

Ae 78

ESTUDOS EM AVALIAÇÃO EDUCACIONAL • N. 2 JUL./DEZ. 1990 • FUNDAÇÃO CARLOS CHAGAS • SÃO PAULO

Em 2006 muda sua periodicidade para quadrimestral. E, a partir do n. 76, passa a ser apenas digital.

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 Popovic (*Brasil, FCC*)

www.fcc.org.br

BBE - Bibliografia Brasileira de Educação (*Brasil, Cibec/Inep/MEC*)

http://pergamum.inep.gov.br/pergamum/biblioteca/index.php?resolution=1024_1

Capex - Portal de Periódicos/Qualis (*Brasil*)

<http://www.periodicos.capes.gov.br/>; <http://qualis.capes.gov.br/webqualis/>

Clase - Citas Latinoamericanas en Ciencias Sociales y Humanidades (*México, Unam*)

<http://biblat.unam.mx/>

Diadorim - Diretório de Políticas de Acesso Aberto das Revistas Científicas Brasileiras (*Brasil, Ibict*)

<http://diadorim.ibict.br/handle/1/1027>

Dialnet - Fundación Dialnet (*Espanha, Universidad de La Rioja*)

<https://dialnet.unirioja.es/servlet/revista?codigo=24462>

DOAJ - Directory of Open Access Journals (*Suécia*)

<http://www.doaj.org/>

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>

Google Scholar - <http://scholar.google.com.br/>

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, el Caribe, España y Portugal (*México, Unam*)

<http://www.latindex.unam.mx>

REDIB - Red Iberoamericana de Innovación y Conocimiento Científico (*Espanha*)

http://www.redib.org/recursos/Serials/Record/oai_revista1741

Sumários.org - Sumários de Revistas Brasileiras (*Fundação de Pesquisas Científicas de Ribeirão Preto - FUNPEC-RP, Brasil*)

http://www.sumarios.org/periodicos?area=5&title_peri=estudos+em+avalia%C3%A7%C3%A3o+educacional

UlrichsWeb Global Serial Directory

<http://ulrichsweb.serialsolutions.com>

VERSÃO ELETRÔNICA

www.fcc.org.br

E-MAIL

eae@fcc.org.br

PROJETO GRÁFICO

Gustavo Piqueira e Samia Jacintho / Casa Rex

DIAGRAMAÇÃO

Núcleo de Design / Fundação Carlos Chagas

ESTUDOS EM AVALIAÇÃO EDUCACIONAL

Periódico da Fundação Carlos Chagas criado em 1990 sucedendo Educação e Seleção (1980-1989).

Tem como objetivo divulgar a produção acadêmica relacionada com o campo da avaliação educacional, acolhendo trabalhos originários de diferentes áreas das ciências humanas, com perspectivas teóricas e metodológicas diversas, propiciando a troca de informações e o debate acadêmico. Publica trabalhos inéditos, apresentados sob a forma de artigos, ensaios teórico-metodológicos, relatos de experiência, entrevistas e resenhas.

A revista não se responsabiliza pelos conceitos emitidos em matérias assinadas.

Direitos autorais reservados: reprodução integral de artigos apenas com autorização específica; citação parcial permitida com referência completa à fonte.

COMITÊ EDITORIAL

Nelson Gimenes (*Editor-chefe*)
Adriana Bauer
Cláudia Oliveira Pimenta
Fabiana Silva Fernandes
Gabriela Miranda Moriconi
Vandré Gomes da Silva

EDITORES ASSOCIADOS

Daniel Abud Seabra Matos
Raquel Rangel de Meireles Guimarães
Rodrigo Rosistolato

CONSELHO EDITORIAL

Bernardete A. Gatti
(Fundação Carlos Chagas, São Paulo, São Paulo, Brasil)

Clarilza Prado de Sousa
(Fundação Carlos Chagas e Pontifícia Universidade Católica de São Paulo, São Paulo, São Paulo, Brasil)

Dalton Francisco de Andrade
(Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil)

Domingos Manuel de Barros Fernandes
(Universidade de Lisboa, Lisboa, Portugal)

Fernando Lang da Silveira
(Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil)

Joaquim José Soares Neto
(Universidade de Brasília, Brasília, Distrito Federal, Brasil)

José Francisco Soares
(Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil)

Lina Kátia Mesquita de Oliveira
(Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil)

Luzia Marta Bellini
(Universidade Estadual de Maringá, Maringá, Paraná, Brasil)

Maria Eugénia Ferrão
(Universidade da Beira Interior, Covilhã, Portugal)

Maria Inês Gomes de Sá Pestana
(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, Distrito Federal, Brasil)

Naura Syria Carapeto Ferreira
(Universidade Tuiuti do Paraná, Curitiba, Paraná, Brasil)

Nícia Maria Bessa
(Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil)

Nigel Pelham de Leighton Brooke
(Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil)

Paulo de Martino Jannuzzi
(Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, Rio de Janeiro, Brasil)

Robert Verhine
(Universidade Federal da Bahia, Salvador, Bahia, Brasil)

Sandra Zákia Sousa
(Universidade de São Paulo, São Paulo, São Paulo, Brasil)

Sérgio Vasconcelos de Luna
(Pontifícia Universidade Católica de São Paulo, São Paulo, São Paulo, Brasil)

Yara Lúcia Esposito
(Fundação Carlos Chagas, São Paulo, São Paulo, Brasil)

NÚCLEO DE DOCUMENTAÇÃO E DIFUSÃO CIENTÍFICA (NDDC)

EQUIPE DE EDIÇÃO

Adélia Maria Mariano da S. Ferreira – Coordenação
Camila M. Camargo de Oliveira – Assistente editorial
Gabriella Fernandes Rampinelli – Assistente administrativo
Lucas dos Santos Medrado – Auxiliar de Edições
Marcia Caxeta – Assistente editorial

NORMALIZAÇÃO BIBLIOGRÁFICA

Biblioteca Ana Maria Poppovic

REVISÃO ESTATÍSTICA

Miriam Bizzocchi
Raquel da Cunha Valle

SUMÁRIO

TABLA DE CONTENIDO

TABLE OF CONTENTS

ARTIGOS ARTÍCULOS ARTICLES

Teste adaptativo informatizado da Provinha Brasil – Leitura: resultados e perspectivas 525

Test adaptativo informatizado del Provinha Brasil – Leitura: resultados y perspectivas
A computerized adaptive test of Provinha Brasil - Leitura: results and perspectives

RODRIGO TRAVITZKI, OCIMAR MUNHOZ ALAVARSE, DOUGLAS DE RIZZO MENEGHETTI,
ÉRICA MARIA DE TOLEDO CATALANI

As provas do Exame Nacional do Ensino Médio são unidimensionais? 554

¿Las pruebas del Examen Nacional de la Secundaria Superior son unidimensionales?
Are the tests of the National Exam of Upper Secondary Education unidimensional?

LEANDRO ARAUJO DE SOUSA, JOSÉ AIRTON DE FREITAS PONTES JUNIOR, ADRIANA EUFRÁSIO BRAGA

O que diferencia escolas com bom desempenho nos anos finais do ensino fundamental? 577

¿Qué diferencia a las escuelas con buen desempeño en los últimos años de la enseñanza primaria?

What differentiates schools with good performance in the final years of elementary education?

LUIZ GUILHERME SCORZAFAVE, ERNESTO MARTINS FARIA, BRUNA ALVES

Evidências empíricas do impacto da alfabetização sobre o crescimento econômico 600

Evidencias empíricas del impacto de la alfabetización sobre el crecimiento económico
Empirical evidence of the impact of literacy on economic growth

VINÍCIUS DE AZEVEDO COUTO FIRME

Heterogeneidade do impacto do trabalho infantil no rendimento escolar do Paraná 626

Heterogeneidad del impacto del trabajo infantil en el rendimiento escolar de Paraná
Heterogeneity of the impact of child labor on school performance in Paraná

CARLOS ALBERTO RAMOS TORRES, MARCOS DE OLIVEIRA GARCIAS

Avaliação e políticas de resultados: condicionantes da atuação do coordenador 651

Evaluación y políticas de resultados: condicionantes de la actuación del coordinador
Evaluation and results policies: conditioning of the coordinator's performance

JEANNY MEIRY SOMBRA SILVA, LAURINDA RAMALHO DE ALMEIDA

Indicadores de matrículas do Público-Alvo da Educação Especial no estado de São Paulo 674

Indicadores de matrículas del Público-Meta de la Educación Especial en el estado de San Pablo

Indicators of enrollment of Special Education students in Sao Paulo state (Brazil)

RÉGIS HENRIQUE DOS REIS SILVA, JOICE MELO VIEIRA, JULIANA SALLES MADEIRA,
LUCIANA CRISTINA SALVATTI COUTINHO

Uma matriz de (auto)avaliação da internacionalização do ensino superior no Brasil 699

Una matriz de (auto)evaluación de la internacionalización de la educación superior en Brasil
A (self)assessment matrix of the internationalization of higher education in Brazil

GABRIEL BRITO AMORIM, KYRIA REBECA FINARDI

Contornos do Estado Avaliador no Brasil 723

Contornos del Estado Evaluador en Brasil

Outline of the Evaluative State in Brazil

MARILDA PASQUAL SCHNEIDER, ELINA RENILDE DE OLIVEIRA RIBEIRO

Uma análise dos resultados educacionais dos estados brasileiros 742

Un análisis de los resultados educacionales de los estados brasileños

Analysis of the educational results of Brazilian states

PEDRO JORGE HOLANDA ALVES, JEVUKS MATHEUS ARAÚJO, GUSTAVO SARAIVA FRIO, LYVIA CABRAL CORDEIRO

<https://doi.org/10.18222/eae.v0ix.7216>

TESTE ADAPTATIVO INFORMATIZADO DA PROVINHA BRASIL – LEITURA: RESULTADOS E PERSPECTIVAS

-  RODRIGO TRAVITZKI^I
 OCIMAR MUNHOZ ALAVARSE^{II}
 DOUGLAS DE RIZZO MENEGHETTI^{III}
 ÉRICA MARIA DE TOLEDO CATALANI^{IV}

