,39	0,36
Média (12 itens)	0,52	0,46
Fidedignidade	0,80	

FATOR 3 – DEPREDAÇÃO		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Dependências internas	0,75	0,67
Paredes	0,70	0,61
Dependências externas	0,69	0,61
Banheiro	0,64	0,58
Pichação nos muros	0,63	0,51
Sinais de depredação	0,46	0,45
Média (6 itens)	0,65	0,57
Fidedignidade	0,81	

FATOR 4 – SEGURANÇA		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Vigia noite	0,90	0,71
Vigia fim de semana	0,73	0,62
Vigia dia	0,53	0,47
Média (3 itens)	0,72	0,60
Fidedignidade	0,76	

Correlação entre fatores	F1-F2 = 0,43	F1-F3 = -0,42	F1-F4 = 0,06	F2-F3 = -0,09	F2-F4 = -0,08	F3-F4 = -0,01

Fonte: Elaboração dos autores.

Na análise do questionário de escola, quatro fatores foram encontrados (Tabela 2). Nota-se que a maior parte dos itens dos fatores, sobretudo dos dois primeiros, possui cargas fatoriais altas e correlações item-resto moderadas. Os coeficientes de fidedignidade dos quatro fatores identificados são satisfatórios. Esses resultados indicam uma boa estrutura fatorial e consistência interna adequada dos fatores. Vale destacar a existência de uma correlação moderada e positiva entre o fator 1 e o fator 2 e uma correlação moderada e negativa entre o fator 1 e o fator 3. Ou seja, quanto mais positiva a percepção da estrutura, mais positiva tende a ser a percepção da qualidade dos equipamentos da escola, e menor a percepção de sinais de depredação.

TABELA 3 – Cargas fatoriais depois da rotação Promax (CF), correlações item-resto (r_{it}) e fidedignidade dos fatores do questionário de professor

FATOR 1 – OCORRÊNCIA DE VIOLÊNCIA

DESCRIÇÃO DO ITEM	CF	R_{IT}
Aluno com porte de arma de fogo	0,93	0,84
Atentado à vida	0,93	0,84
Professor vítima de roubo	0,92	0,82
Agressão física	0,92	0,83
Aluno com porte de arma branca	0,84	0,77
Professor vítima de furto	0,82	0,76
Alunos sob efeito de álcool	0,74	0,69
Professor ameaçado por aluno	0,73	0,70
Alunos sob efeito de drogas ilícitas	0,70	0,66
Agressão Verbal	0,43	0,39
Média (10 itens)	0,80	0,73
Fidedignidade	0,92	

FATOR 2 – ATUAÇÃO DO DIRETOR

DESCRIÇÃO DO ITEM	CF	R_{IT}
Professor confia no diretor	0,88	0,80
Diretor motiva para o trabalho	0,86	0,79
Diretor estimula ativ. inovadoras	0,84	0,78
Comprometimento diretor (prof.)	0,79	0,75
Média (4 itens)	0,84	0,78
Fidedignidade	0,90	

FATOR 3 – USO DE MATERIAIS NAS PRÁTICAS PEDAGÓGICAS

DESCRIÇÃO DO ITEM	CF	R_{IT}
Computador	0,66	0,49
Vídeo-DVD	0,65	0,51
Internet	0,63	0,46
Revistinhas	0,46	0,41
Livros de literatura	0,44	0,38
Jornais	0,44	0,39
Média (6 itens)	0,55	0,44
Fidedignidade	0,73	

FATOR 4 – CAUSAS PROBLEMAS DE APRENDIZAGEM: ESCOLA

DESCRIÇÃO DO ITEM	CF	R_{IT}
Sobrecarga trab. dos professores	0,55	0,42
Baixo salário dos professores	0,52	0,41
Carência de infraestrutura física/ped.	0,52	0,41
Poucas oportunidades - desen. aluno	0,50	0,41
Insegurança física da escola	0,45	0,35
Conteúdos curriculares inadequados	0,41	0,32
Não cumprimento conteúdo curricular	0,38	0,30
Média (7 itens)	0,47	0,37
Fidedignidade	0,68	

FATOR 5 – EXPERIÊNCIA E CONDIÇÕES DE TRABALHO

DESCRIÇÃO DO ITEM	CF	R_{IT}
Professor-salário bruto	0,75	0,58
Professor-tempo que está formado	0,61	0,38
Professor-tempo lecionando	0,46	0,30
Professor-carga horária	0,43	0,38
Média (4 itens)	0,56	0,41
Fidedignidade	0,65	

FATOR 6 – FORMAÇÃO DO PROFESSOR

DESCRIÇÃO DO ITEM	CF	R_{IT}
Nível de escolaridade	0,66	0,40
Modalidade do curso superior	0,64	0,43
Natureza da instituição em que formou	0,56	0,36
Tipo de escola do curso superior	0,47	0,41
Média (4 itens)	0,58	0,40
Fidedignidade	0,58	

Correlação entre fatores	F1-F2 = -0,04	F1-F3 = -0,02	F1-F4 = -0,03	F1-F5 = 0,01	F1-F6 = 0,03
	F2-F3 = 0,20	F2-F4 = 0,32	F2-F5 = -0,02	F2-F6 = -0,04	F3-F4 = 0,22
	F3-F5 = 0,19	F3-F6 = 0,09	F4-F5 = 0,10	F4-F6 = 0,00	F5-F6 = 0,18

Fonte: Elaboração dos autores.

Os resultados da Análise Fatorial do questionário de professor (Tabela 3) demonstram, de modo geral, coerência de conteúdo dos itens com os fatores, uma estrutura fatorial aceitável e uma consistência interna razoável dos fatores. Os fatores 4, 5 e 6 apresentam consistência interna fraca (valores de λ_2 entre 0,58 e 0,68), mas optou-se por mantê-los porque esses resultados podem auxiliar no aprimoramento do questionário para aplicações futuras.

No questionário de professor, das 15 correlações entre fatores, nove têm valor menor que 0,10. Nota-se, portanto, que em geral as correlações são baixas. Todavia, considerando que se trata de uma área pouco pesquisada, optou-se por apresentar as correlações, pois podem contribuir para o refinamento do instrumento. O mesmo ocorreu para os questionários de diretor e aluno.

TABELA 4 – Cargas fatoriais depois de rotação Promax (CF), correlações item-resto (r_{it}), fidedignidade dos fatores e a correlação entre os fatores do questionário de diretor

FATOR 1– RECURSOS		
DESCRIÇÃO DO ITEM	**CF**	**R_{IT}**
Internet para professores	0,78	0,64
Computadores (professores)	0,75	0,61
Internet para alunos	0,73	0,58
Computadores para alunos	0,66	0,53
Computadores (administrativo)	0,58	0,54
Impressora	0,56	0,52
Telefone	0,53	0,59
Fax	0,50	0,54
Biblioteca	0,49	0,50
Aparelho de som	0,48	0,49
Retroprojetor	0,47	0,50
Fitas de vídeo ou DVD (*laser*)	0,46	0,46
Quadra de esportes	0,44	0,47
Vídeo cassete ou DVD	0,43	0,42
Fita de vídeo/DVD (educativo)	0,41	0,41
Laboratório	0,41	0,38
Televisor	0,41	0,40
Máquina copiadora	0,39	0,38
Projetor de *slides*	0,37	0,37
Média (19 itens)	0,52	0,49
Fidedignidade	0,88	

FATOR 3 – EXPERIÊNCIA DO DIRETOR		
DESCRIÇÃO DO ITEM	**CF**	**R_{IT}**
Tempo como diretor	0,91	0,65
Tempo de direção na atual escola	0,79	0,59
Tempo da escolaridade atual	0,40	0,40
Tempo de trabalho com educação	0,39	0,43
Média (4 itens)	0,62	0,52
Fidedignidade	0,74	

FATOR 4 – FORMAÇÃO		
Descrição do Item	CF	r_{it}
Área do título	0,72	0,63
Título	0,71	0,65
Forma de realização do curso	0,54	0,35
Escolaridade	0,51	0,40
Média (4 itens)	0,62	0,51
Fidedignidade	0,74	

FATOR 5 – PARTICIPAÇÃO EM EVENTOS COM A COMUNIDADE		
DESCRIÇÃO DO ITEM	**CF**	**R_{IT}**
Mutirão: limpeza	0,56	0,33
Mutirão: manutenção da estr. física	0,54	0,34
Manutenção: hortas, pomar, jardins	0,41	0,28
Camp. solidariedade: comunidade	0,41	0,35
Eventos (terceiros) na escola	0,35	0,30
Eventos da comunidade na escola	0,33	0,29
Eventos da esc. para comunidade	0,32	0,28
Camp. solidariedade: escola	0,31	0,27
Média (8 itens)	0,40	0,31
Fidedignidade	0,62	

(continua)

(continuação)

FATOR 2 – VIOLÊNCIA		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Agressão verbal a alunos	0,67	0,57
Agressão verbal a funcionários	0,65	0,56
Agressão verbal a professores	0,63	0,55
Agressão física a alunos	0,62	0,54
Problemas disciplinares (alunos)	0,49	0,42
Porte de arma branca	0,37	0,34
Ação de gangues na escola	0,36	0,35
Média (7 itens)	0,54	0,48
Fidedignidade	0,77	

FATOR 6 – CARÊNCIAS RECURSOS HUMANOS		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Insuficiência de professores	0,40	0,49
Carência de pessoal administrativo	0,37	0,40
Rotatividade do corpo docente	0,33	0,40
Alto índice de faltas (professores)	0,32	0,38
Falta de apoio pedagógico	0,30	0,28
Média (5 itens)	0,34	0,39
Fidedignidade	0,64	

Correlação entre fatores	F1-F2 = 0,19	F1-F3 = 0,32	F1-F4 = 0,33	F1-F5 = 0,08	F1-F6 = -0,05
	F2-F3 = 0,07	F2-F4 = 0,23	F2-F5 = -0,08	F2-F6 = 0,02	F3-F4 = 0,27
	F3-F5 = 0,03	F3-F6 = 0,00	F4-F5 = 0,11	F4-F6 = -0,03	F5-F6 = 0,05

Fonte: Elaboração dos autores.

Nota-se que a maior parte dos itens do questionário de diretor possui cargas fatoriais e correlações item-resto moderadas e índices de fidedignidade razoáveis. Os fatores 5 e 6 apresentam consistência interna fraca (valores de λ_2 0,62 e 0,64, respectivamente).

TABELA 5 – Cargas fatoriais depois de rotação Promax (CF), correlações item-resto (r_{it}), fidedignidade dos fatores e a correlação entre os fatores do questionário de aluno

FATOR 1- NÍVEL SOCIOECONÔMICO		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Computador	0,74	0,60
Internet	0,70	0,57
Televisão em cores	0,64	0,57
Carro	0,55	0,49
Banheiro	0,53	0,48
Lavadora de roupas	0,51	0,47
Freezer	0,45	0,41
Rádio	0,44	0,40
Geladeira	0,40	0,38
DVD	0,34	0,33
Média (10 itens)	0,53	0,47
Fidedignidade	0,80	

FATOR 2 – DEVER DE CASA E CORREÇÕES		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Faz dever de Matemática	0,67	0,46
Faz dever de Língua Portuguesa	0,61	0,44
Prof. corrige dever de Matemática	0,45	0,36
Prof. corrige dever de Português	0,40	0,32
Professor elogia	0,37	0,31
Média (5 itens)	0,50	0,38
Fidedignidade	0,63	

FATOR 3 - INCENTIVO DOS PAIS		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Pais incentivam dever de casa	0,69	0,53
Pais incentivam a estudar	0,58	0,42
Pais incentivam a ler	0,53	0,45
Pais incentivam a ir às aulas	0,50	0,36
Média (4 itens)	0,57	0,44
Fidedignidade	0,65	

FATOR 4 – HÁBITO DE LEITURA DOS PAIS		
DESCRIÇÃO DO ITEM	CF	R_{IT}
Vê mãe lendo	0,70	0,38
Mãe sabe ler	0,58	0,44
Pai sabe ler	0,49	0,44
Vê pai lendo	0,45	0,44
Média (4 itens)	0,55	0,43
Fidedignidade	0,66	

Correlação entre fatores	F1-F2 = -0,09	F1-F3 = 0,06	F1-F4 = 0,34
	F2-F3 = 0,25	F2-F4 = 0,08	F3-F4 = 0,24

Fonte: Elaboração dos autores.

A Análise Fatorial do questionário de aluno aponta uma estrutura fatorial aceitável, com resultados semelhantes aos dos demais questionários: cargas fatoriais e correlações item-resto moderadas e índices de fidedignidade razoáveis (Tabela 5). Os fatores 2, 3 e 4 apresentam consistência interna fraca (valores de λ_2 entre 0,63 e 0,66).

Em todos os questionários, pode-se discutir a manutenção de alguns itens na composição dos fatores, todavia, considerando o caráter exploratório das análises e a intenção desse estudo de dar subsídios para o aprimoramento dos questionários, buscou-se manter a maior quantidade possível de itens.

Em resumo, na Análise Fatorial dos questionários do Saeb 2009, foi constatado que entre 45% e 71% dos itens não entraram na estrutura fatorial. O questionário com a maior proporção de itens excluídos foi o de professor (Tabela 6).

TABELA 6 – Resumo das variáveis excluídas nas análises por questionário

QUESTIONÁRIO	TOTAL DE VARIÁVEIS	TOTAL DE VARIÁVEIS EXCLUÍDAS NAS ANÁLISES	% DE VARIÁVEIS EXCLUÍDAS DAS ANÁLISES
Escola	66	30	45%
Professor	121	86	71%
Diretor	134	86	64%
Aluno	47	24	51%

Fonte: Elaboração dos autores.

DISCUSSÃO

Em síntese, a Análise Fatorial dos questionários do Saeb 2009 resultou em quatro fatores no questionário de escola (*qualidade da estrutura, qualidade dos equipamentos, depredação* e *segurança*), seis no de professor (*violência na escola, avaliação da atuação do diretor, uso de equipamentos nas práticas pedagógicas, experiência e condições de trabalho, fatores relativos à escola relacionados aos problemas de aprendizagem* e *formação do professor*), seis no de diretor (*recursos da escola, violência, experiência do diretor, formação do diretor* e *faltas e presença da comunidade*) e quatro no de aluno (*nível socioeconômico, lições e correções, incentivo* e *hábito de leitura dos pais*).

Muitos fatores são convergentes com aqueles encontrados por Rauter (2003), uma vez que os questionários, embora sejam de edições diferentes (2001 e 2009), mantêm muitas similaridades. Além disso, diversos fatores estão relacionados

com os eixos a serem avaliados pelos questionários segundo Pestana (1998): o eixo contexto engloba a percepção de violência de professores e diretores, a presença de equipes de segurança nas escolas e o nível socioeconômico dos alunos; o eixo de processo envolve os fatores relacionados com a percepção do professor quanto à atuação do diretor, as causas de problemas de aprendizagem atribuídos às escolas e aos alunos e o uso de equipamentos nas práticas pedagógicas, além do fator dever de casa e correções segundo a percepção dos alunos; e o eixo insumo contempla os fatores relacionados com a formação de professores e diretores, as condições de trabalho e qualidade de equipamentos e a infraestrutura.

Considerando a concepção teórica apresentada pelo Inep (BRASIL, 2001) e por Franco *et al.* (2003) acerca do que se pretende medir nos questionários do Saeb 2001, nota-se que nem todos os construtos foram contemplados nas análises realizadas neste artigo com os questionários de 2009. Entre os seis construtos relacionados ao aluno, observa-se que as características sociodemográficas são mensuradas e, no que concerne à Análise Fatorial, verifica-se que o capital social, o capital cultural e as práticas de estudo estão de certa forma contemplados. Motivação e autoestima e trajetória escolar não aparecem em nenhum dos fatores. Quanto aos construtos relacionados à sala de aula, os questionários do Saeb 2009 contemplam todos aqueles definidos em 2001, com exceção do construto miscelânea. Por fim, no que se refere aos construtos relacionados à escola, vários não são contemplados, tais como liderança, clima acadêmico, clima disciplinar, atividades extracurriculares e miscelânea (que englobaria itens como admissão de alunos, critérios de formação de turmas, acesso aos resultados do Saeb, etc.). Por outro lado, os construtos violência e segurança na escola aparecem em vários questionários, o que não estava previsto nos questionários de 2001.

Nota-se que construtos importantes estão sendo avaliados nos questionários do Saeb 2009. Todavia, ainda há fatores que precisam ser mais bem trabalhados, pois apresentam itens com baixas cargas fatoriais ou pouco itens. Vários fatores mostraram baixo coeficiente de fidedignidade: dos

20 fatores identificados, oito tinham um coeficiente de fidedignidade inferior a 0,70. Além disso, há vários fatores associados ao desempenho dos alunos relevantes para a análise e que não são contemplados.

No Naep, por exemplo, os questionários respondidos pelos alunos apresentam itens relacionados com disciplinas, práticas pedagógicas, equipamentos utilizados e percepção de dificuldade de conteúdos e provas. Os professores das disciplinas avaliadas no Naep respondem a diversos itens referentes a formas de aperfeiçoamento e práticas pedagógicas utilizadas no dia a dia, enquanto no Saeb poucos itens associam-se a esses fatores, e na análise realizada, tais itens não se agrupam em fatores. Já o Pisa utiliza escalas no questionário dos estudantes que avaliam o clima escolar, as estratégias de aprendizagem e os hábitos de leitura, além da relação com o professor e práticas pedagógicas empregadas por estes. No Pisa, as escolas também são avaliadas por meio de escalas quanto à participação de professores nas decisões e ao envolvimento do diretor com a rotina da escola, aspectos esses também pouco abordados no Saeb.

Além das avaliações educacionais estrangeiras e internacionais, há estudos na área de fatores associados que indicam importantes aspectos que devem ser investigados no contexto brasileiro, pois interferem no desempenho (ANDRADE; LAROS, 2007; SOARES, 2007, 2005). A literatura tem mostrado a importância de fatores como capital social (BONAMINO et al., 2010; GOODMAN; GREGG; WASHBROOK, 2011), capital cultural (BONAMINO et al., 2010; MARTELETO; ANDRADE, 2013), atitudes e comportamentos dos alunos e pais (GABLE; WOLF, 1993; GOODMAN; GREGG; WASHBROOK, 2011; LEE, 2013), interações entre alunos e professores e clima escolar (STEWART, 2008; KONISHI et al., 2010), características de professores (WAYNE; YOUNGS, 2003) e liderança de diretores (LOUIS; DRETZKE; WAHLSTRON, 2010).

Ainda temos que nos perguntar sobre o que fazer com os itens que não entraram na Análise Fatorial. Cerca de 50% dos itens dos questionários foram retirados das análises por apresentarem alto índice de respostas omissas ou durante a Análise Fatorial. Aparentemente esses itens não fazem parte

de um construto, ou seja, com esses itens não é possível formar medidas com vários itens (*multiple-item measures*), sendo que em geral medidas com vários itens são melhores por serem mais fidedignas e terem mais evidências de validade.

Portanto, a partir deste estudo, verificou-se que, além de uma avaliação e definição clara dos construtos que se pretende medir, é necessária uma revisão na construção das questões dos questionários, pois várias não atendem às orientações para construção de instrumentos de medidas (PERES-DOS-SANTOS; LAROS, 2007; HOGAN, 2006; PASQUALI, 1999; CLARK; WATSON, 1995). Deve-se investir tempo na padronização das escalas das respostas, na construção de orientação para os pesquisadores quanto à possibilidade de utilização das variáveis como ordinais ou intervalares e na revisão das perguntas constantes nos questionários, de modo a torná-las mais simples e claras. Com essas ações, entende-se que os questionários tornar-se-iam mais eficientes.

Também é preciso investir em outras formas de coletas das informações. O uso de questionários em papel e a limitação das respostas a poucos professores têm reduzido a participação dos mesmos e inibido respostas sinceras devido ao medo de serem identificados. Assim, sugere-se o uso de questionários *on-line* com a participação de todos os professores das escolas, e não somente daqueles das séries avaliadas. É importante ainda um trabalho de conscientização sobre a importância dos questionários.

Ainda em relação a outras formas de coleta de informações, atualmente todos os participantes respondem a um questionário específico, o que gera uma enorme quantidade respostas acerca dos mesmos construtos. Para fins de pesquisa, trabalhar com um número maior de tipos de questionários pode ser mais eficiente, pois possibilitaria a coleta de informações de um número maior de construtos e cada participante responderia a um número menor de questões. Sabe-se que questionários muito longos tendem a não ser bem respondidos.

Entre as limitações deste estudo, é preciso apontar que as análises foram concentradas nas variáveis que poderiam ser trabalhadas de forma ordinal ou intervalar. Apesar de al-

guns pesquisadores considerarem as variáveis ordinais com mais de dois valores como variáveis contínuas (PASQUALI, 2012b), sugere-se que outros estudos sejam realizados utilizando outros tipos de correlação. De todo modo, para os propósitos deste estudo, a análise fatorial empregada se mostra válida, pois permite apontar os fatores gerais que parecem ser medidos nos questionários do Saeb e gerar a reflexão sobre o que se pretende medir. O objetivo do artigo não foi criar medidas em si. Além disso, é preciso que outros estudos sejam realizados tendo como foco as variáveis nominais. Outra análise interessante que não foi realizada refere-se à verificação da reprodução dos construtos encontrados para o 9º ano nos questionários do 5º ano ou de edições posteriores do Saeb.

Por fim, espera-se que os resultados desta pesquisa possam motivar outros pesquisadores e apoiar o Inep a rever os questionários do Saeb a fim de que possam ser aprimorados. Os questionários dessa importante avaliação devem ser foco de contínuos debates sobre a pertinência dos referenciais teóricos, dos construtos almejados, das questões que compõem os questionários, dos indicativos de validade das medidas, enfim, da qualidade dos instrumentos como um todo.

REFERÊNCIAS

ANDRADE, Josemberg Moura de. *Construção de um modelo explicativo de desempenho escolar*: um estudo psicométrico e multinível com dados do SAEB. 2005. Dissertação (Mestrado em Psicologia Social e do Trabalho) – Instituto de Psicologia da Universidade de Brasília, Brasília, 2005.

ANDRADE, Josemberg Moura de; LAROS, Jacob Arie. Fatores associados ao desempenho escolar: estudo multinível com dados do SAEB/2001. *Psicologia: Teoria e Pesquisa,* v. 23, n. 1, p. 33-42, 2007.

BONAMINO, Alicia; ALVES, Fátima; FRANCO, Creso; CAZELLI, Sibele. Os efeitos das diferentes formas de capital no desempenho escolar: um estudo à luz de Bourdieu e de Coleman. *Revista Brasileira de Educação*, v. 15, n. 45, p. 487-499, 2010.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais. *SAEB 2001*: novas perspectivas. Brasília: Inep, 2001.

_____. *SAEB – 2005 primeiros resultados*: médias de desempenho do SAEB/2005 em perspectiva comparada. Brasília: Inep, 2007.

_____. *Microdados Prova Brasil 2009*. Brasília: Inep, 2009. Disponível em: <http://portal.inep.gov.br/basica-levantamentos-acessar>. Acesso em: 01 jun. 2012.

BRASIL. Ministério da Educação e do Desporto. Secretaria de Educação Fundamental. *Referenciais para formação de professores*. Documento Técnico. Brasília: SEF, 1998. 103p.

_____. Portaria n. 482, de 7 de junho de 2013. Dispõe sobre o Sistema de Avaliação da Educação Básica – SAEB. Brasília: MEC, 2013.

CLARK, Lee Anna; WATSON, David. Constructing validity: basic issues in objective scale development. *Psychological Assessment*, v. 7, n. 3, p. 309-319, 1995.

FRANCO, Creso; FERNANDES, Cristiano; SOARES, José Francisco; BELTRÃO, Kaizô; BARBOSA, Maria Eugênia; ALVES, Maria Tereza Gonzaga. O referencial teórico na construção dos questionários contextuais do Saeb 2001. *Estudos em Avaliação Educacional,* n. 28, p. 39-74, 2003.

GABLE, Robert K.; WOLF, Marian B. *Instrument development in the affective domain:* measuring attitudes and values in corporate and school settings. 2. ed. Amsterdam: Kluwer Academic, 1993.