^I Universidade São Francisco (USF), Campinas-SP, Brasil; r.travitzki@gmail.com

^{II} Universidade de São Paulo (USP), São Paulo-SP, Brasil; ocimar@usp.br

^{III} Centro Universitário da Fundação Educacional Inaciana “Padre Sabóia de Medeiros” (Centro Universitário FEI), São Bernardo do Campo-SP; douglasrizzo@fei.edu.br

^{IV} Universidade de São Paulo (USP), São Paulo-SP, Brasil; ericamtc@usp.br

RESUMO

Este artigo descreve um Teste Adaptativo Informatizado (TAI) da Provinha Brasil – Leitura, com base na Teoria da Resposta ao Item. Detalham-se o funcionamento e o desenvolvimento do algoritmo. O TAI foi aplicado com o uso de tablets a 1.983 alunos dos 1o e 2o anos do ensino fundamental, em 15 escolas da Rede Municipal de Ensino de São Paulo. Os resultados confirmam a qualidade dos itens da Provinha Brasil, do trabalho realizado nas escolas e, sobretudo, do TAI. Em relação à gestão do tempo de prova, conclui-se que há uma associação positiva entre proficiência e tempo, mas só até certo ponto; os alunos tendem a demorar mais nos itens mais difíceis; essa tendência é mais intensa nos alunos mais proficientes, confirmando a hipótese de que eles tendem a gerir melhor o tempo de prova.

PALAVRAS-CHAVE TESTE ADAPTATIVO INFORMATIZADO • AVALIAÇÃO DE COMPETÊNCIA • PROVINHA BRASIL • TEORIA DA RESPOSTA AO ITEM.

TEST ADAPTATIVO INFORMATIZADO DEL PROVINHA BRASIL - LEITURA: RESULTADOS Y PERSPECTIVAS

RESUMEN

Este artículo describe un Test Adaptativo Informatizado (TAI) de la Provinha Brasil– Leitura, con base en la Teoría de la Respuesta al Ítem. Se detalla el funcionamiento y el desarrollo del algoritmo. El TAI fue aplicado con el uso de tablets a 1.983 alumnos del primero y segundo año de la enseñanza primaria, en 15 Escuelas de la Red Municipal de Enseñanza en San Pablo, Brasil. Los resultados confirman la calidad de los ítems del la Provinha Brasil, del trabajo realizado en las escuelas y, especialmente, del TAI. En relación con la administración del tiempo del examen, se concluye que hay una asociación positiva entre proficiencia y tiempo, pero solo hasta cierto punto; los alumnos tienden a demorarse más en los ítems más difíciles; esta tendencia es más intensa en los alumnos más proficientes, confirmando la hipótesis de que ellos tienden a administrar mejor el tiempo del examen.

PALAVRAS-CHAVE TEST ADAPTATIVO INFORMATIZADO • EVALUACIÓN DE COMPETENCIA • PROVINHA BRASIL • TEORÍA DE LA RESPUESTA AL ÍTEM.

A COMPUTERIZED ADAPTIVE TEST OF PROVINHA BRASIL - LEITURA: RESULTS AND PERSPECTIVES

ABSTRACT

This article describes a Computerized Adaptive Test (CAT) of Provinha Brasil – Leitura,¹ based on Item Response Theory. We detail the operation and development of the algorithm. The CAT was administered by means of tablet computers to 1,983 students in the 1st and 2nd grades of primary education, in 15 schools of the Municipal Education System of São Paulo. Results confirm the quality of Provinha Brasil's items, of the work done in schools and, mainly, of the CAT. As to the management of test time, we found a positive association between proficiency and time, but only to a certain extent; students tend to take longer on the more difficult items; this tendency is stronger in more proficient students, thus confirming the hypothesis that they tend to manage test time better.

KEYWORDS COMPUTERIZED ADAPTIVE TESTING • COMPETENCY ASSESSMENT • PROVINHA BRASIL • ITEM RESPONSE THEORY.

1 Provinha Brasil – Leitura is a Brazilian standardized reading assessment test for primary education students.

INTRODUÇÃO²

Os testes padronizados, sobretudo por seu emprego nas avaliações externas, têm se tornado cada vez mais presentes em escolas públicas e, progressivamente, pode se aventar, inserem-se na cultura escolar brasileira como expressão de políticas educacionais que, mediante usos variados dos resultados, utilizam-se desses testes. Contudo, para que esse tipo de política possa contribuir para a melhoria da qualidade da educação, há diversos desafios a serem enfrentados. Alguns dizem respeito à interpretação dos resultados e sua utilização em estratégias de gestão de redes de ensino e de escolas e mais ainda no cotidiano da sala de aula. Outros se referem a dificuldades logísticas relacionadas à segurança e ao gerenciamento de grandes quantidades de papel. Um terceiro tipo de desafio diz respeito à qualidade técnica dos testes enquanto instrumentos de medida. O Teste Adaptativo Informatizado (TAI) é uma tecnologia de avaliação da aprendizagem capaz de contribuir para a superação desses desafios, principalmente os dos dois últimos tipos.

A ideia de um teste que se adapte à proficiência de cada indivíduo remonta à década de 1970, embora ainda não tenha se popularizado plenamente, sendo muito incipiente no Brasil. Nesse tipo de teste, cada indivíduo responde a um conjunto de itens, selecionado em função de seu domínio – sua proficiência – daquilo que está sendo avaliado no decorrer da aplicação do teste de tal modo que o teste só se define completamente ao final de sua realização, podendo-se, portanto, encontrar testes que diferem de respondente para respondente em decorrência de suas diferenças de proficiência. No início do teste, a cada respondente é apresentado um item, o qual pode, por exemplo, ser selecionado aleatoriamente ou escolhido de acordo com uma condição pré-determinada. Posteriormente, durante a realização do teste, ele se adapta a cada indivíduo, dependendo de seu desempenho (representado por seus acertos e erros), principalmente através da seleção de itens mais difíceis ou mais fáceis, e tem o objetivo tanto de otimizar a precisão da estimação de proficiência de cada indivíduo quanto de reduzir o tempo de aplicação, em contraste com os testes em papel ou apresentados em computador, considerados lineares, que desde o início estão prontos com uma determinada quantidade de itens, mesmo que apresentados sequencialmente. Essa é a essência de um teste adaptativo, que pode apresentar diversas variações e níveis de complexidade. Quando esse tipo de teste é aplicado em formato digital, mediante processos informatizados, denomina-se Teste Adaptativo Informatizado.

2 Parte deste trabalho foi apresentada no congresso “Teaching, Learning and E-learning (IAC-TLEI)” de 2018, em Viena. Agradecemos aos diversos profissionais da educação que colaboraram para o projeto piloto, ao Núcleo Técnico de Avaliação da Secretaria Municipal de Educação de São Paulo, à Faculdade de Educação da Universidade de São Paulo (USP), ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e à Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco).

Comparado a um teste linear convencional, o TAI apresenta algumas vantagens, como permitir: 1) igualar a precisão de medida para diferentes níveis de proficiência;³ 2) manter a precisão do teste reduzindo significativamente o número de itens; e 3) aumentar a precisão do teste caso se mantenha o número de itens (BARRADA, 2012). Evidentemente, a viabilidade de um TAI depende, como qualquer metodologia de avaliação, de um banco de itens com grande variabilidade e densidade de dificuldade em relação à escala de proficiência.

Em um teste adaptativo, após a resposta dada ao primeiro item, são selecionados itens de acordo com as respostas anteriores. Assim, a cada resposta, adiciona-se alguma informação sobre a medida de proficiência do respondente, processo pelo qual se diminui a incerteza (erro) da estimativa de proficiência. Em outras palavras, o instrumento de medida – o teste – torna-se mais preciso a cada resposta. Adicionalmente, a escolha de itens em tempo real permite a coleta de informações mais qualificadas sobre a proficiência dos alunos em pontos menos explorados da escala, como os extremos superior e inferior.

Para se compreender melhor essa ideia, basta imaginar um teste linear apresentado a um aluno que tenha baixa proficiência e tenha acertado seis itens fáceis que lhe foram apresentados no conjunto de itens do teste. Nessa situação, o acerto de mais um item fácil adicionaria pouca informação ao conjunto, uma vez que já se tenha considerável certeza de que o aluno tem um domínio menor dentre as habilidades do construto em questão – o objeto do qual se busca estimar o domínio. Esse novo item não deverá alterar a estimativa da proficiência, tampouco reduzirá a incerteza da estimativa. Esta é uma das grandes vantagens dos testes adaptativos: otimizar a coleta de informação sobre a proficiência do aluno, evitando realizar medidas, mediante a apresentação de mais itens, desnecessárias, pois

Em um formato TAI, a seleção de itens e a estimativa de proficiência seguem de mãos dadas. A eficiência na estimativa de proficiência está fortemente relacionada à seleção de itens apropriados para um indivíduo. De maneira circular, a adequação dos itens para um indivíduo depende em grande parte da qualidade das estimativas de proficiência intermediárias.⁴ (LINDEN; GLAS, 2010, p. 4, tradução nossa)

3 Tendo em vista que, em testes convencionais, há uma imprecisão maior nas estimativas de proficiência nos extremos inferior e superior da escala.

4 Do original: “*Within a CAT format, item selection and ability estimation proceed hand in hand. Efficiencies in ability estimation are heavily related to the selection of appropriate items for an individual. In a circular fashion, the appropriateness of items for an individual depends in large part on the quality of interim ability estimates*”.