GEORGE, Darren; MALLERY, Paul. *SPSS for Windows step by step*: a simple guide and reference. 11.0 update. 4. ed. Boston: Allyn & Bacon, 2003.

GOODMAN, Alissa; GREGG, Paul; WASHBROOK, Elizabeth. Children's educational attainment and behaviours of parents and children through childhood. *Longitudinal and Life Course Studies*, v. 2, n. 1, p. 1-18, 2011.

HOGAN, Thomas P. *Introdução à prática de testes psicológicos.* Rio de Janeiro: LTC, 2006.

KAPLAN, Robert M.; SACCUZZO, Dennis P. *Psychological testing*: principles, applications and issues. 5. ed. Belmont, CA: Wadsworth, 2001.

KONISHI, Chiaki.; HYMEL, Shelley.; ZUMBO, Bruno. D.; LI, Zhen. Do school bullying and student-teacher relationships matter for academic achievement? A multilevel analysis. *Canadian Journal of School Psychology*, v. 25, n. 1, p. 19-39, 2010.

LAROS, Jacob Arie. O uso da análise fatorial: algumas diretrizes para pesquisadores. In: PASQUALI, Luiz (Ed.). *Análise fatorial para pesquisadores.* Brasília: LabPAM, 2012. p. 141-160.

LEE, Jihyun. Can writing attitudes and learning behavior overcome gender difference in writing? Evidence from NAEP. *Written Communication*, v. 30, n. 2, p. 164-193, 2013.

LEE, Valerie E.; BRYK, Anthony S.; SMITH, Julia B. The organization of effective secondary schools. In: DARLING-HAMMOND, L. *Review of Research in Education*, v. 19, p. 171-267, 1993.

LOUIS, Karen Seashore; DRETZKE, Beverly; WAHLSTRON, Kyla. How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, v. 21, n. 3, p. 315-336, 2010.

MARTELETO, Leticia; ANDRADE, Fernando. The educational achievement of Brazilian adolescents: cultural capital and the interaction between families and schools. *Sociology of Education*, v. 86, n. 4, p. 1-20, 2013.

MELLO, Guiomar Namo de. Escolas eficazes: um tema revisitado. In: XAVIER, Antonio Carlos da R. et al. (Org.). *Gestão escolar*: desafios e tendências. Brasília: Ipea, 1994. p. 329-369.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *PISA 2009 Technical Report*, PISA, OECD, 2012.

PASQUALI, Luiz. *Instrumentos psicológicos*: manual prático de elaboração. Brasília: LabPAM, 1999.

_____. Rotação dos fatores. In: PASQUALI, Luiz (Ed.). *Análise fatorial para pesquisadores.* Brasília: LabPAM, 2012a. p. 85-102.

_____. A correlação e a regressão. In: PASQUALI, Luiz (Ed.). *Análise fatorial para pesquisadores.* Brasília: LabPAM, 2012b. p. 85-102.

PERES-DOS-SANTOS, Luiz Francisco Bastos; LAROS, Jacob Arie. Avaliação da prática pedagógica do professor de ensino superior. *Estudos em Avaliação Educacional*, v. 18, n. 36, p. 75-95, 2007.

PESTANA, Maria Inês. O sistema de avaliação brasileiro. *Revista Brasileira de Estudos Pedagógicos*, v. 79, n. 191, p. 65-73, 1998.

RAUTER, Raíssa. *Estrutura fatorial das questões do SAEB 2001 relacionadas a características da escola.* Brasília, 2003, 124f. Dissertação (Mestrado em Psicologia Social e do Trabalho) – Instituto de Psicologia, Universidade de Brasília, Brasília, 2003.

RAUTER, Raíssa; LAROS, Jacob Arie; JESUS, Girlene Ribeiro de. Questionários contextuais do SAEB: análise qualitativa baseada nos resultados da Análise Fatorial. *Revista Eletrônica Ibero-americana sobre Qualidade, Eficácia e Mudança em Educação – REICE*, v. 5, n. 2, p. 56-68, 2007. Disponível em: <http://www.rinace.net/reice.htm>. Acesso em: mar. 2013.

SAMMONS, Pam; HILLMAN, Josh; MORTIMORE, Peter. *Key characteristics of effective schools*: a review of school effectiveness research. London: Office for Standards in Education OFSTED, 1995.

SOARES, José Francisco. O efeito da escola no desempenho cognitivo de seus alunos. In: MELLO E SOUZA, Alberto (Ed.). *Dimensões da avaliação educacional*. Petrópolis: Vozes, 2005. p. 174-204.

_____. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de Pesquisa*, v. 37, n. 130, p. 130-160, 2007.

STEWART, Endya B. School structural characteristics, student effort, peer associations, and parental involvement the influence of school-and individual-level factors on academic achievement. *Education and Urban Society*, v. 40, n. 2, p. 179-204, 2008.

TABACHNICK, Barbara G.; FIDELL, Linda S. *Using multivariate statistics*. 5. ed. Boston: Pearson Education, 2007.

URBINA, Susana. *Fundamentos da testagem psicológica*. Porto Alegre: Artmed, 2007.

WAYNE, Andrew J.; YOUNGS, Peter. Teacher characteristics and student achievement gains: a review. *Review of Educational Research*, v. 73, n. 1, p. 89-122, 2003.

WILLMS, J. Douglas. *Monitoring school performance*: a guide for educator. Washington, D.C.: The Falmer, 1992.

CAMILA AKEMI KARINO
Mestre em Psicologia Social, do Trabalho e das Organizações pela Universidade de Brasília (UnB). Doutoranda do Programa de Pós-Graduação do Instituto de Psicologia da UnB, Brasília, Distrito Federal, Brasil
camilaakarino@gmail.com

LUÍS GUSTAVO DO AMARAL VINHA
Mestre em Estatística pela Universidade de São Paulo (USP). Professor Assistente do Departamento de Estatística da Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil
lgvinha@gmail.com

JACOB ARIE LAROS
Doutor PhD em Psicologia pela University of Groningen, Holanda. Professor Associado do Instituto de Psicologia da Universidade de Brasília (IP-UnB), Brasília, Distrito Federal, Brasil
jalaros@gmail.com

Recebido em: NOVEMBRO 2013
Aprovado para publicação em: NOVEMBRO 2014

CONFIABILIDADE E CONCORDÂNCIA ENTRE JUÍZES: APLICAÇÕES NA ÁREA EDUCACIONAL

DANIEL ABUD SEABRA MATOS

RESUMO

Os objetivos desta pesquisa foram: (1) investigar as estratégias de verificação da confiabilidade e concordância entre juízes, enfatizando as aplicações na área educacional; (2) realizar uma revisão da literatura nacional sobre as técnicas de confiabilidade e concordância entre juízes e suas áreas de aplicação; e (3) ilustrar a aplicação das técnicas de confiabilidade e concordância entre juízes por meio da análise das correções das redações do vestibular de uma universidade pública de Minas Gerais. Utilizamos o coeficiente de correlação intraclasse para analisar a confiabilidade e concordância entre juízes na correção das redações no período de 2005 a 2010. Identificamos pouco uso, nas pesquisas educacionais, de técnicas de concordância entre juízes. Quanto à análise da correção das redações, alguns resultados foram satisfatórios (exemplo: confiabilidade média dos juízes para as notas totais das redações) e outros insatisfatórios (exemplo: concordância baixa em alguns critérios de correção).

PALAVRAS-CHAVE TAXA DE CONFIABILIDADE • CONCURSO VESTIBULAR • JUÍZES • REDAÇÃO.

RESUMEN

Los objetivos de la presente investigación fueron los siguientes: (1) investigar las estrategias de verificación de la confiabilidad y concordancia entre jueces, enfatizando las aplicaciones en el área educativa; (2) realizar una revisión de la literatura nacional sobre las técnicas de confiabilidad y concordancia entre jueces y sus áreas de aplicación; y (3) ilustrar la aplicación de las técnicas de confiabilidad y concordancia entre jueces por medio del análisis de las correcciones de las redacciones del examen de ingreso a una universidad pública de Minas Gerais. Utilizamos el coeficiente de correlación intraclase para analizar la confiabilidad y concordancia entre jueces en la corrección de las redacciones en el periodo de 2005 a 2010. Identificamos poco uso, en las investigaciones educativas, de técnicas de concordancia entre jueces. En lo que se refiere al análisis de la corrección de las redacciones, algunos de los resultados fueron satisfactorios (ejemplo: confiabilidad media de los jueces para las notas totales de las redacciones) y otros insatisfactorios (ejemplo: baja concordancia en algunos criterios de corrección).

PALABRAS CLAVE TASA DE CONFIABILIDAD • EXAMEN PARA INGRESO A LA UNIVERSIDAD• JUECES • REDACCIÓN.

ABSTRACT

The aims of this study were to: (1) investigate the strategies for verifying reliability and agreement among evaluators, focusing on the applications in the educational area; (2) conduct a review of the national literature on the techniques of reliability and agreement among judges and their areas of application; and (3) illustrate the application of the techniques of reliability and agreement among evaluators by analyzing the corrections of the Vestibular (college entrance exam) essays from one public university in Minas Gerais. We used the intraclass correlation coefficient to analyze the reliability and agreement among evaluators in the correction of the essays from 2005 to 2010. We identified little use, in the educational research, of agreement techniques among evaluators. As for the analysis of the correction of essays, some results were satisfactory (example: mean reliability of the evaluators for total scores of the essays) and others were unsatisfactory (example: low agreement in some criteria of correction).

KEYWORDS RATE OF RELIABILITY • VESTIBULAR EXAMINATION • EVALUATORS • ESSAY.

INTRODUÇÃO

Nos últimos anos, tem acontecido um aumento do interesse na área de avaliação e suas aplicações no campo educacional. Esse contexto produziu um crescimento significativo de pesquisas sobre a avaliação educacional. Também tem se configurado atualmente um maior compromisso do governo, das universidades, das escolas e dos profissionais da educação com a realização de avaliações confiáveis e de boa qualidade.

Algumas possíveis definições de avaliação incluem tanto definições mais gerais quanto definições mais específicas do campo educacional. São alguns exemplos: um processo de delineamento, obtenção e fornecimento de informações que permitam julgar alternativas de decisão (STUFFLEBEAM, 1971) e um procedimento sistemático e compreensivo em que se utilizam estratégias diversas para avaliar a trajetória acadêmica e pessoal do estudante (QUINTANA, 2003). Ainda com relação à definição de avaliação, para Jorba e Sanmartí (2003), toda atividade de avaliação é um processo com três etapas: 1) coleta de informação, que pode ser ou não instrumentada; 2) análise dessa informação e conclusão sobre o

resultado da análise; e 3) tomada de decisões de acordo com a conclusão.

Além disso, para cumprir seus propósitos, a avaliação precisa atender a alguns requisitos. Assim, para que uma avaliação seja considerada de boa qualidade precisa ter, por exemplo, validade (*validity*) e confiabilidade (*reliability*). Validade pode ser definida como "o grau em que todas as evidências acumuladas corroboram a interpretação pretendida dos escores de um teste para os fins propostos" (AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; AMERICAN PSYCHOLOGICAL ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, 1999). Já a confiabilidade está relacionada com a consistência e precisão dos resultados do processo de mensuração (URBINA, 2007).

Portanto, dentre as várias possibilidades de abordagem da avaliação (inclusive as diferentes etapas do processo avaliativo), podemos afirmar que o principal foco desta pesquisa envolve a etapa de análise da informação e dos resultados, com ênfase no critério da confiabilidade. Mais especificamente, investigamos estratégias de verificação da confiabilidade e concordância entre juízes (*inter-rater agreement and reliability*) e suas aplicações na área educacional. O termo juiz é usado aqui como sinônimo de avaliador, ou seja, em situações nas quais mais de um juiz (avaliador) participa de um processo avaliativo, existem estratégias para verificar qual o grau de concordância desses juízes (avaliadores). Em situações como essas, é crucial verificar se existe um nível mínimo de concordância entre os juízes. A importância disso reside no fato de explicitar a confiabilidade de um processo avaliativo. Se dois ou mais juízes discordam muito em uma avaliação, isso pode indicar uma falta de confiabilidade nos resultados.