No TAI da Provinha Brasil – Leitura, objeto deste artigo, utilizou-se um banco de 39 itens fornecido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e buscou-se construir um teste com menos itens e maior precisão para cada aluno, em média, do que o teste linear original aplicado em papel composto de 20 itens de múltipla escolha. Ademais, além de diminuir a quantidade de itens e aumentar a precisão, procurou-se resolver um problema que não se encontrava abordado na literatura, ou seja, que, em uma avaliação da aprendizagem com função formativa, o critério de parada do teste levasse em conta não apenas a incerteza relativa à proficiência estimada, mas também a incerteza relativa ao nível de proficiência atribuído ao examinando, ou seja, sua classificação em um dos segmentos da escala de proficiência. Sobretudo porque a cada nível de proficiência corresponde uma interpretação pedagógica, e a perspectiva do teste é que seu resultado seja utilizado por professores em suas atividades cotidianas, quando é mais importante saber a descrição do que o aluno é capaz de realizar naquele nível do que receber um número – a proficiência estimada –, que apenas permitiria comparar o desempenho dos alunos entre si. A partir dessa informação mais qualitativa, os professores podem planejar atividades didáticas para atingir patamares superiores. Adicionalmente, a avaliação – julgamento sobre o desempenho do aluno – é feita segundo critérios que relacionam desempenho e ano de escolarização com esses níveis de interpretação pedagógica.

Se a ciência dos testes adaptativos ainda está evoluindo, revelando seu grande potencial, há também consideráveis desafios quanto à complexidade dos procedimentos estatísticos, em comparação a dos testes lineares (LINDEN; GLAS, 2010). É necessário, por exemplo, que haja um ou mais critérios para a seleção de itens, e a escolha de um critério pode depender de fatores como os objetivos do teste, a extensão do banco de itens e a distribuição das dificuldades de seus itens. É importante considerar, ainda, questões relativas à segurança e sustentabilidade do banco de itens, que podem ser otimizadas com diferentes técnicas de controle da taxa de exposição do item (BARRADA, 2010). Além disso, o TAI precisa de um ou mais critérios de parada. Os critérios normalmente utilizados são: 1) alcançar um número predeterminado de itens; 2) alcançar um mínimo de incerteza na estimativa da proficiência; e 3) alcançar limite mínimo de informação que um novo item adicionaria à estimativa da proficiência (BARRADA, 2012).

Como ponderado anteriormente, a qualidade e o tamanho do banco de itens são um aspecto importante e desafiador, cuja solução depende dos objetivos do teste e de iniciativas específicas de elaboração de itens. Para Barrada (2012), há quatro objetivos gerais para o TAI, aos quais é possível dar maior ou menor importância: 1) confiabilidade da estimativa da proficiência; 2) segurança do banco de itens; 3) restrições de conteúdo; e 4) manutenção do banco de itens. Alguns desses objetivos

se contrapõem, como, por exemplo, (1) em relação a (2) e (4). Com efeito, há um *trade-off* entre precisão da medida e segurança do banco de itens,⁵ que pode ser minimizado com outros métodos de seleção de itens, em detrimento da seleção do item mais informativo a cada momento do teste (GEORGIADOU; TRIANTAFILLOU; ECONOMIDES, 2007).

O TAI descrito neste artigo é uma versão adaptativa informatizada da Provinha Brasil, um instrumento padronizado disponibilizado pelo Ministério da Educação (MEC), criado em 2007 (BRASIL, 2007a), para avaliação da proficiência em leitura, cujos resultados são expressos numa escala de cinco níveis, numerados de 1 a 5. O TAI aqui descrito foi aplicado em 15 escolas de ensino fundamental da Rede Municipal de Ensino de São Paulo. O artigo tem uma primeira seção na qual há uma breve descrição da Provinha Brasil – Leitura, depois é focalizado o funcionamento do algoritmo do TAI, que é o componente responsável pela parte “adaptativa” do teste. Na terceira seção, são tratadas as simulações computacionais utilizadas na construção do algoritmo e no ajuste inicial de alguns parâmetros. Na seção seguinte, são apresentados os resultados da aplicação do TAI, que confirmaram a validade do instrumento. Na quinta seção, é tratada a aplicação experimental do TAI, na qual é analisada, entre outros resultados, uma questão mais geral a partir dos dados coletados, relativa à forma de os alunos gerirem o tempo de prova. Na conclusão, estão condensados os resultados do TAI da Provinha Brasil – Leitura e algumas indicações para a continuidade de pesquisas assemelhadas.

A PROVINHA BRASIL - LEITURA

O processo de alfabetização e letramento é especialmente importante nos anos iniciais do ensino fundamental, momento em que professores destinam grande parte do ensino para o desenvolvimento da competência leitora e escritora das crianças (SOARES, 2016). Os indicadores nacionais têm mostrado resultados insatisfatórios no desenvolvimento dessas competências para a totalidade da população brasileira em idade escolar. Nesse contexto, a construção do TAI voltou-se para a Provinha Brasil – Leitura,⁶ desenvolvida pelo Inep para os alunos do 2º ano do ensino fundamental a partir do Plano de Desenvolvimento da Educação (PDE) (BRASIL, 2007b), em consonância com as recomendações de organismos internacionais para a “década da alfabetização” (2003 a 2012) (GONTIJO, 2012).

5 A segurança do banco de itens se refere, por exemplo, à taxa de exposição dos itens. Quanto mais um item é apresentado a uma população, menos informativo ele tende a ser. Especialmente se o teste proporciona recompensas (*high stakes*).

6 A Provinha Brasil incorporou a competência em Matemática a partir de 2011, mas o projeto do TAI concentrou-se na leitura, considerando sua importância no processo de escolarização.

A construção de uma versão adaptativa e informatizada da Provinha Brasil – Leitura decorreu de projeto desenvolvido pelos pesquisadores do Grupo de Estudos e Pesquisas em Avaliação Educacional (Gepave), vinculado à Faculdade de Educação da Universidade de São Paulo (Feusp), em parceria, no momento inicial do projeto, com o Inep e, durante todo seu desenvolvimento, com a Secretaria Municipal de Educação de São Paulo (SME-SP), mediante seu Núcleo Técnico de Avaliação (NTA), envolvendo gestores centrais, regionais (supervisores escolares), diretores e coordenadores pedagógicos de escola, professores e alunos.⁷ Dar a devida importância à questão da leitura no desenvolvimento de um projeto pedagógico consequente requer, entre outros elementos, cuidar dos procedimentos avaliativos (SOARES, 2016), considerando a perspectiva formativa do uso de seus resultados, sendo a Provinha Brasil – Leitura um instrumento construído nessa perspectiva, no qual, com base numa Matriz de Referência, são consideradas as habilidades concernentes à alfabetização, compreendida como desenvolvimento da compreensão das regras de funcionamento do sistema de escrita alfabética, e ao letramento inicial, entendido como apreensão de possibilidades de usos e funções sociais da linguagem escrita. Alfabetização e letramento, no escopo da Provinha Brasil, são abordados como processos complementares e paralelos, e os itens que compõem os testes buscam abarcá-los no espectro de dificuldade do teste. Importante salientar que os itens são pré-testados nacionalmente, além de a matriz estar referenciada em vários documentos do MEC voltados para a formação de docentes dos anos iniciais do ensino fundamental. Com isso, buscou-se garantir tanto a fidedignidade dos resultados quanto a validade para fins de utilização de seus resultados, condição para que possam ser integrados plenamente no processo de ensino nos primeiros anos da escolarização.

Aplicada desde 2008 e descontinuada em 2016, a Provinha Brasil compõe-se de dois testes, sendo o primeiro para aplicação em março (início do ano letivo) e o segundo, em outubro (final do ano letivo), ambos contendo 20 itens de múltipla escolha elaborados segundo uma matriz de especificações para leitura, pré-testados, conforme normas estatísticas, calibrados por especialistas do Inep e disponibilizados aos professores do país, conjuntamente com orientação para aplicação, “correção” – na verdade, processamento das respostas – e interpretação dos resultados.

Esses resultados iniciais, constituídos pela quantidade de acertos dos alunos nos testes, são, posteriormente, expressos numa escala de proficiência em leitura com cinco níveis de proficiência, com uma interpretação pedagógica para cada nível e com sugestões de trabalho docente para avanços dos alunos. A qualidade desse processo está apoiada no pré-teste dos itens para que se possa efetuar essa

7 Mais detalhes do projeto podem ser encontrados em Alavarse *et al.* (2018) e Catalani (2019).

associação entre quantidade de acertos e proficiência, tendo sido confirmada, no caso do TAI em epígrafe, em discussões com mais de 100 professores das escolas participantes do projeto.

Na versão em papel da Provinha Brasil, os próprios professores aplicam o teste, tabulam as respostas e interpretam os resultados, sendo todos os itens do teste disponibilizados para os professores e gestores a cada aplicação. Essa disponibilização, segundo opiniões de professores, permite maior compreensão dos resultados dos alunos, como se apreende num estudo de caso no município de Camaragibe, que descreveu uma interessante apropriação do instrumento, destacando ainda a importância do trabalho constante dos professores e gestores para os bons resultados relatados (MORAIS; LEAL; ALBUQUERQUE, 2009). Ainda que não seja o foco deste trabalho, é importante registrar dissensos em relação à concepção de leitura implícita na matriz de referência e na própria estrutura do teste da Provinha Brasil, como encontrado em Gontijo (2012). Entretanto, como apontado, são constatadas potencialidades da Provinha Brasil no apoio ao trabalho docente, pois o resultado da aplicação não se resume a um número, como seria a contagem de acertos ou mesmo o nível de proficiência do aluno, pois, no material de aplicação de cada teste, cada nível é acompanhado de uma descrição do que o aluno é capaz de fazer em termos de alfabetização e letramento inicial, e, ainda, são feitas algumas sugestões de intervenção docente para que o aluno possa avançar para o nível superior.

Importante salientar que, no banco de itens utilizado, cada item estava associado a um descritor da Matriz de Leitura, isto é, a cada elemento dessa matriz que, em conjunto, descreve o que seria o construto “leitura”, cuja proficiência se pretende avaliar e que, nos documentos da Provinha Brasil, é considerada uma

[...] atividade que depende de processamento individual, mas se insere num contexto social e envolve [...] capacidades relativas à decifração, à compreensão e à produção de sentido. A abordagem dada à leitura abrange, portanto, desde capacidades necessárias ao processo de alfabetização até aquelas que habilitam o(a) estudante à participação ativa nas práticas sociais letradas, ou seja, aquelas que contribuem para o seu letramento. Isso implica que o(a) estudante desenvolva, entre outras habilidades, as de ler palavras e frases, localizar informações explícitas em frases ou textos, reconhecer o assunto de um texto, reconhecer finalidades dos textos, realizar inferências e estabelecer relações entre partes do texto. (BRASIL, 2016, p. 9)

Nesses termos, a Provinha Brasil, além de se constituir em material sob maior controle dos professores para fins de sua aplicação e tratamento das respostas,

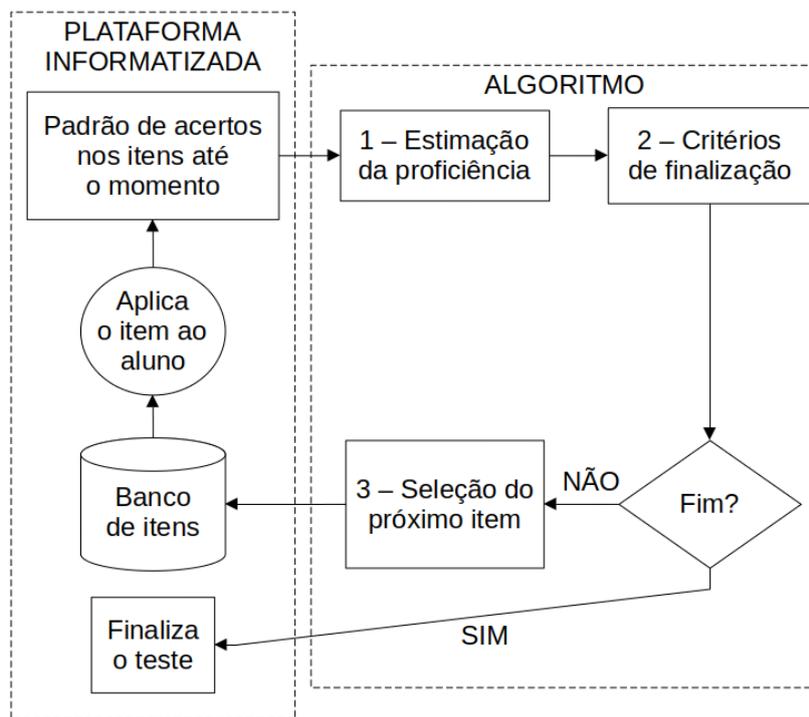
explicita seus fundamentos, o escopo de seus resultados interpretados e, ainda, fornece alternativas de desdobramentos didáticos, o que favorece o debate nas escolas, inclusive por ser mais transparente do que outras avaliações da aprendizagem externas.

O FUNCIONAMENTO DO ALGORITMO DO TAI

São diversos os desafios para viabilizar um teste de avaliação da aprendizagem. No caso de um TAI, além dos tradicionais, relacionados à qualidade dos itens e validade do teste como um todo, dois outros são adicionados: ser informatizado e adaptativo, visando a promover benefícios de ordem logística e de precisão de medida de proficiência. Para enfrentar esses dois desafios, optamos por construir dois módulos relativamente independentes: uma interface amigável para o usuário (criada com tecnologia *web* utilizando Java como principal linguagem de programação) e um sistema de processamento estatístico e psicométrico em tempo real (desenvolvido na linguagem R). O primeiro módulo permite ao teste ser informatizado, enquanto o segundo o torna adaptativo. A interface será aqui referida como “plataforma informatizada”, enquanto o processamento estatístico será denominado de “algoritmo”.

A plataforma informatizada é responsável pela exibição de itens aos respondentes e pela captura automática das respostas. Funciona *online* e foi acessada através de uma *intranet* para todas as escolas participantes. A apresentação dos itens foi realizada através do uso de *tablets* conectados à rede *Wi-Fi* das escolas. A leitura das comandas dos itens aos alunos, que na versão original da Provinha deve ser feita pelo professor aplicador, foi feita mediante essa plataforma e disponibilizada individualmente aos alunos, ao usarem os *tablets*, equipados com fones de ouvido e API de Fala do Google. A Figura 1 representa, de forma simplificada, os componentes do TAI da Provinha Brasil – Leitura e suas inter-relações.

FIGURA 1 - Representação simplificada dos componentes do TAI da Provinha Brasil - Leitura e de suas inter-relações



Fonte: Elaboração dos autores.

O algoritmo tem como fundamento teórico a Teoria da Resposta ao Item (TRI), descrita, entre outros autores, por Baker (2001), e seu objetivo é proporcionar uma dinâmica adaptativa à plataforma informatizada de testes. Mais especificamente, buscou-se otimizar a precisão da medida e minimizar o número de itens do teste, evitando perdas na validade do instrumento e no tempo de processamento computacional.

O algoritmo engloba três componentes, cada qual caracterizado por uma sequência de ações, descritas a seguir.

1) Estimação de proficiência:

- a) recebe como *input* da plataforma as respostas de cada aluno – configurando um padrão de acertos – e a distribuição *a priori* da proficiência, além dos parâmetros do banco de itens;
- b) estima a proficiência e o erro padrão, utilizando um método bayesiano que incorpora a distribuição *a priori*.

2) Critério de finalização do teste:

- a) verifica se o teste alcançou o limite máximo de itens;
- b) verifica se o erro padrão é menor do que o limite máximo definido;
- c) verifica se o nível de proficiência foi identificado de forma confiável;
- d) se ao menos uma das três condições anteriores for verdadeira, e se o teste já ultrapassou o limite mínimo de itens, envia um *output* para a plataforma terminar o teste; caso contrário, segue com o terceiro passo.

3) Seleção do próximo item do teste:

- a) identifica o descritor da matriz menos representado entre os itens aplicados ao aluno até o momento;
- b) busca os itens desse descritor no banco de itens;
- c) busca o item mais informativo desse subconjunto, levando em conta a proficiência estimada na etapa 1;
- d) retorna o item selecionado como *output* para a plataforma (juntamente com a proficiência e erro padrão estimados, para que a plataforma possa incluí-los como informação na próxima vez que acionar o algoritmo, depois que o item selecionado for respondido).

A estimação da proficiência é realizada com base na distribuição esperada *a posteriori* – *Expected a Posteriori* (EAP) (BOCK; MISLEVY, 1982) – com 21 pontos de quadratura. Os critérios para seleção de itens incluem: 1) a Máxima Informação de Fisher (MFI) (BARRADA, 2010); e 2) a seleção equilibrada de itens de cada eixo de conteúdo da matriz.

A seleção equilibrada de itens entre os descritores da matriz de referência tem o objetivo de evitar a perda de validade da prova devido à seleção dos itens promovida pelo TAI. De fato, trata-se de um cuidado importante para evitar um efeito colateral do critério de seleção de itens por máxima informação de Fisher. O algoritmo mantém sempre uma proporção semelhante de itens de cada eixo de conteúdo para garantir que o teste adaptativo represente a matriz desejada. Contudo, embora tenha sido implementada no algoritmo, a seleção equilibrada de itens não foi aplicada nessa fase piloto, em virtude do número reduzido de itens no banco.

Para determinação do fim de teste, o algoritmo utiliza um critério misto, levando em conta três regras:

- a) número de itens do teste (mínimo de oito e máximo de 20 itens);
- b) limite permitido de incerteza (erro padrão menor que 35 pontos);
- c) grau de confiança na determinação do nível de proficiência da escala da Provinha Brasil – Leitura (85% de confiança, segundo nossas simulações).

Os dois primeiros critérios são amplamente utilizados em testes adaptativos (BARRADA, 2012). O terceiro critério (c) desenvolvido para o TAI da Provinha Brasil – Leitura constitui uma modificação do critério de parada utilizado nas avaliações com finalidade de classificar os sujeitos. A classificação é usada em situações de avaliação somativas, nas quais uma decisão do tipo sim ou não deve ser tomada, mais comum em exames e certificações (BABCOCK; WEISS, 2012; WEISS; KINGSBURY, 1984). Trata-se, aparentemente, de uma contribuição significativa desse projeto para o estado da arte do TAI.

SIMULAÇÕES DO TAI DA PROVINHA BRASIL - LEITURA

Esta seção apresenta as tecnologias utilizadas e a metodologia de simulação empregada no desenvolvimento do TAI da Provinha Brasil – Leitura.

Software e hardware

O algoritmo foi escrito na linguagem de programação R, especializada em estatística, de código aberto e livre. As simulações também foram realizadas nessa linguagem. Para ambas as finalidades, foram testados os pacotes: *catR* (MAGIS; RAÏCHE, 2012), *PP* (REIF, 2019) e *irtoys* (PARTCHEV, 2016). As simulações foram inicialmente realizadas em 2016 e refeitas em fevereiro de 2018, com novas versões dos pacotes, a fim de preservar a qualidade dos resultados apresentados. As versões dos pacotes utilizadas nas simulações de 2018, aqui apresentadas, foram: *irtoys* 0.2.0; *PP* 0.6.1; *catR* 3.13.