A literatura especializada aponta diversas maneiras de medir o nível de confiabilidade e concordância entre juízes, incluindo técnicas como porcentagem, correlação, coeficiente *Kappa* de Cohen, dentre outras (HANEY et al., 2004). No entanto, existe uma lacuna nas pesquisas educacionais brasileiras, pois estratégias de verificação do nível de confiabilidade e concordância entre juízes têm sido pouco estudadas e aplicadas no campo educacional. Essa, porém, parece ser uma realidade diferente da área de ciências da saúde no

Brasil, em que se verifica uma maior aplicação das técnicas de confiabilidade e concordância entre juízes (ANDRADE; SHIRAKAWA, 2006; BRUSCATO; IACOPONI, 2000; DEL-BEN et al., 2001; FRAGA-MAIA; SANTANA, 2005; PERROCA; GAIDZINSKI, 2002, 2003; POLANCZYK et al., 2003; VENTURA; BOTTINO, 2001).

Em face da relevância dessa área de trabalho, realizamos a presente pesquisa cujos objetivos foram: (1) investigar as estratégias de verificação da confiabilidade e concordância entre juízes, enfatizando as aplicações na área educacional; (2) realizar uma revisão da literatura nacional sobre as técnicas de confiabilidade e concordância entre juízes e suas áreas de aplicação (ciências humanas e ciências da saúde); e (3) ilustrar a aplicação das técnicas de confiabilidade e concordância entre juízes por meio da análise das correções das redações do vestibular de uma universidade pública do estado de Minas Gerais.

CONFIABILIDADE E CONCORDÂNCIA ENTRE JUÍZES

Pesquisadores e profissionais geralmente usam o termo confiabilidade entre juízes (*inter-rater reliability*) como uma expressão genérica para a consistência entre avaliadores. No entanto, alguns especialistas em avaliação adotam uma definição mais precisa para o termo. A confiabilidade entre juízes pode ser, assim, definida como uma medida da consistência entre avaliadores na ordenação ou posição relativa de avaliações de desempenho, independentemente do valor absoluto da classificação de cada avaliador. Já a concordância entre juízes (*inter-rater agreement*) pode ser definida como o grau em que dois ou mais avaliadores, utilizando a mesma escala de avaliação, fornecem igual classificação para uma mesma situação observável. Dessa maneira, ao contrário da confiabilidade entre juízes, a concordância entre juízes é uma medida da consistência entre o valor absoluto das classificações dos avaliadores (GRAHAM et al, 2012). Além disso, é possível que dois avaliadores tenham pouca ou nenhuma concordância e ainda assim apresentem uma confiabilidade alta (TINSLEY; WEISS, 2000). O Quadro 1 ilustra essa possibilidade:

QUADRO 1 – Diferença entre confiabilidade e concordância

	CONCORDÂNCIA BAIXA, CONFIABILIDADE ALTA		CONCORDÂNCIA ALTA, CONFIABILIDADE ALTA	
	JUIZ 1	JUIZ 2	JUIZ 3	JUIZ 4
PROFESSOR A	1	2	1	1
PROFESSOR B	2	3	2	2
PROFESSOR C	3	4	3	3
PROFESSOR D	4	5	4	4
CONCORDÂNCIA	0,0		1,0	
CONFIABILIDADE	1,0		1,0	

Fonte: Tinsley, Weiss (2000).

A concordância mede com que frequência dois ou mais avaliadores atribuem exatamente a mesma classificação. A confiabilidade mede a semelhança relativa entre dois ou mais conjuntos de classificações. Nesse sentido, o Quadro 1 exemplifica a diferença entre confiabilidade e concordância. Os juízes 1 e 2 estão de acordo sobre o desempenho relativo dos quatro professores, pois ambos atribuíram classificações que aumentam gradativamente (o professor A recebe o menor escore e o Professor D recebe o maior escore). No entanto, embora eles concordem sobre o ranqueamento relativo dos quatro professores, não concordaram nenhuma vez sobre o nível absoluto de desempenho. Consequentemente, o nível de confiabilidade entre os juízes 1 e 2 foi perfeito (1.0), mas não existiu nenhuma concordância. Já os juízes 3 e 4 concordaram tanto sobre o nível absoluto quanto sobre a ordem relativa do desempenho dos professores. Dessa forma, eles tiveram a confiabilidade (1.0) e a concordância perfeitas (1.0) entre juízes (GRAHAM et al., 2012).

Uma das possíveis explicações para a diferença entre confiabilidade e concordância é a utilização de pontos de ancoragem distintos. Por exemplo: juízes realizando classificações em uma escala de 1 a 10. Imaginemos que o juiz 1 avalia todos os sujeitos com escores altos na parte superior da escala (entre 5 e 10) e o juiz 2 avalia todos com escores baixos (entre 1 e 5 na escala). Numa situação como essa, precisamos definir se a variabilidade individual do juiz é

importante para os fins de uma dada pesquisa. Se não for, simplesmente testamos se cada juiz classificou cada observação de uma maneira semelhante (ordenação ou posição relativa de avaliações – consistência/confiabilidade). Mas se a variabilidade individual do juiz for importante, então testamos se cada juiz deu para cada observação exatamente o mesmo escore (diferença no valor absoluto – concordância). Assim, nessa situação, podemos encontrar alta consistência/confiabilidade e pouca ou nenhuma concordância entre os juízes.

Normalmente, a concordância entre juízes é mais importante para os educadores quando eles tomam decisões de alto impacto (*high-stakes decisions*), como retenção ou promoção. Isso ocorre porque, muitas vezes, precisamos tomar decisões com base num limiar de pontuação com um critério de corte. A concordância entre juízes também é importante quando informa os resultados de avaliação com o intuito de fornecer *feedback*. A confiabilidade entre juízes é mais frequentemente utilizada em pesquisas ou onde o único interesse é na consistência das decisões dos avaliadores sobre os níveis relativos de desempenho. Com base nessas definições, a concordância entre juízes pode ser considerada a medida de maior interesse para avaliações educacionais (GRAHAM et al., 2012).

Ainda com relação à nomenclatura empregada na literatura, destacamos que neste trabalho usamos preferencialmente a expressão "concordância entre juízes". No entanto, quando necessário, efetuamos a distinção entre os conceitos.

Quanto aos métodos para calcular a concordância entre juízes, a literatura indica vários, sendo que a porcentagem de concordância absoluta (*percentage of absolute agreement*) é a técnica mais simples utilizada. Ela consiste unicamente em calcular o número de vezes em que os avaliadores concordam e dividir pelo número total de avaliações (varia entre 0 e 100%). Para Stemler (2004), o valor de 75% é considerado o mínimo de concordância aceitável, já valores a partir de 90% são considerados altos. Uma desvantagem dessa técnica reside no fato de ela não levar em consideração a proporção de concordância devido ao acaso.

Nesse sentido, as tentativas de estimativas melhores da concordância entre juízes começaram com o desenvolvimento

do coeficiente *kappa* de Cohen – K (*Cohen's kappa coefficient*) (COHEN, 1960). O coeficiente *kappa* é um procedimento estatístico que leva em consideração no seu cálculo a probabilidade de concordância ao acaso (CROCKER; ALGINA, 2009). Assim, esse coeficiente pode ser definido como a proporção de concordância entre os juízes após ser retirada a proporção de concordância devido ao acaso (FONSECA et al., 2007). O *kappa* varia entre 0 e 1, podendo ser interpretado da seguinte forma: $K<0,4$ é pobre; $0,4 \leq K < 0,75$ é satisfatório a bom; $K \geq 0,75$ é excelente (FLEISS, 1981). No entanto, esse critério de corte não é totalmente consensual na literatura. Um dos primeiros critérios foi o proposto por Landis e Koch (1977): $K< 0$: sem concordância; $0 \leq K < 0,21$: presença de ligeira concordância; $0,21 \leq K < 0,41$: concordância fraca; $0,41 \leq K < 0,61$: concordância moderada; $0,61 \leq K < 0,81$: concordância substancial; $0,81 \leq K \leq 1,00$: concordância quase perfeita. Altman (1991) considera a concordância alta a partir de 0,80. Destacamos, ainda, que a literatura descreve o *kappa* como o método mais utilizado quando as variáveis são nominais (FONSECA et al., 2007).

Para variáveis ordinais, existe o chamado *kappa* ponderado (kw), que permite atribuir pesos diferentes às concordâncias e aos desacordos. Assim, esse coeficiente, desenvolvido por Cohen (1968), é um índice de concordância preferível quando classificamos um conjunto de dados em categorias ordenadas, pois o k distingue apenas entre acordo e desacordo em categorias nominais (SCHUSTER, 2004). Em 1981, Fleiss elaborou uma extensão do *kappa* e a denominou *Fleiss' generalized kappa*, para incluir casos em que existem três ou mais juízes (KING, 2004). Dessa maneira, a limitação do *kappa* original (medir a concordância apenas entre dois juízes) foi superada pela inclusão de múltiplos juízes.

Nos últimos anos, existe uma preocupação crescente com relação às limitações do *kappa*. As principais críticas são: o k é afetado pela distribuição heterogênea de categorias (problema de prevalência) e pela extensão na qual os juízes discordam (problema de viés) (BLOOD; SPRATT, 2007). Como uma forma de superar essas limitações, Gwet (2001) desenvolveu duas novas estatísticas: *first-order agreement coefficient* (AC1)

e *second-order agreement coefficient* (AC2) (GWET, 2001). O AC1 é utilizado com dois ou mais juízes e uma escala de classificação com duas ou mais categorias. Já o AC2 também é utilizado com dois ou mais juízes, mas com uma escala de classificação ordenada contendo duas ou mais categorias (BLOOD; SPRATT, 2007). Assim como o *kappa*, ambos os coeficientes AC1 e AC2 variam entre 0 e 1 e possuem a mesma interpretação: quanto mais próximo de 1, melhor (menor a probabilidade de a concordância acontecer devido ao acaso). Entretanto, as pesquisas com as estatísticas AC1 e AC2 ainda são poucas e as conclusões baseadas nesses índices ainda precisam ser cautelosas (BLOOD; SPRATT, 2007). Dessa forma, o coeficiente *kappa* continua sendo largamente utilizado para aferir a concordância entre juízes.

Para dados contínuos, existe um tipo especial de coeficiente de correlação denominado intraclasse – CCI (*intra-class correlation coefficient*), que é a medida de concordância mais utilizada para variáveis contínuas (LU, 2007). A correlação de Pearson mede a intensidade da associação interclasse (entre variáveis de classes diferentes, entre construtos diferentes). Já o coeficiente de correlação intraclasse mede essa intensidade dentro de uma mesma classe (diferentes medidas de um mesmo construto), que podem ser medidas repetidas de um mesmo participante ou medidas de várias pessoas dentro de um mesmo grupo (exemplo: estudantes em uma sala de aula). Assim, o CCI é aplicado em dados estruturados em grupos, sendo obtido dividindo o valor da variação entre os indivíduos pela variação total. O coeficiente de correlação intraclasse é uma medida de concordância corrigida pela concordância esperada ao acaso (BLAND; ALTMAN, 1990). Além disso, existem tipos diferentes de CCI e ele pode ser estimado de diversas formas. Alguns tipos de CCI são bastante conhecidos na literatura, o Coeficiente Alfa de Cronbach, por exemplo, é muito utilizado na área de psicometria para avaliar a consistência interna (confiabilidade), a similaridade entre os itens de um teste.

Uma das vantagens do CCI é que ele representa a concordância entre dois ou mais juízes ou entre várias medidas feitas pelo mesmo juiz. O CCI é equivalente à estatística

kappa para variáveis contínuas: também varia entre 0 e 1, podendo ser interpretado da seguinte forma: CCI<0,4 é pobre; 0,4≤CCI<0,75 é satisfatório a bom; CCI ≥0,75 é excelente (FLEISS, 1981). Esses critérios também podem variar na literatura, inclusive dependendo da técnica empregada. Para Hair *et al.* (2005), por exemplo, os valores de Alfa de Cronbach de 0,60 a 0,70 são considerados como limite inferior da aceitabilidade. Entretanto, segundo Graham *et al.* (2012), existe pouco consenso sobre um valor suficiente para o CCI. Enquanto 0,70 pode ser suficiente para uma medida utilizada para fins de pesquisa, alguns pesquisadores defendem um valor de 0,8 a 0,9 como o mínimo para a tomada de decisões importantes sobre pessoas (HAYS; REVICKI, 2005).