O computador utilizado para a realização das simulações foi um *notebook* com processador i7 de 4 núcleos de 2.50 GHz e 8 Gb de memória, com o Sistema Operacional Linux Mint. Não foi utilizado processamento paralelo ou aceleração via GPU.

Simulações

Para desenvolvimento do algoritmo, foram testados, via simulação, cinco métodos de seleção de itens (além da seleção aleatória) e sete métodos para estimação da proficiência. Os métodos foram testados quanto à precisão e velocidade. Além disso, as simulações permitiram o ajuste de dois parâmetros no algoritmo: o erro padrão máximo e o valor crítico do intervalo de confiança.

As simulações se fundamentam na função logística de dois parâmetros da TRI (ANDRADE; TAVARES; VALLE, 2000), que descreve a probabilidade de um indivíduo com proficiência conhecida acertar um item com parâmetros conhecidos. Embora tenham sido testados diferentes bancos, os resultados aqui descritos se referem ao banco de itens fornecido pelo Inep, composto de 39 itens com dois parâmetros definidos (dificuldade e discriminação).

Para cada situação, foram realizadas 1.000 simulações, cada uma com 1.000 participantes (com distribuição normal de proficiência, de média 500 e desvio padrão 100) respondendo a 20 dos 39 itens do banco original do Inep.

Para estimação da proficiência foram comparados quatro métodos, sendo um deles testado em vários pacotes, totalizando sete métodos na prática. Dois métodos se fundamentam no princípio da verossimilhança: a busca da máxima verossimilhança (LORD, 1980) e da verossimilhança ponderada (WARM, 1989). Os outros dois métodos utilizam estatística bayesiana: a distribuição esperada *a posteriori* (EAP, do inglês *Expected a Posteriori*) (BOCK; MISLEVY, 1982) e o estimador modal (BIRNBAUM, 1969).

A seguir, listamos os sete métodos comparados, provenientes dos três pacotes:

- 1) ML: máxima verossimilhança (pacote *catR*);
- 2) WL: verossimilhança ponderada (do pacote *catR*);
- 3) BM: estimador bayesiano modal (do pacote *catR*);
- 4) EAP: método EAP (função *thetaEst* do pacote *catR*);
- 5) eapC: método EAP (função *eapEst* do pacote *catR*);
- 6) eapI: método EAP (do pacote *irtoys*);
- 7) eapP: método EAP (do pacote *PP*).

Os métodos para seleção de itens testados, provenientes do pacote *catR* (MAGIS; RAÏCHE, 2012), foram os seguintes:

- a) *random*: seleção aleatória de itens do banco;
- b) MFI: seleciona o item com maior informação para a proficiência estimada até então, a partir da função de informação do item (ANDRADE; TAVARES; VALLE, 2000);
- c) *bOpt* (A regra de Urry): seleciona o item com nível de dificuldade mais próximo da proficiência estimada até então;
- d) *thOpt* (estratificação por Máxima Informação): adaptação do método MFI com objetivo de aumentar a segurança do banco de itens;
- e) O método progressivo (REVUELTA J.; PONSODA, 1998): o item é selecionado segundo dois elementos, um relativo à Máxima Informação e outro aleatório. Ao longo do teste, o elemento aleatório vai perdendo a importância. Isso promove maior segurança do banco de itens; e
- f) O método proporcional (BARRADA, 2010): o item é selecionado segundo probabilidades relacionadas à Informação de Fisher, também com objetivo de promover maior segurança do banco de itens.

Critérios para finalização de teste

O principal objetivo na elaboração do critério para finalização foi proporcionar um teste que tenha menos itens e maior precisão do que um teste similar, porém não adaptativo. Para tanto, foram considerados três critérios simultaneamente.

Em primeiro lugar, foi definido de antemão um limite máximo de 20 itens e um mínimo de oito. O limite mínimo garante a aplicação de ao menos quatro itens de cada um dos dois eixos da Provinha.⁸ Já o limite máximo garante a finalização do teste nos patamares do teste em papel da Provinha Brasil, mesmo que os demais critérios não sejam alcançados.

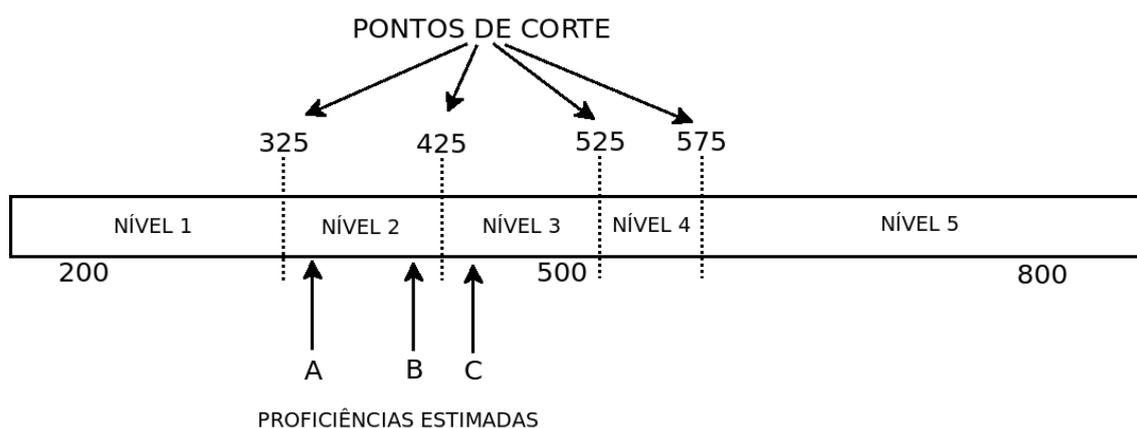
8 Os dois eixos utilizados na Provinha Brasil são: 1) apropriação do sistema de escrita; e 2) leitura. Contudo, em virtude do tamanho do banco disponível, optamos por não utilizar esse critério no piloto.

Em segundo lugar, foi definido um erro máximo na estimativa da proficiência, mensurado pelo erro padrão. Esse erro começa alto e vai diminuindo ao longo do teste, conforme mais itens vão sendo respondidos. O critério de finalização consiste em determinar um limite máximo permitido para o erro padrão, de acordo com os itens disponíveis, a população-alvo e os objetivos. E, finalmente, o critério “confiabilidade do nível de proficiência”, descrito a seguir.

Confiabilidade do nível de proficiência

Este critério busca otimizar o tamanho da prova, garantindo a alocação correta do aluno no nível de proficiência. Ele determina o encerramento do teste a partir do momento em que a proficiência e seu intervalo de confiança estão inteiramente contidos em um único dos cinco níveis de proficiência definidos para a Provinha Brasil (Figura 2). Vale ressaltar que os pontos que dividem a escala da Provinha Brasil em cinco níveis, também denominados pontos de corte, resultaram de um processo psicométrico (ancoragem) associado à análise pedagógica dos itens realizada por especialistas e educadores. É a descrição dos níveis – resultante desse processo – que permite a interpretação da nota do aluno (proficiência) com base na Teoria da Resposta ao Item. Embora esse método tenha sido identificado posteriormente na literatura como método de parada com objetivo de classificação, não foi encontrado o uso para fins de avaliação formativa na literatura sobre testes adaptativos.

FIGURA 2 – Pontos de corte e níveis de proficiência na escala da Provinha Brasil – Leitura



Fonte: Elaboração dos autores.

Com a adição desse critério de fim de teste, o TAI da Provinha Brasil pode ser encerrado mesmo que o erro na estimativa da proficiência de um aluno seja relativamente alto, desde que o intervalo que contenha a proficiência esteja completamente contido em um dos cinco níveis da Provinha Brasil – Leitura. Afinal, a Provinha não é um teste de seleção ou certificação, em que a precisão da

proficiência é mais relevante. O mais importante é saber se o nível de proficiência do aluno foi adequadamente identificado, priorizando assim o diagnóstico pedagógico e, conseqüentemente, a intervenção para melhoria da alfabetização. Em outras palavras, esse critério contribui para a interpretação e avaliação, que vão além da simples mensuração.

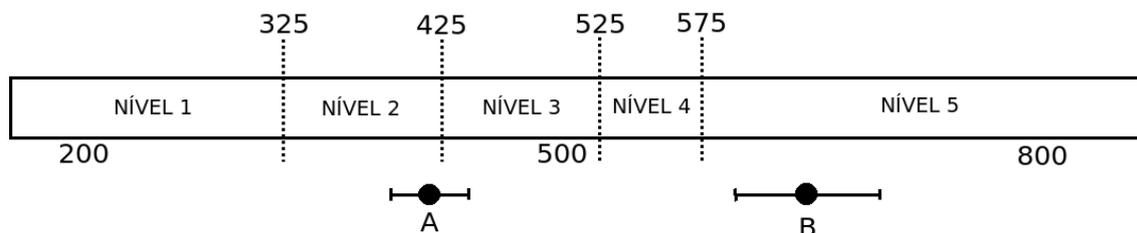
A título de exemplo, na Figura 2 é possível observar que a proficiência aferida para o aluno B está mais próxima da proficiência aferida para o aluno C do que a aferida para o aluno A. Contudo, levando-se em conta os pontos de corte definidos pela interpretação pedagógica da escala, os alunos A e B estão no mesmo nível da escala, enquanto o aluno C pertence ao nível seguinte. Em termos pedagógicos, isso significa que, supostamente, os alunos A e B demonstraram domínio de habilidades que exigem intervenções similares, enquanto o aluno C demonstrou domínio mais amplo que exige outras intervenções.

Cabe ressaltar que há sempre alguma incerteza na estimação da proficiência de um aluno, qualquer que seja o método utilizado. Tal incerteza depende principalmente da quantidade de itens respondidos e também da proximidade entre a dificuldade do item e a proficiência do respondente. Supondo a normalidade da distribuição de proficiência, a confiabilidade da estimativa (seu intervalo de confiança) pode ser determinada a partir do erro padrão. O intervalo de confiança é definido por um limite mínimo e um máximo e pode ser obtido com diferentes graus de confiança, multiplicando-se o erro padrão pelo valor crítico, que, por sua vez, depende do grau de confiança desejado. Para o algoritmo foi definido um grau de confiança de 85%, que corresponde a um valor crítico de 1,44 (FERREIRA, 2005). Isso significa que, quando o teste termina por esse critério, há 85% de chance de a proficiência verdadeira do aluno estar dentro do nível identificado pelo teste, segundo a Teoria da Resposta ao Item.

A Figura 3 ilustra a utilidade do critério de finalização do teste por confiabilidade do nível de proficiência como complemento ao critério de erro padrão máximo. O aluno B tem um erro (intervalo de confiança) maior do que o aluno A, porém seu nível de proficiência já foi estimado com segurança (nível 5) depois que ele acertou cinco itens em um teste adaptativo. O aluno A, por sua vez, respondeu a 11 itens nesse teste adaptativo, mas ainda não foi classificado de forma confiável, podendo pertencer aos níveis 2 ou 3. Afinal, o erro da estimativa não depende apenas do número de itens apresentados, mas também do padrão de acertos de cada aluno e dos parâmetros dos itens respondidos. Mais ainda, a quantidade e posicionamento dos pontos de corte interferem fortemente nesse critério. Com efeito, o aluno B foi beneficiado pela adição desse terceiro critério de finalização de teste: terminando mais rapidamente sem perda na precisão do teste, dado que a finalidade prática da Provinha Brasil – Leitura é fornecer uma medição confiável da proficiência

do aluno para que os professores, considerando os cinco níveis de proficiência, possam tomar decisões pedagógicas a partir de uma informação mais fidedigna. Destaca-se, sem entrar no mérito dessa decisão, que o Inep definiu como desejável que cada aluno esteja pelo menos no nível 4 ao final do 2º ano.

FIGURA 3 – Representação dos intervalos de confiança de duas proficiências estimadas



Fonte: Elaboração dos autores.

RESULTADOS DAS SIMULAÇÕES

Esta seção apresenta os resultados obtidos via simulação na etapa de construção do algoritmo e ajuste dos parâmetros básicos.

Métodos de seleção de itens e estimação de proficiência

Os métodos de seleção de itens e estimação de proficiência foram testados quanto à sua precisão e velocidade de processamento. Em relação à seleção de itens, todos os métodos se mostraram suficientemente rápidos. O Teste t revelou que todos os critérios apresentam menor erro do que a ausência (*random*, seleção aleatória de itens), mas que as diferenças entre os métodos não são significativas ($p < 0,05$) na maioria das simulações. Assim sendo, levando em conta a literatura especializada, optamos por incluir no algoritmo a seleção pelo método da Máxima Informação de Fisher. Contudo, se o TAI da Provinha Brasil – ou outro TAI – vier de fato a se consolidar como política pública, será necessário rever essa escolha técnica, pois ela não leva em conta a segurança e sustentabilidade do banco de itens. Os métodos conhecidos como progressivo ou proporcional poderiam ser mais adequados nesse caso.

Em relação aos métodos de estimação da proficiência, não foi observada diferença significativa ($p < 0,05$) na precisão dos métodos com menor erro (BM, EAP, eapC, eapI, WL) para um teste de 20 itens, com população de proficiência média igual a 500. Contudo, vale notar que, em populações com média diferente da esperada (600 ou 400, por exemplo), o método WL se mostrou mais preciso do que os outros, o que se deve à sua menor dependência de uma estimativa *a priori* da população.

Observou-se ainda que o método EAP do pacote PP (eapP) apresentou erro muito maior do que os outros. Isso ilustra outra função importante das simulações

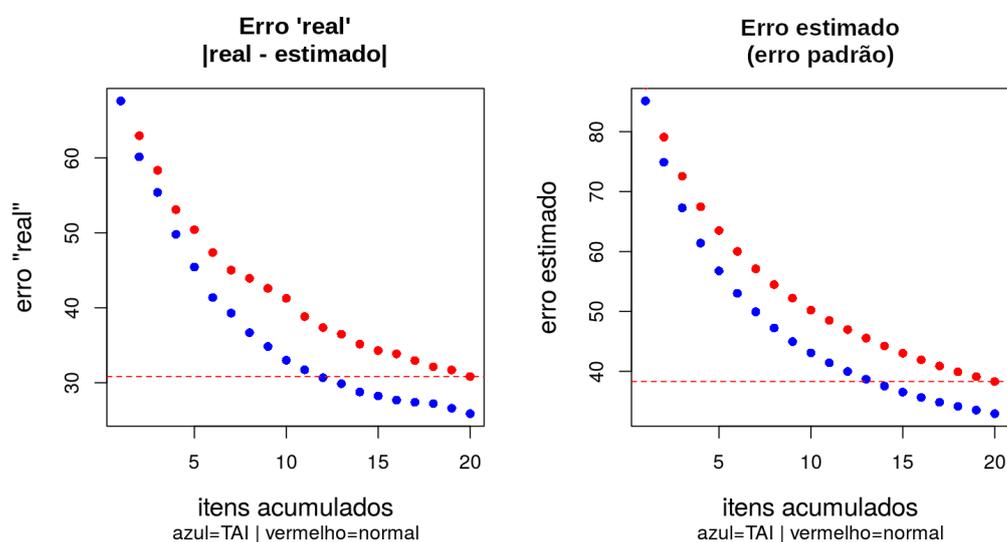
para o desenvolvimento de algoritmos, que é prevenir o uso de pacotes de qualidade duvidosa, com resultados inconsistentes. Tal cuidado é especialmente importante quando se trabalha com *software* livre, mas não deixa de ser necessário com *softwares* proprietários.

Em suma, levando em conta a precisão e a velocidade simultaneamente, o método escolhido para a estimação da proficiência foi o EAP do pacote *irtoys*. Para a seleção do próximo item, foi escolhido o método MFI, embora seja necessária uma revisão dessa escolha, caso haja necessidade de se preservar a segurança do banco de itens.

Critérios para fim de teste: erro padrão máximo

Para determinar o erro padrão máximo a ser aceito pelo TAI, consideramos o objetivo de produzir um teste que tenha, em média, maior precisão e menor tamanho do que um teste convencional, não adaptativo. A Figura 4 mostra os erros de estimativa em testes convencionais e adaptativos, segundo as simulações realizadas. O erro estimado é aquele que pode ser obtido pelo algoritmo do TAI a cada novo padrão de respostas do examinando. O erro “real”, por sua vez, não pode ser obtido pelo algoritmo. Só o conhecemos no contexto da simulação.

FIGURA 4 - Erros de estimativa em testes convencionais (normal) e adaptativos (TAI) de diferentes tamanhos de até 20 itens, segundo as simulações realizadas



Fonte: Elaboração dos autores.

Nota: A linha vermelha pontilhada indica o erro do teste normal com 20 itens.

Nos dois gráficos da Figura 4, a linha pontilhada de cor vermelha marca o erro alcançado pelo teste convencional depois de 20 itens. Isso corresponde a um erro padrão de 38 e uma diferença de proficiências de 31. Com efeito, os gráficos mostram que um teste adaptativo entre 12 e 13 itens tende a produzir estimativas

tão precisas quanto um teste convencional de 20 itens. Com base nessa observação, é possível delimitar um ponto médio; ou seja, um erro padrão correspondente a um teste convencional de 15 a 19 itens. Tal ajuste depende do equilíbrio desejado entre precisão e tamanho do teste em cada situação de avaliação.

É importante destacar que nos dois métodos de detecção do erro houve essa semelhança, entre 12 e 14 itens, confirmando a qualidade da estimativa obtida pelo *software* utilizado. Caso não houvesse tal semelhança, o algoritmo utilizado não teria qualquer garantia de corresponder à realidade.

Enfim, com o objetivo de proporcionar um teste menor e mais preciso, foi escolhido o limite máximo para o erro padrão: 35 pontos na escala da Provinha Brasil, que corresponderia a um teste adaptativo com 16 itens. Posteriormente, os resultados da aplicação do teste confirmaram essa previsão.

Confiabilidade do nível de proficiência

Para delimitação dos níveis de proficiência, foram utilizados os pontos de corte da Provinha Brasil – Leitura (Figura 2). Para o mesmo conjunto de dados pode-se obter intervalos de confiança mais ou menos largos, dependendo do nível de confiança que se deseja. A cada nível de confiança corresponde um valor crítico. Para ajuste do melhor valor crítico para o algoritmo da Provinha Brasil, foram testados quatro níveis de confiança.

Nota-se na Tabela 1 que, como seria esperado, quanto maior o nível de confiança desejado, menos testes serão finalizados segundo o critério da confiabilidade do nível de proficiência. É importante saber, contudo, quantos testes foram corretamente finalizados (comparando o nível de proficiência “real” e o estimado) em cada caso. Verificamos quantos dos testes finalizados por esse critério teriam sido bem-sucedidos na determinação do nível de proficiência do aluno.

TABELA 1 – Testes finalizados pelo critério de confiabilidade do nível de proficiência segundo o nível de confiança, em 1.000 simulações

NÍVEL DE CONFIANÇA	VALOR CRÍTICO	TESTES FINALIZADOS	FINALIZAÇÕES CORRETAS	TAXA DE ACERTO (%)
80%	1,28	354	297	83,9
85%	1,44	242	207	85,5
90%	1,645	114	109	95,6
95%	1,96	62	62	100,0

Fonte: Elaboração dos autores.

Esses resultados mostram que, na escala da Provinha Brasil – Leitura, há razoável correspondência entre o nível de confiança definido pelo valor crítico e o nível de confiança do critério confiabilidade do nível de proficiência. No algoritmo, definimos o valor crítico de 1,44.