Destacamos, ainda, que a literatura aponta outros métodos para verificar o nível de concordância entre juízes. São exemplos disso: o coeficiente *Bennett's S,* a Teoria de Resposta ao Item (TRI) (EMBRETSON; REISE, 2000), dentre outros. Primi *et al.* (2007), por exemplo, analisaram o nível de confiabilidade de juízes com relação ao Teste de Criação de Metáforas (TCM). Esse teste é composto por nove itens, aos quais os participantes deram um total de 513 respostas. Cada resposta foi avaliada de forma independente pelos juízes, em uma escala de 0 a 3, correspondente ao nível de elaboração da metáfora. A concordância foi aferida por meio do modelo de Rasch, assumindo cada ideia como sendo um caso e cada juiz como sendo um item de um teste hipotético. Os autores utilizaram um procedimento indicado por Linacre (1998), denominado rede de juízes ancorados (*judge-linking network*). Na matriz de dados, foi aplicado o modelo de Rasch de créditos parciais (PRIMI et al., 2007).

MÉTODO

Inicialmente, realizamos uma busca na literatura brasileira sobre trabalhos que abordam as técnicas de verificação do nível de concordância entre juízes. Bancos de dados do *Scientific Electronic Library Online* (SciELO) e da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) foram pesquisados na internet. Combinamos os seguintes descritores usando

tanto o singular quanto o plural: nível, concordância, confiabilidade, juízes, avaliadores. Também utilizamos descritores em inglês (*inter-rater, intercoder, agreement, realiability*), pois alguns periódicos nacionais publicam artigos em inglês. Em alguns casos, foram encontrados somente os resumos dos estudos, que não foram incluídos nessa revisão da literatura porque não atendiam ao critério de fornecer dados suficientes para a análise das pesquisas. Nossa revisão sobre as técnicas de concordância entre juízes é analisada pela área de aplicação (ciências da saúde ou ciências humanas) e pela identificação dos métodos mais utilizados.

Após o levantamento bibliográfico, ilustramos a aplicação das técnicas de confiabilidade e concordância entre juízes por meio da análise das correções das redações do vestibular de uma universidade pública do estado de Minas Gerais. Cada redação foi corrigida por dois avaliadores. No caso de discrepância nas notas, um terceiro avaliador participava da correção. Analisamos todas as redações da segunda etapa do vestibular entre os anos de 2005 e 2010. No período compreendido entre 2005 e 2009, o valor máximo da redação era de 20 pontos. Em 2010, o valor total passou a ser de 30 pontos e a diferença (discrepância) de nota permitida entre as correções era de 3 pontos. Não temos informação sobre o critério de discrepância para os anos anteriores. Outra limitação dos nossos dados foi que, nos casos de discrepância, não tivemos acesso à nota do terceiro corretor.

A literatura descreve dois métodos geralmente utilizados para a correção de redações em larga escala: o holístico e o analítico. A escolha adequada de um ou outro método depende de muitos fatores como, por exemplo, dos objetivos estabelecidos. Se o objetivo principal é a seleção dos estudantes, o método holístico pode ser suficiente. Entre os testes em larga escala que empregam esse método, podemos citar o TOEFL (*Test of English as a Foreign Language*). O método holístico consiste basicamente em atribuir um escore único com todos os critérios da avaliação considerados conjuntamente (exemplo: tema, coesão e ortografia). Já no método analítico, as redações não recebem apenas uma nota geral. Os avaliadores atribuem notas em diversos critérios, que combinados

geram um escore final. Um exemplo de avaliação em larga escala que emprega esse método é o Exame Nacional do Ensino Médio (Enem). Nesse exame, os examinadores avaliam a redação de acordo com cinco critérios (competências). Cada uma das competências recebe uma nota entre 0 e 200 pontos, podendo atingir um escore máximo de 1.000 pontos. No entanto, destacamos que o método analítico não tem sido muito utilizado para a correção de redações em larga escala por diversos fatores, como o alto custo (BACHA, 2001). Entretanto, algumas pesquisas indicam que, caso o objetivo principal seja mais educacional (*feedback* para o aluno dos seus pontos fortes e fracos), o método analítico pode ser considerado mais apropriado (HAMP-LYONS, 1991), o que pode ser visto nas redações da segunda etapa da universidade pública que analisamos, já que se enquadram no método de correção analítico. O Quadro 2 mostra com detalhes os critérios empregados na correção.

QUADRO 2 – Critérios de correção das redações de 2010

FATORES AVALIADOS	PONTUAÇÃO MÁXIMA	CORREÇÃO DO EXAMINADOR
1. Adequação à proposta	6	
Tema	2	
Gênero / tipo textual	2	
Condições de circulação do texto	2	
2. Estratégias de textualização	13	
Coesão / coerência	4	
Progressão / paragrafação	3	
Informatividade	3	
Argumentação	3	
3. Aspectos Morfossintáticos	8	
Construção do período/ pontuação	4	
Concordância / regência / colocação	4	
4. Ortografia	3	
Total	30	

Fonte: Elaboração do autor.

Como indicado no Quadro 2, foi utilizado na correção das redações de 2010 um método de correção analítico com quatro critérios: adequação à proposta (máximo de 6 pontos), estratégias de textualização (máximo de 13 pontos), aspectos morfossintáticos (máximo de 8 pontos) e ortografia (máximo de 3 pontos). A nota final do candidato podia atingir 30 pontos, como resultado da soma de pontos obtidos nos quatro critérios de avaliação da redação.

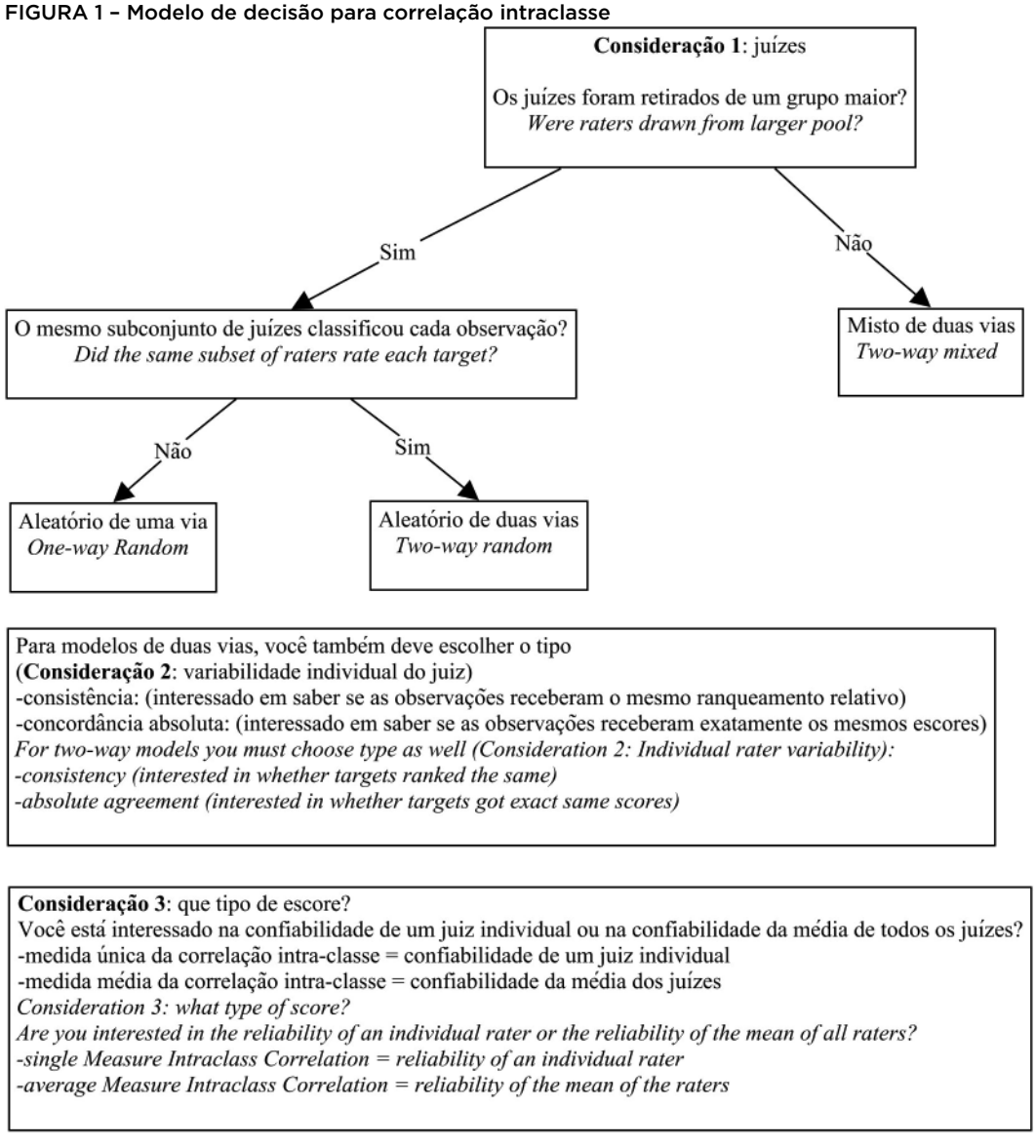
Destacamos que para os anos entre 2005 e 2009 não tivemos acesso aos critérios de correção (só obtivemos as notas totais de cada corretor). Além disso, não foi possível identificar os corretores pelos nomes. Dessa forma, a falta de informações para o referido período limitou a possibilidade de trabalho com os dados.

Como a nota das redações é uma variável contínua, utilizamos, nesta pesquisa, o coeficiente de correlação intraclasse – CCI, e, para realizar o cálculo das técnicas do nível de concordância entre juízes, adotamos o *software SPSS 20*. Para os anos de 2005 a 2010, produzimos uma análise geral da confiabilidade dos juízes considerando o processo seletivo como um todo. Além disso, no ano de 2010, fizemos também análises mais detalhadas do processo de correção para explicitar a gama variada de aplicações que as técnicas de concordância entre juízes possuem.

COEFICIENTE DE CORRELAÇÃO INTRACLASSE

Já afirmamos anteriormente que existem tipos diferentes de coeficiente de correlação intraclasse (CCI) e que ele pode ser estimado de diversas formas. Nesse sentido, indicamos uma série de questões que precisam ser respondidas para a escolha correta da forma de estimação do CCI (FIgura 1).

FIGURA 1 – Modelo de decisão para correlação intraclasse



Fonte: Romberg (2009, tradução nossa).

Como apontado na Figura 1, diversas questões devem ser respondidas para a utilização correta do CCI. Primeiramente, com relação aos juízes (consideração 1). Por exemplo: o mesmo subconjunto de juízes classificou cada observação? Em segundo lugar, para modelos de duas vias, a variabilidade individual do juiz está em jogo (consideração 2); em outras palavras, devemos estabelecer se nosso interesse está na

confiabilidade (consistência) ou na concordância (concordância absoluta) dos juízes. Por fim, determinamos o tipo de escore de nosso interesse: confiabilidade de um juiz individual ou confiabilidade da média de todos os juízes (consideração 3).

Destacamos que nossos resultados serão apresentados seguindo esse modelo de decisão.

RESULTADOS

REVISÃO DA LITERATURA

No total, encontramos na literatura brasileira 62 estudos que abordam as técnicas de verificação do nível de concordância entre juízes (25 dissertações/teses e 37 artigos). Nossa hipótese de que existia uma lacuna nas pesquisas educacionais brasileiras foi confirmada, pois 52 estudos foram realizados na área das ciências da saúde e apenas 10 nas ciências humanas (6 eram trabalhos relacionados à psicologia e apenas 4 se referiam a aplicações na educação). Santos *et al.* (2010), por exemplo, investigaram a confiabilidade dos julgamentos dos examinadores de prova escrita do processo seletivo do mestrado em Educação de uma universidade privada do Centro-Oeste. No total, foram 10 processos de seleção, entre os anos de 1994 e 2006. As atribuições de notas dos examinadores (juízes) foram analisadas ano a ano por meio do coeficiente de correlação de Pearson. Os resultados foram muito irregulares, com índices de correlação variando entre baixos (r = 0,15) e elevados (r = 0,89). Os autores finalizaram o estudo sugerindo procedimentos adicionais para melhorar a confiabilidade na avaliação das provas escritas do processo seletivo do mestrado.