RESULTADOS DA APLICAÇÃO EXPERIMENTAL DO TAI

Foram realizadas 1.983 aplicações do TAI da Provinha Brasil – Leitura, abrangendo 823 alunos do 1º ano e 1.160 alunos do 2º ano do ensino fundamental, distribuídos em 80 turmas de 15 escolas de ensino fundamental da Rede Municipal de Ensino de São Paulo.

TABELA 2 – Estatísticas descritivas da aplicação experimental do TAI da Provinha Brasil – Leitura a alunos do 1º e 2º ano do ensino fundamental da Rede Municipal de Ensino de São Paulo

ESTATÍSTICAS		PROFICIÊNCIA	ACERTOS	ITENS	DURAÇÃO	APLICAÇÕES
Total	Média	462,72	0,55	17,33	11,52	1983
	Erro padrão	1,84	0,00	0,05	0,12	
	Desvio padrão	81,93	0,20	2,18	5,53	
1º Ano	Média	416,83	0,44	16,74	10,88	823
	Erro padrão	2,14	0,01	0,06	0,20	
	Desvio padrão	61,41	0,15	1,60	5,75	
2º Ano	Média	495,28	0,63	17,75	11,98	1160
	Erro padrão	2,32	0,01	0,07	0,16	
	Desvio padrão	79,04	0,19	2,43	5,32	

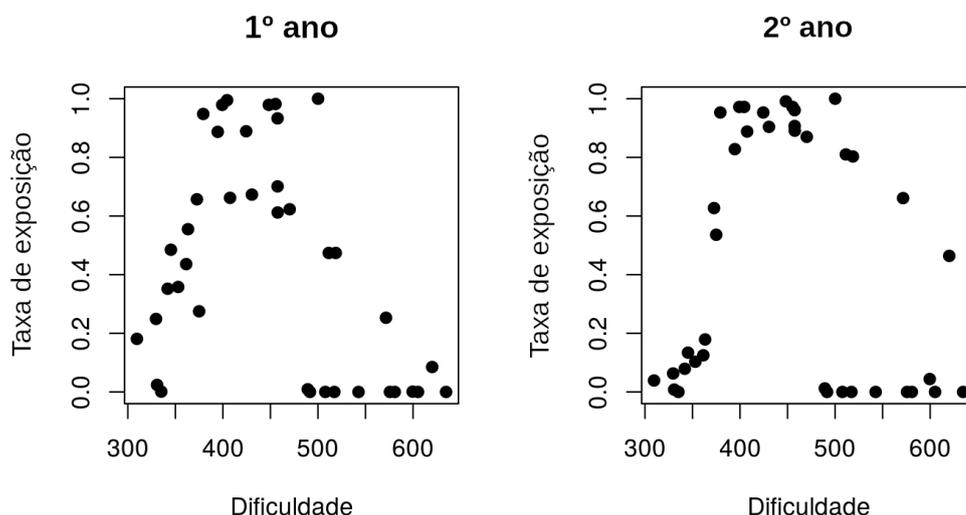
Fonte: Elaboração dos autores.

A Tabela 2 descreve quatro aspectos importantes da aplicação do TAI. Nota-se que a proficiência média estimada para o 2º ano (495,28 pontos) apresenta magnitude adequada, considerando-se que a escala da Provinha Brasil (projetada para o 2º ano) tem média 500 e desvio padrão 100. O desvio padrão (79,04), por sua vez, foi menor do que 100, o que reflete uma variância menor da população analisada em relação à população para qual a Provinha Brasil foi delineada, ou seja, alunos de todo Brasil. Nota-se, ainda, que a média do 1º ano foi menor do que a média da escala, o que também seria esperado, dado que os itens foram produzidos e calibrados para o 2º ano. Esses resultados confirmam, portanto, não apenas a qualidade do algoritmo utilizado no TAI, mas também a qualidade dos itens produzidos pelo Inep.

Em relação à taxa de exposição dos itens, o TAI de fato utilizou mais alguns itens do que outros. No total, foram aplicados apenas 31 itens para os alunos. Levando em conta que o banco possui 39 itens, conclui-se que houve aproveitamento de 79,5% do banco de itens. A Figura 5 busca relacionar a taxa de exposição e a dificuldade do item. Mais uma vez, os resultados empíricos corroboram os resultados esperados via simulação: além de uma taxa maior nos itens de dificuldade

média, observa-se também que, no 2º ano, houve um uso maior dos itens mais difíceis, e um uso menor dos itens mais fáceis. Nota-se, ainda, que a grande maioria dos itens não utilizados se encontra na porção superior da escala.

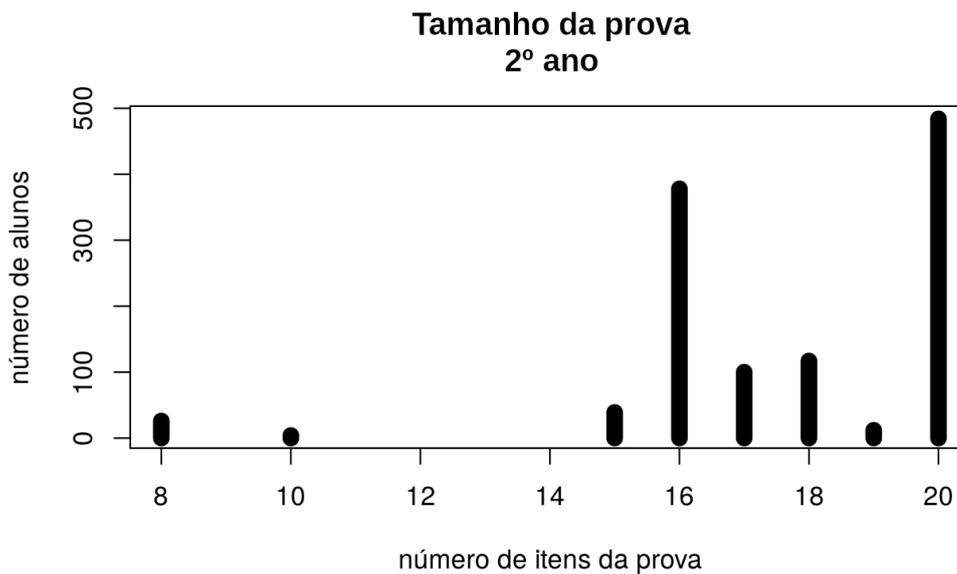
FIGURA 5 - Relação entre taxa de exposição e dificuldade dos itens para alunos do 1º e 2º anos do ensino fundamental



Fonte: Elaboração dos autores.

Em relação ao número de itens apresentados aos alunos, nota-se que boa parte dos testes terminou com 16 itens (Figura 6), como esperado pelas simulações e pela escolha do limite máximo de 35 pontos para o erro padrão da estimativa da proficiência. Além disso, alguns testes terminaram com dez ou menos itens, o que ocorreu em virtude do outro critério para finalização de teste, a confiabilidade do nível de proficiência. Nesse sentido, os resultados confirmam a importância desse critério como complemento ao critério do erro padrão máximo. Por outro lado, o elevado número de testes com 20 itens não foi previsto pelas simulações, revelando limitações no modelo utilizado. Levando em conta que o algoritmo visa a manter certo grau mínimo de incerteza, é possível que haja efeitos reais não previstos pelo modelo unidimensional de dois parâmetros da TRI, aqui utilizado como base das simulações.

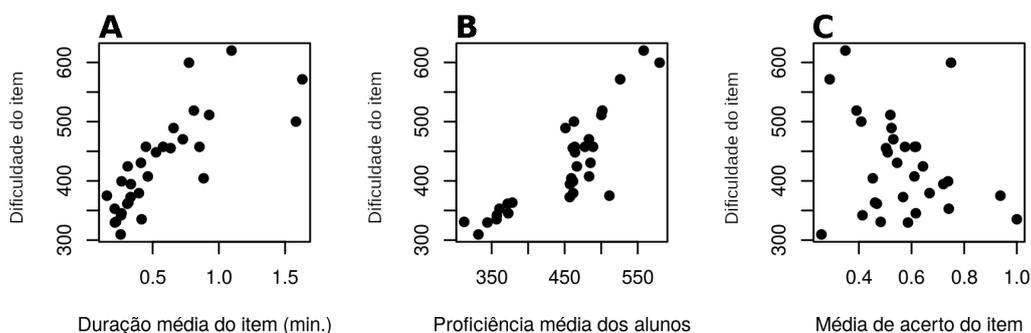
FIGURA 6 - Número de aplicações do TAI da Provinha Brasil - Leitura a alunos de 2º ano do ensino fundamental, segundo o tamanho da prova



Fonte: Elaboração dos autores.

Como esperado, os itens mais difíceis foram apresentados a alunos com maior proficiência, como mostra a tendência da Figura 7B. Da mesma forma, os itens mais difíceis demoraram mais para ser respondidos pelos alunos (Figura 7A). Nota-se ainda que não houve relação linear entre dificuldade e média de acerto nos itens (Figura 7C), uma propriedade dos testes adaptativos marcadamente diferente dos testes convencionais, nos quais se observam alta correlação entre a média de acerto no item e sua dificuldade estimada via TRI. De modo geral, os três resultados confirmam o que seria esperado em um TAI.