Ainda com relação a nossa revisão da literatura, as estratégias de verificação do nível de concordância entre juízes utilizadas com maior frequência nos estudos foram: coeficiente *kappa* (23), coeficiente alfa de Cronbach (18), correlação de Pearson (12), coeficiente de correlação intraclasse (11), correlação de Spearman (7) e *kappa* ponderado (6). Outras técnicas foram empregadas com menor frequência, sendo alguns exemplos: *Kappa* múltiplo – *Fleiss' generalized kappa* (2), porcentagem de concordância absoluta (2), Teoria

de Resposta ao Item – modelo de Rasch (1) e Coeficiente Alfa de Krippendorff (1). Destacamos aqui dois aspectos: alguns estudos utilizam mais de uma técnica; e, apesar de a literatura apontar como método mais adequado para estimar a concordância entre juízes o coeficiente de correlação intraclasse, diversos trabalhos empregam outros tipos de correlação. Ainda assim, se somarmos todas as pesquisas que utilizaram correlação, essa passa a ser a técnica mais frequente.

CORREÇÕES DAS REDAÇÕES

Para os resultados das correções das redações, organizamos nossos dados a partir de diferentes perguntas de pesquisa. Destacamos que cada uma dessas perguntas exige uma aplicação diferente das técnicas de concordância entre juízes. Além disso, nas análises de todas as perguntas, excluímos as redações que tiveram nota zero, pois acreditamos que isso aumentaria artificialmente a concordância entre os juízes. Os critérios para o aluno receber nota zero eram muito claros (exemplo: entregar a redação em branco). De qualquer forma, o número de redações com nota zero no banco de dados foi pequeno.

PERGUNTA 1

Consideradas como um todo, qual foi a qualidade das correções das redações para o período entre 2005 e 2010? Em outras palavras: para fins de processo seletivo (vestibular), existiu uma confiabilidade mínima entre os juízes?

Consideração 1: juízes

Os juízes foram retirados de um grupo maior? Sim, dentre um grupo grande de corretores, somente dois eram designados para corrigir cada redação.

O mesmo subconjunto de juízes classificou cada observação? Não. Duplas diferentes de juízes corrigiram cada redação.

Isso implica um modelo aleatório de uma via. Portanto, não podemos realizar a escolha da consideração 2. Nesse caso, temos o resultado apenas da consistência (confiabilidade).

Consideração 3: Que tipo de escore?

Estamos interessados na confiabilidade da média de todos os juízes. Portanto, escolhemos a medida média da correlação intraclasse.

A Tabela 1 exibe, para todos os anos, o número de redações corrigidas, a menor e a maior nota, a média, o desvio padrão e a correlação intraclasse com o respectivo intervalo de confiança.

TABELA 1 – Análise das notas totais das redações no período entre 2005 e 2010

ANO	N	MIN.	MAX.	MÉDIA	DESVIO PADRÃO	CORRELAÇÃO INTRACLASSE (MÉDIA)	INTERVALO DE CONFIANÇA	
							LIMITE INFERIOR	LIMITE SUPERIOR
2005	4.266	1,5	20,0	11,8	3,1	0,86	0,858	0,874
2006	4.335	3,3	20,0	12,2	2,9	0,89	0,883	0,896
2007	2.324	1,0	18,5	11,6	2,5	0,87	0,864	0,885
2008	6.857	2,0	20,0	12,2	2,6	0,92	0,924	0,930
2009	10.223	1,4	20,0	11,6	2,7	0,91	0,911	0,918
2010	10.108	2,7	29,7	16,2	4,3	0,95	0,953	0,957

Fonte: Elaboração do autor.

Nota: Modelo aleatório de uma via. Intervalo de confiança 95 %.

A partir da Tabela 1, percebemos que um número grande de redações era corrigido todos os anos, sendo 2009 o ano em que mais redações foram corrigidas (10.223). A nota média das redações e a variabilidade das notas possuem valores próximos entre os anos de 2005 e 2009. Já no ano de 2010, a nota média das redações torna-se maior pelo fato de o valor total da redação ter sido alterado para 30 pontos, e, da mesma forma, a variabilidade dos dados também se torna maior. Quanto aos coeficientes de correlação intraclasse, eles variaram entre 0,86 e 0,95. Como essa análise se enquadra em uma decisão de alto impacto (aprovar pessoas em um vestibular de uma universidade pública), acreditamos que um critério de corte rigoroso precisa ser adotado. Nesse sentido, tendo como referência critérios para a tomada de decisões importantes sobre pessoas (valor mínimo de 0,8 a 0,9) (HAYS;

REVICKI, 2005), podemos afirmar que a confiabilidade da média de todos os juízes quanto à nota total das redações foi satisfatória. Também chama atenção o fato de o menor valor da correlação intraclasse ser de 2005 e o maior valor, de 2010. Isso provavelmente está relacionado à manutenção de membros da equipe de examinadores que, com o decorrer do tempo, aperfeiçoaram suas habilidades de correção.

Os resultados da pergunta 1 são importantes, mas apresentam algumas limitações, pois temos apenas dados sobre a confiabilidade (consistência) média de todos os juízes para a nota total da redação. Lembrando-se de que na nossa definição apresentada anteriormente, a confiabilidade entre juízes é uma medida da consistência entre avaliadores na ordenação ou posição relativa de avaliações de desempenho, independentemente do valor absoluto da classificação de cada avaliador. Assim, quando temos diferentes subconjuntos de juízes classificando cada participante, não podemos particionar a variância devido a juízes individuais. Portanto, com esses dados, não conseguimos identificar, por exemplo, membros da equipe que estejam com um padrão de confiabilidade abaixo do mínimo esperado. Isso fica diluído nesse tipo de análise geral. Além disso, também não podemos obter resultados sobre a concordância entre juízes (medida da consistência entre o valor absoluto das classificações dos avaliadores).

PERGUNTA 2

Analisando uma mesma dupla de juízes classificando as mesmas redações, qual a confiabilidade e a concordância da média dos dois examinadores?

Consideração 1: juízes

Os juízes foram retirados de um grupo maior? Sim, dentre um grupo grande de corretores, somente dois foram selecionados.

O mesmo subconjunto de juízes classificou cada observação? Sim. A mesma dupla de juízes corrigiu cada redação.

Isso implica um modelo aleatório de duas vias.

Consideração 2: variabilidade individual do juiz

Estamos interessados tanto na consistência (confiabilidade) quanto na concordância absoluta.

Consideração 3: Que tipo de escore?

Estamos interessados na confiabilidade da média dos juízes. Portanto, escolhemos a medida média da correlação intraclasse.

A Tabela 2 apresenta a confiabilidade e concordância média de juízes em 2010, tanto para a nota total da redação quanto para os quatro critérios de correção separadamente.

TABELA 2 – Confiabilidade e concordância média de juízes em 2010

JUÍZES	COEFICIENTES DE CORRELAÇÃO		CRITÉRIOS DE CORREÇÃO DAS REDAÇÕES				
			NOTA TOTAL	ADEQUAÇÃO À PROPOSTA (1)	ESTRATÉGIAS DE TEXTUALIZAÇÃO (2)	ASPECTOS MORFOSSINTÁTICOS (3)	ORTOGRAFIA (4)
Dalila e Eloisa	Alfa de Cronbach		0,93	0,81	0,75	0,68	0,77
	CCI	Consistência	0,93	0,81	0,75	0,68	0,77
		Concordância	0,86	0,81	0,70	0,60	0,77
Elisangela e Karine	Alfa de Cronbach		0,88	0,80	0,78	0,86	0,90
	CCI	Consistência	0,88	0,80	0,78	0,86	0,90
		Concordância	0,88	0,80	0,78	0,86	0,90

Fonte: Elaboração do autor.

Nota: Modelo aleatório de duas vias. Dalila e Eloisa, N= 374; Elisangela e Karine, N= 107 (nomes fictícios).

CCI: Coeficiente de correlação intraclasse.

Como já destacamos, a concordância mede com que frequência dois ou mais avaliadores atribuem exatamente a mesma classificação. A confiabilidade mede a semelhança relativa entre dois ou mais conjuntos de classificações. Nesse sentido, a Tabela 2 ilustra a diferença entre confiabilidade e concordância. No caso dos coeficientes de correlação intraclasse dos juízes Dalila e Eloisa, a concordância foi menor do que a confiabilidade para a nota total, o critério 2 e o critério 3. Como geralmente a concordância entre juízes é mais importante na educação quando tomamos decisões de alto impacto, podemos realizar uma série de considerações. Quanto à nota total, tomando novamente como critério de corte um valor

mínimo de 0,8 ou 0,9 (HAYS; REVICKI, 2005), a concordância pode ser considerada satisfatória. Além disso, como o método de correção da redação é analítico, é importante analisar também os quatro critérios separadamente. Assim, para a confiabilidade, os coeficientes variaram entre 0,68 e 0,81 (somente o critério 1 está acima do ponto de corte de 0,8). Já para a concordância (resultado mais importante), os coeficientes variaram entre 0,60 e 0,81 (os valores diminuem ainda mais nos critérios 2 e 3).

A dupla Elisangela e Karine, quando comparada à dupla Dalila e Eloisa, apresenta resultados mais confiáveis. Isso porque os coeficientes de confiabilidade e concordância são altos e possuem os mesmos valores. Quanto à nota total, apesar de indicar menor confiabilidade (0,88) do que Dalila e Eloisa (0,93), a dupla apresentou maior concordância (0,88). Isso se reflete também nos critérios da correção analítica, já que os coeficientes variaram entre 0,78 e 0,90, e foram, no geral, mais altos do que o valores da dupla Dalila e Eloisa, sendo que somente um dos critérios está abaixo do critério de corte de 0,8.

As análises dos quatro critérios separadamente são importantes por um motivo: geralmente, no método analítico, é recomendado um terceiro corretor tanto no caso de uma discrepância na nota total quanto de discrepância nos critérios adotados. O Enem, por exemplo, considera discordância entre avaliadores 200 pontos de discrepância na nota total ou 80 pontos em cada uma das cinco competências. Assim, a partir da Tabela 2, percebemos que os critérios 2 e 3 provavelmente foram problemáticos no caso da dupla Dalila e Eloisa. No entanto, como a universidade só trabalhava com a tolerância de até 3 pontos de diferença na nota total, esses fatores não foram considerados. Nesse caso, poderia ser solicitado um terceiro avaliador para esses critérios. Interessante destacar aqui que dados como esses podem ser utilizados, por exemplo, para identificar membros da equipe que estejam com um padrão de concordância abaixo do mínimo esperado. Isso pode resultar em um melhor treinamento da equipe.

Por fim, como já apontamos previamente, existem tipos diferentes de CCI e ele pode ser estimado de diversas formas.

O Coeficiente Alfa de Cronbach é muito citado na literatura. Por ser uma medida de consistência interna (confiabilidade), seu valor sempre é coincidente com o valor do coeficiente de correlação intraclasse de consistência/confiabilidade (Tabela 2).

PERGUNTA 3

Analisando uma mesma dupla de juízes classificando as mesmas redações, qual a confiabilidade de qualquer examinador individual? Eles estão utilizando pontos de ancoragem diferentes em suas classificações?

Consideração 1: juízes

Os juízes foram retirados de um grupo maior? Sim, dentre um grupo grande de corretores, somente dois foram selecionados.

O mesmo subconjunto de juízes classificou cada observação? Sim. A mesma dupla de juízes corrigiu cada redação.

Isso implica um modelo aleatório de duas vias.

Consideração 2: variabilidade individual do juiz

Estamos interessados apenas na concordância absoluta, pois queremos verificar se os juízes estão usando pontos de ancoragem diferentes em suas classificações.

Consideração 3: Que tipo de escore?

Estamos interessados na concordância de um juiz individual. Portanto, escolhemos a medida única (*single measures*) da correlação intraclasse.

A Tabela 3 mostra a concordância individual e média de juízes em 2010, tanto para a nota total da redação quanto para os quatro critérios de correção separadamente.

Em primeiro lugar, apesar de a pergunta 3 necessitar apenas da medida única da correlação intraclasse , também incluímos na Tabela 3 a média das correlações por um motivo: ilustrar a importância de utilizar múltiplos juízes em avaliações de alto impacto. Dessa forma, o CCI individual e o CCI médio indicam a segurança que temos no processo avaliativo se estivermos baseados em apenas um juiz ou na

média dos juízes. Obviamente, para fins de seleção, utilizaremos sempre o CCI médio. No entanto, o CCI individual nos fornece outro tipo de informações igualmente importantes.

Dessa maneira, a análise dos resultados da dupla Felipe e Wallace (nomes fictícios) indica algumas questões interessantes. Primeiramente, apesar de a concordância média dos dois juízes estar em patamar aceitável para avaliações de alto impacto na nota total e no critério 1, o CCI individual indica que os juízes provavelmente estão usando pontos de ancoragem diferentes em alguns critérios de correção, como no critério 3, que parece ser o mais problemático, por apresentar uma concordância muito baixa. Por isso, tomados em conjunto, esses resultados mostram que uma concordância média satisfatória na nota total pode esconder alguns problemas, nesse caso, referentes à falta de concordância em critérios do método de correção analítico (mesma crítica apresentada na pergunta 2) e à utilização de pontos de ancoragem diferentes pelos juízes em alguns critérios.

Tabela 3 - Concordância individual e média de juízes em 2010

CRITÉRIOS DE CORREÇÃO DAS REDAÇÕES	INDIVIDUAL			MÉDIO		
	CCI	INTERVALO DE CONFIANÇA		CCI	INTERVALO DE CONFIANÇA	
		LIMITE INFERIOR	LIMITE SUPERIOR		LIMITE INFERIOR	LIMITE SUPERIOR
Nota total	0,76	0,550	0,864	0,86	0,710	0,927
Adequação à proposta (1)	0,70	0,635	0,758	0,82	0,777	0,863
Estratégias de textualização (2)	0,64	0,524	0,739	0,78	0,688	0,850
Aspectos morfossintáticos (3)	0,10	-0,017	0,216	0,18	-0,034	0,355
Ortografia (4)	0,63	0,554	0,699	0,77	0,713	0,823

Fonte: Elaboração do autor.

Nota: Modelo aleatório de duas vias. Intervalo de confiança 95 %. Felipe e Wallace. N= 265 (nomes fictícios).

CCI: Coeficiente de Correlação Intraclasse.

Por fim, a Tabela 3 também evidencia falta de precisão dos resultados, pois alguns intervalos de confiança são muito amplos. Esse erro de medida provavelmente está ligado à amostra pequena de redações analisadas (N= 265). Esse fato sugere-nos que, para avaliar a qualidade da correção da dupla de juízes, é

interessante ter o maior número possível de observações.

PERGUNTA 4

O critério de 3 pontos de discrepância na nota total da redação é suficiente?

Para responder essa questão, selecionamos aleatoriamente no banco de dados casos em que os juízes discordaram em mais de 3 pontos (N=276) na nota total das redações e casos nos quais a discordância dos juízes estava num intervalo entre 2 e 3 pontos (N=423). Na primeira situação (discordância em mais de 3 pontos), a confiabilidade média dos juízes foi de apenas 0,19. Na segunda situação, a confiabilidade média foi de 0,88. Lembramos aqui, mais uma vez, que, como temos duplas diferentes de juízes corrigindo cada redação, apenas a confiabilidade/consistência média é analisada (modelo aleatório de uma via).

Portanto, o critério de 3 pontos de discrepância na nota total da redação parece ser sensível o suficiente para indicar baixa confiabilidade entre os juízes e a necessidade de um terceiro corretor. No entanto, algumas limitações que apontamos em outras perguntas permanecem aqui, entre elas a impossibilidade de obter resultados sobre a concordância entre juízes e a falta de indicadores por parte da universidade para analisar a discrepância nos quatro critérios do método de correção analítica.

Tomadas em conjunto, as quatro perguntas de pesquisa discutidas anteriormente ilustram a complexidade da análise de um processo avaliativo. Dizer se os resultados de uma avaliação são confiáveis ou não, envolve muitas variáveis.

CONSIDERAÇÕES FINAIS

Encontramos na literatura brasileira 62 pesquisas que utilizaram estratégias de verificação da concordância entre juízes. Nossa hipótese de que existia uma lacuna nas pesquisas educacionais brasileiras foi confirmada, pois a maioria dos estudos foi realizada na área das ciências da saúde. Assim, recomendamos tanto uma maior utilização das técnicas de

concordância de juízes quanto à realização de mais pesquisas na área educacional.

Quanto à correção das redações, para os anos de 2005 a 2010, realizamos uma análise geral da confiabilidade dos juízes considerando o processo seletivo como um todo. No ano de 2010, fizemos também análises mais detalhadas do processo de correção para explicitar a gama variada de aplicações que as técnicas de concordância entre juízes possuem. A partir disso, percebemos que alguns resultados foram satisfatórios (exemplo: confiabilidade média dos juízes para as notas totais das redações) e outros insatisfatórios (exemplo: concordância baixa em alguns critérios de correção). Isso ilustra a complexidade da análise de um processo avaliativo, de modo que afirmar se os resultados de uma avaliação são confiáveis ou não envolve muitos fatores.

Com relação à concordância entre juízes, indicamos um ponto importante para pensar o critério de corte de todas as técnicas apresentadas anteriormente: de um modo geral, os pesquisadores afirmam que quanto maior as consequências resultantes da avaliação, maior a necessidade de uma concordância entre juízes alta (LEBRETON; SENTER, 2008). Portanto, não existe uma regra geral única, e devemos estar especialmente atentos para situações que envolvem decisões de alto impacto. Nesse sentido, as perguntas de pesquisa que usamos exemplificam essa questão.

Concluímos apontando uma questão central para os educadores hoje: qual o nível de confiabilidade e concordância dos diversos processos avaliativos que acontecem na área educacional? Se não conseguirmos responder a essa questão, não saberemos se as diversas avaliações estão atingindo o resultado esperado. Como já discutimos, a eficácia dos processos avaliativos está condicionada, dentre outros fatores, pelo nível de concordância entre juízes.

REFERÊNCIAS

ALTMAN, D. *Practical statistics for medical research. Boca Raton*, FL: CRC, 1991.

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; AMERICAN PSYCHOLOGICAL ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT

IN EDUCATION. *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association, 1999.

ANDRADE, M.; SHIRAKAWA, I. Versão brasileira do Defense Style Questionnaire (DSQ) de Michael Bond: problemas e soluções. *Revista de Psiquiatria do Rio Grande do Sul*, Porto Alegre, v. 28, n. 2, p. 144-160, 2006.

BACHA, N. Writing evaluation: What can analytic versus holistic essay scoring tell us? *System*, v. 29, p. 371-383, 2001.

BLAND, J. M.; ALTMAN, D. G. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput. Biol. Med.*, v. 20, n. 5, p. 337-340, 1990.

BLOOD, E.; SPRATT, K. F. *Disagreement on Agreement: Two Alternative Agreement Coefficients.* SAS Global Forum, 2007.

BRUSCATO, W. L.; IACOPONI, E. Validade e confiabilidade da versão brasileira de um inventário de avaliação de relações objetais. *Rev. Bras. Psiquiatr*., São Paulo, v. 22, n. 4, p. 172-177, 2000.

COHEN, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, v. 20, p. 37-46, 1960.

_____. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, v. 70, p. 213-220, 1968.

CROCKER, L.; ALGINA, J. *Introduction to Classical and Modern Test Theory.* Belmont, CA: Wadsworth Group, 2009.

DEL-BEN, C. M.; VILELA, J. A. A.; CRIPPA, J. A. S.; HALLAK, J. E. C.; LABATE, C. M.; ZUARDI, A. W. Confiabilidade da Entrevista Clínica Estruturada para o DSM-IV – Versão Clínica traduzida para o português. *Rev. Bras. Psiquiatr.,* São Paulo, v. 23, n. 3, p. 156-159, 2001.

EMBRETSON, S. E.; REISE, S. P. *Item response theory for psychologists*. New York: Routledge, 2000.

FLEISS, J. *Statistical methods for rates and proportions.* New York: John Wiley & Sons, 1981.

FONSECA, R.; SILVA, P.; SILVA, R. Acordo inter-juízes: O caso do coeficiente kappa. *Laboratório de Psicologia*, Lisboa, v. 5, n.1, p. 81-90, 2007.

FRAGA-MAIA, H.; SANTANA, V. S. Concordância de informações de adolescentes e suas mães em inquérito de saúde. *Revista de Saúde Pública*, São Paulo, v. 39, n. 3, p. 430-437, 2005.

GRAHAM, M.; MILANOWSKI, A.; MILLER, J. *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*. Research Report, 2012.

GWET, K. *Handbook of inter-rater reliability:* How to estimate the level of agreement between two or multiple raters. Gaithersburg: Stataxis, 2001.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Análise multivariada de dados*. 5. ed. Porto Alegre: Bookman, 2005. 593 p.

HAMP-LYONS, L. Scoring procedures for ESL contexts. In: HAMP-LYONS, L. (Ed.). *Assessing second language writing in academic contexts.* Norwood, NJ: Ablex, 1991. p. 241–276.

HANEY, W.; RUSSELL, M.; BEBELL, D. Drawing on education: using drawings to document schooling and support changes. *Harvard Educational Review*, v. 74, n. 3, 241-271, 2004.

HAYS, R. D.; REVIKI, D. A. Reliability and validity (including responsiveness). In: FAYERS, P. M.; HAYS, R. D. (Ed.). *Assessing quality of life in clinical trials:* Methods and practice. NY: Oxford University Press, 2005.

JORBA, J.; SANMARTÍ, N. A função pedagógica da avaliação. In: BALLESTER, M. (Org.). *Avaliação como apoio à aprendizagem.* Porto Alegre: Artmed, 2003. cap 2, p.23-45.

KING, J. E. Software Solutions for Obtaining a Kappa-Type Statistic for Use with Multiple Raters. In: ANNUAL MEETING OF THE SOUTHWEST EDUCATIONAL RESEARCH ASSOCIATION, 2004, Dallas, EUA. *Anais*… Dallas: 2004.

LANDIS, J. R.; KOCH, G. G. A one way components of variance model for categorical data. *Biometrics*, v. 33, p. 671–679, 1977.

LEBRETON, J. M.; SENTER, J. L. Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, v. 11, n. 4, p. 815-852, 2008.

LINACRE, J. M. Rating, judges and fairness. *Rasch Measurement Transactions*, v. 12, n. 2, p. 630-1, 1998.

LU, L.; SHARA, N. *Reliability analysis:* calculate and compare intra-class correlation coefficients (ICC) in SAS. NESUG, 2007.

PERROCA, M. G.; GAIDZINSKI, R. R. Instrumento de classificação de pacientes de perroca: teste de confiabilidade pela concordância entre avaliadores – correlação. *Rev. Esc. Enferm.*, São Paulo, v. 36, n. 3, p. 245-252, 2002.

_____. Avaliando a confiabilidade interavaliadores de um instrumento para classificação de pacientes – coeficiente Kappa. *Rev. Esc. Enferm.*, São Paulo, v. 37, n. 1, p. 72-80, 2003.

POLANCZYK, G. V.; EIZIRIK, M.; ARANOVICH, V.; DENARDIN, D.; SILVA, T. L.; CONCEIÇÃO, T. V.; PIANCA, T. G.; ROHDE, L. A. Interrater agreement for the schedule for affective disorders and schizophrenia epidemiological version for school-age children (K-SADS-E). *Rev. Bras. Psiquiatr.,* São Paulo, v. 25, n. 2, p. 87-90, 2003.

PRIMI, R.; MIGUEL, F. K.; COUTO, G.; MUNIZ, M. Precisão de avaliadores na avaliação da criatividade por meio da produção de metáforas. *Psico-USF*, Itatiba, v. 12, n. 2, p. 197-210, 2007.

QUINTANA, H. E. O portfólio como estratégia para a avaliação. In: BALLESTER, M. (Org.). *Avaliação como apoio à aprendizagem.* Porto Alegre: Artmed, 2003. cap 16, p.163-173.

ROMBERG, A. *Intraclass correlation coefficients.* Reliability and more. 2009. Disponível em: <http://www.docstoc.com/docs/112692917/Intraclass-correlation-coefficients>. Acesso em: 07 jan. 2012.

SANTOS, T. M. de B. M. dos; MONTEIRO, V. R. V.; JUNIOR, J. F. R. Confiabilidade dos julgamentos de avaliadores de prova escrita na seleção para o mestrado. *Est. Aval. Educ.*, São Paulo, v. 21, n. 46, p. 363-374, maio/ago. 2010.

SCHUSTER, C. A note on the interpretation of weighted kappa and its relations to other rater agreement statistics for metric scales. *Educational and Psychological Measurement*, v. 64, p. 243-253, 2004.

STEMLER, S. E. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, v. 9, n. 4, 2004.

STUFFLEBEAM, D. L. (Org.). *Educational evaluation & decision making*. Bloomington: Phi Delta Kappa, 1971.

TINSLEY, H. E. A.; WEISS, D. J. Interrater reliability and agreement. In: TINSLEY, H. E. A.; BROWN, S. D. (Ed.). *Handbook of applied multivariate statistics and mathematical modeling*. New York: Academic Press, 2000. p. 95-124.

URBINA, S. *Fundamentos da testagem psicológica*. Porto Alegre: Artmed, 2007.

VENTURA, M. M.; BOTTINO, C. M. C. Estudo de confiabilidade da versão em português de uma entrevista estruturada para o diagnóstico de demência. *Revista da Associação Médica Brasileira*, São Paulo, v. 47, n. 2, p. 110-116, 2001.

DANIEL ABUD SEABRA MATOS

Doutor em Educação pela Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brasil. Doutorado sanduíche na University of Florida, Estados Unidos. Professor do Departamento de Educação da Universidade Federal de Ouro Preto (Ufop), Ouro Preto, Minas Gerais, Brasil

danielmatos@ichs.ufop.br

Recebido em: MAIO 2014
Aprovado para publicação em: OUTUBRO 2014

AGRADECIMENTOS

Aos pareceristas que contribuíram, durante o ano de 2014, avaliando e selecionando artigos para publicação na revista, para garantir a qualidade de seu padrão científico.

Adelina de Oliveira Novaes
(Fundação Carlos Chagas – FCC, São Paulo, São Paulo, Brasil)

Adolfo Ignácio Calderon
(Pontifícia Universidade Católica de Campinas –PUC-Campinas, Campinas, São Paulo, Brasil)

Alda Luiza Carlini
(Pontifícia Universidade Católica de São Paulo – PUC/SP, São Paulo, São Paulo, Brasil)

Amélia Cristina Artes
(Fundação Carlos Chagas – FCC, São Paulo, São Paulo, Brasil)

Ana Lucia Manrique
(Pontifícia Universidade Católica de São Paulo – PUC/SP, São Paulo, São Paulo, Brasil)

Ana Lúcia Souza de Freitas
(Pontifícia Universidade Católica do Rio Grande do Sul – UFRGS, Porto Alegre, Rio Grande do Sul, Brasil)

Ana Maria Martins da Costa Santos
(Universidade Estadual Paulista Júlio de Mesquita Filho – Unesp, Presidente Prudente, São Paulo, Brasil)

Ana Maria Saul
(Pontifícia Universidade Católica de São Paulo – PUC/SP, São Paulo, São Paulo, Brasil)

Anamérica Prado Marcondes
(Pontifícia Universidade Católica de São Paulo – PUC/SP, São Paulo, São Paulo, Brasil)

Tufi Machado Soares
(Universidade Federal de Juiz de Fora – UFJF, Juiz Fora, São Paulo, Brasil)

Valéria Virgínia Lopes
(Universidade de São Paulo – USP, São Paulo, São Paulo, Brasil)

Vanda Mendes Ribeiro
(Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária – Cenpec; Universidade da Cidade de São Paulo – Unicid, São Paulo, São Paulo, Brasil)

Vanderli Fava de Oliveira
(Universidade Federal de Juiz de Fora – UFJF, Juiz de Fora, Minas Gerais, Brasil)

Yara Lúcia Esposito
(Fundação Carlos Chagas – FCC, São Paulo, São Paulo, Brasil)

INSTRUÇÕES A COLABORADORES

NORMAS GERAIS

Estudos em Avaliação Educacional publica trabalhos inéditos referentes à educação, apresentados sob a forma de relatos de pesquisa, ensaios teóricos, metodologias, revisões críticas, artigos e resenhas.

Excepcionalmente, serão aceitos trabalhos que tenham sido publicados em periódicos estrangeiros (com a indicação da fonte), os quais serão submetidos à mesma avaliação dos artigos inéditos. O autor deverá apresentar a autorização da revista em que seu artigo tenha sido originalmente publicado.

Os originais recebidos são apreciados por especialistas da área e pelo Comitê Editorial, mantendo-se em sigilo a autoria dos textos.

Os autores recebem comunicação relativa aos pareceres emitidos. O Comitê Editorial reserva-se o direito de recusar o artigo ao qual foram solicitadas ressalvas, caso não sejam atendidas a contento.

Se a matéria for aceita para publicação, a revista permite-se introduzir pequenas alterações formais no texto, respeitando o estilo e a opinião dos autores.

Os trabalhos não deverão ser publicados em qualquer outra forma antes de decorridos seis meses de sua publicação em *Estudos em Avaliação Educacional*.

Artigos de um mesmo autor só será publicado com intervalo de, pelo menos, seis meses.

Solicita-se do(s) autor(es): nome completo, vínculo institucional: última ocupação profissional: cargo e filiação (empresa, instituição ou organização); unidade de referência (da ocupação profissional): Faculdade/Instituto, Programa de Graduação/Pós-graduação, Departamento; titulação (graduado, especialista, mestre/mestrando, doutor/doutorando), endereço, telefone, celular e correio eletrônico. Pede-se, ainda, que o autor indique como seu nome deve constar da publicação, bem como o nome completo da instituição à qual está vinculado e seu e-mail de contato.

Os autores receberão três exemplares impressos da revista em que seus textos forem publicados.

APRESENTAÇÃO DOS ORIGINAIS

Para submeter um artigo a *Estudos em Avaliação Educacional*, é necessário:

- Estar cadastrado no Portal de Periódicos da Fundação Carlos Chagas.. Se não estiver, acesse para fazer o cadastro: <http://publicacoes.fcc.org.br/ojs/> (o login e a senha serão a chave para o acesso ao sistema).

- Encaminhar duas versões do artigo: uma para avaliação, em PDF, sem informações que permitam identificar a autoria; e outra, em Microsoft Word, com todas as informações.

Caso o artigo seja em coautoria, o ideal é que todos os autores estejam cadastrados no sistema. É possível, no entanto, o envio do texto apenas com o cadastro de um dos autores. Em ambos os casos, a pessoa que envia o arquivo precisa incluir os coautores no Passo 3. Metadados da Submissão → Incluir Autor.

Para acompanhar o *status* da submissão, deve-se acessar o sistema → menu → acesso → login e senha.

A primeira página do texto deve trazer o título do trabalho e omitir o nome do autor e a filiação institucional, a fim de assegurar o anonimato no processo de avaliação.

Na **extensão**, os artigos não podem exceder 25 páginas (incluídos os anexos) e a extensão máxima das resenhas é de seis páginas, e devem ter o seguinte formato obrigatório: 3 cm de margem superior, 3 cm de margem inferior, 3 cm de margem esquerda e 2 cm de margem direita; parágrafo 1,25; com espaçamento de 1,5 entre as linhas; sem espaço (anterior ou posterior) entre os parágrafos, páginas enumeradas (após a folha de rosto, na margem inferior à direita), fonte em *Times New Roman*, no corpo 12.

Títulos e subtítulos devem ter, no máximo, 11 palavras (incluindo artigos, preposições, conjunções etc.).

Toda matéria, à exceção de resenhas, precisa vir acompanhada de **resumo** contendo no máximo 11 linhas, com espaçamento simples entre as linhas e sem espaço entre parágrafos, sem conter siglas nem referências, trazendo, em seu início, o título do trabalho. Ao final do resumo, indicar quatro palavras-chave (descritores) do conteúdo do texto.

Citações, remissões, notas e siglas devem obedecer às regras da ABNT (NBR 10520, 2002). As **citações** diretas (textuais), com até três linhas, devem ser incorporadas ao texto, entre aspas, sendo necessário indicar o sobrenome do autor, ano e número da página.

Ex.: Em função desses indicadores, "chegou-se à organização de cinco grandes grupos de escolas denominadas azul, verde, amarelo, laranja e vermelho" (SÃO PAULO, 2001, p. 55).

Citações com mais de três linhas deverão ir em um bloco abaixo do texto, sem aspas, com recuo de 1,25 cm, a partir da margem esquerda, com espaçamento simples entre as linhas e sem espaço entre parágrafos, em fonte *Times New Roman* e corpo 10, sendo necessário indicar o sobrenome do autor, ano e página. Ex.:

> [...] a sofisticação técnica da avaliação nacional – que hoje ocupa a atenção da cúpula decisória e de seus assessores – apresenta-se como entrave para a compreensão; tanto pelos atores dos sistemas e escolas como pela população em geral, do processo avaliativo realizado. (FREITAS, 2004, p. 685)

Na **citação de citação** deve ser empregada a expressão latina "apud" (citado por) para identificar a fonte que foi efetivamente consultada, a qual deve ter a referência completa no rodapé; e na lista de Referências incluir apenas a obra consultada (CARONE et al., 2003).

Ex.: Para Watson (apud CARONE et al., 2003) [...].

As **remissões bibliográficas** indiretas são incorporadas ao texto entre parênteses (ano).

Ex.: Segundo João Barroso (2006), todos...

As **notas explicativas** devem ser evitadas e utilizadas apenas quando for estritamente necessário, preferencialmente sem ultrapassar três linhas. Devem figurar sempre no rodapé da página, numeradas sequencialmente.

As **siglas** devem ser desdobradas quando mencionadas à primeira vez no artigo.

Ex.: Exame Nacional do Ensino Médio (Enem).

Tabelas, gráficos, quadros e figuras (assim como os **títulos** e as **fontes**) devem ser apresentados no corpo do texto, e não em caixas de texto, alinhados à esquerda, em sua página correspondente, numerados com algarismos arábicos, com títulos (posicionados acima, em corpo 12) padronizados quanto ao formato e termos utilizados. Abaixo destes, sem estarem em caixas de texto, deve, obrigatoriamente, ser indicada a fonte dos dados (remetida às referências bibliográficas), com autoria e ano, inclusive se for de elaboração própria dos autores, em corpo 10, alinhada à esquerda, espaço 1,5 entre linhas.

Tabelas, gráficos e quadros devem, ainda, ser enviados em um arquivo separado, em *software* compatível com o ambiente Windows, de preferência em Excel, e as figuras (ilustrações, imagens, mapas, fotos etc.) em arquivo com alta resolução (300 dpi), todos produzidos em preto e branco, em tamanho máximo de 10 cm de largura.

Referências de cunho bibliográfico devem vir ao final do texto, por ordem alfabética de sobrenome do autor, e, quando possível, fazer constar por extenso o prenome dos autores. Os títulos das obras devem vir em itálico.

Quando houver dois ou três autores, separa-se o primeiro autor e os demais por ponto e vírgula; ultrapassando três autores, faz-se a entrada pelo autor principal (referenciado no texto) e substitui-se os outros pela expressão "et al.".

Sua apresentação deve seguir as normas da ABNT (NBR 6023, 2002). A exatidão das referências e a correta citação no texto são de responsabilidade do(s) autor(es) dos artigos, sendo uma exigência para a publicação do trabalho.

DIREITO DE RESPOSTA

Estudos em Avaliação Educacional acolhe comentário(s) a artigo publicado na revista. Se o comentário for aceito para publicação, a revista oferecerá ao autor igual espaço para réplica, que poderá ser publicada no mesmo número do comentário ou no número subsequente. Ambos estão sujeitos ao mesmo processo de avaliação dos demais textos. Não são aceitos comentários ou réplicas a resenhas.