FIGURA 7 - Relação entre a dificuldade (parâmetro b) dos itens e três aspectos de sua aplicação: A) o tempo médio para resolução do item; B) a proficiência média dos alunos que receberam o item; C) a média de acerto no item. Alunos de 1º e 2º ano do ensino fundamental



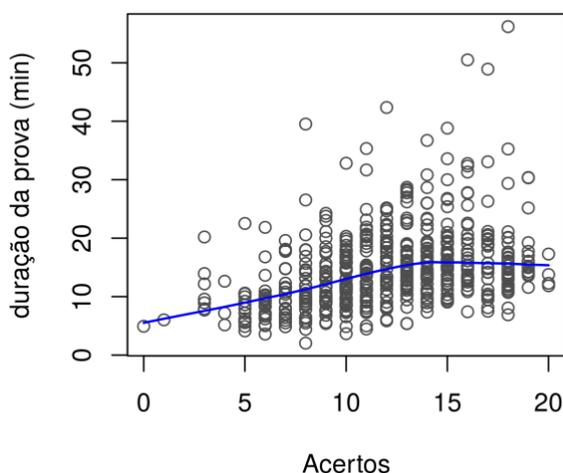
Fonte: Elaboração dos autores.

Gestão do tempo e proficiência

O aspecto temporal da execução do teste também foi investigado, comparando alunos de diferentes proficiências. A hipótese inicial é que os alunos mais proficientes tendem a gerir melhor o tempo de prova.

Em primeiro lugar, os resultados dos testes não adaptativos⁹ mostram uma associação positiva entre a proficiência e a duração da prova (Figura 8). Ou seja, há certa tendência de que alunos mais proficientes fiquem mais tempo fazendo a prova. A regressão linear confirma que essa relação é significativa ($p < 0,001$). Por outro lado, parece haver um limite para essa tendência, pois, a partir de 14 acertos, os alunos parecem não precisar de mais tempo para obter bons resultados. Observa-se inclusive uma fraca tendência de queda a partir desse ponto, mas a regressão linear indica que ela não é significativa ($p = 0,42$).

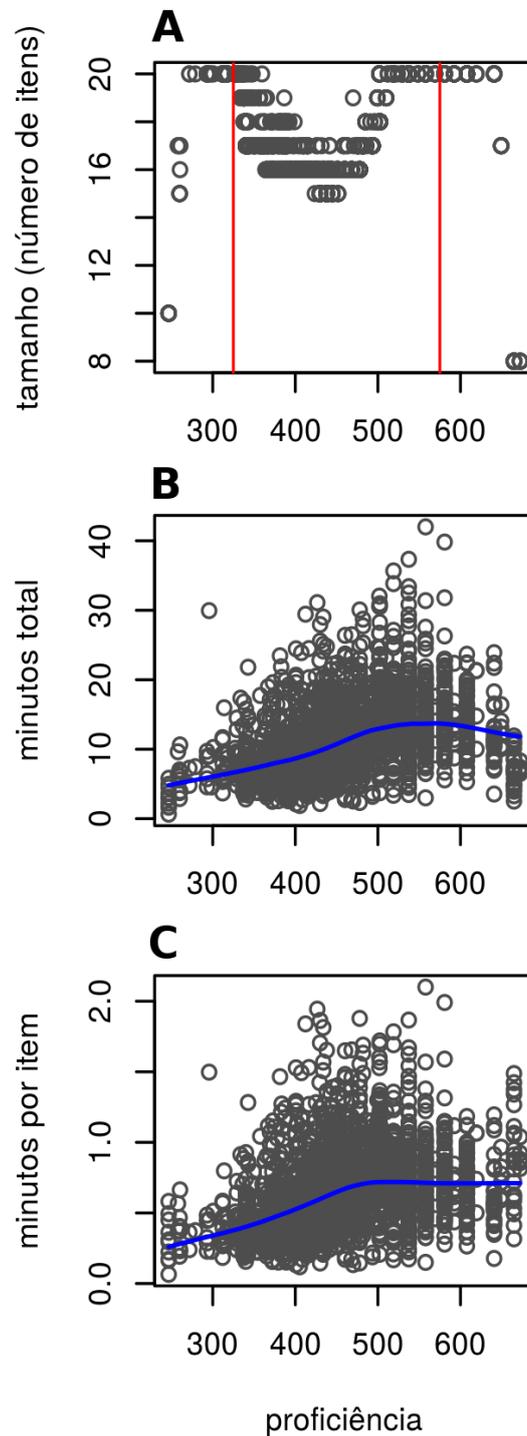
FIGURA 8 – Relação entre o número de acertos e a duração da Provinha Brasil - Leitura (versão eletrônica não adaptativa). Alunos de 2º ano do ensino fundamental



Fonte: Elaboração dos autores.

⁹ Os resultados e amostra dos testes eletrônicos não adaptativos - realizados nesse projeto como uma espécie de controle - estão detalhados em Alavarse *et al.* (2018).

FIGURA 9 - Relação entre proficiência e (A) tamanho, (B) duração total e (C) duração ponderada. Alunos de 1º e 2º ano do ensino fundamental



Fonte: Elaboração dos autores.

Nota: Linhas vermelhas: pontos de corte mínimo e máximo. Linhas azuis: tendência não linear calculada com estimador M.

Os resultados do TAI são semelhantes. É importante levar em conta que o TAI tem tamanhos variados (em número de itens), diferentemente do teste não adaptativo, sempre com 20 itens. Assim sendo, a análise da relação entre proficiência

e duração da prova é um pouco mais complexa no caso do TAI. A Figura 9 resume as principais informações necessárias nesse sentido.

Nota-se que os menores testes (com oito ou dez itens) correspondem a alunos nos extremos da escala apresentada na Figura 9A. Nesses casos, a prova finalizou pelo critério de *confiabilidade do nível de proficiência*. Mas a maior parte das finalizações geradas pelo algoritmo adaptativo se baseou no critério de *confiabilidade da proficiência*, que acabou surtindo efeito na faixa mais mediana da escala de proficiência – em que o erro de estimativa costuma ser menor. Essa distribuição na escala de proficiência confirma o caráter complementar dos dois critérios de finalização de teste em um TAI.

Na Figura 9B, nota-se uma tendência de crescimento até certo ponto. A partir de 600 pontos, há um declínio suave. Contudo, essa figura representa a duração total do teste, mas o TAI tem diferentes tamanhos. É possível que os alunos com mais de 600 pontos tenham ficado pouco tempo na prova, pois o algoritmo finalizou o teste rapidamente. De fato, quando observamos a Figura 9C, essa hipótese se confirma, já que o declínio desaparece quando analisamos o tempo por item, em vez do tempo total de prova.

Nota-se ainda que a tendência observada na Figura 9C é semelhante à observada na Figura 8. Com efeito, essa relação entre proficiência e tempo investido na prova (uma associação positiva até certo ponto, a partir do qual se estabiliza) se mostra consistente nos dois tipos de teste, adaptativo e não adaptativo. Um detalhe importante é que o tempo de prova deixa de aumentar quando a proficiência é um pouco maior do que a média observada: 463 pontos na Figura 9C e 13 acertos na Figura 8. A interpretação desses fatos não é clara, mas sua repetição nos dois casos indica um terreno fértil para investigação.

Foi analisado também um segundo aspecto da relação entre gestão do tempo e proficiência. Para tanto, selecionaram-se dois estratos da amostra: o terço inferior e o terço superior dos alunos em termos de proficiência. A Tabela 3 compara os dois estratos em algumas características dos itens aplicados.

TABELA 3 – Média e coeficiente de variação da duração de aplicação dos itens, segundo os estratos definidos pelo terço inferior e superior dos alunos de 1º e 2º ano do ensino fundamental em termos da proficiência

ESTRATOS	DURAÇÃO MÉDIA (MINUTOS)	COEFICIENTE DE VARIAÇÃO DA DURAÇÃO	NÚMERO TOTAL DE ITENS RESPONDIDOS
Terço inferior	0,45	0,91	29
Terço superior	0,68	0,74	22
P valor (Teste t pareado)	0,01	0,01	---

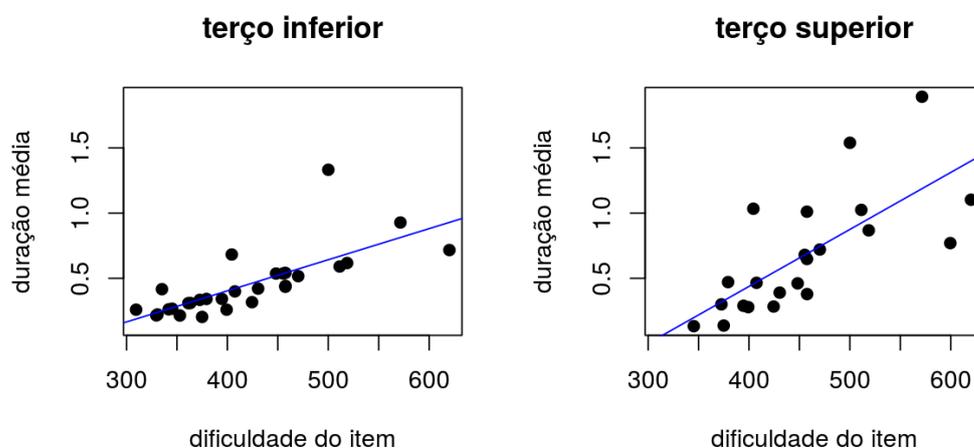
Fonte: Elaboração dos autores.

Nota: A significância da diferença entre os estratos é apontada pelo P valor do Teste t pareado.

Para verificar se as diferenças nas médias dos dois estratos descritas na Tabela 3 são significativas, foi realizado o Teste t pareado com 95% de confiança. Nota-se que os alunos do estrato superior tendem a ficar mais tempo resolvendo os itens, o que confirma os resultados anteriores. Além disso, o coeficiente de variação indica que há menor variabilidade no estrato superior; ou seja, esses alunos tendem a gerir o tempo de forma mais homogênea em cada item, sugerindo certa racionalidade quanto ao uso do tempo, e talvez também uma capacidade de concentração mais constante.

Por fim, há um ponto a ser verificado, partindo da relação observada entre a duração média e a dificuldade dos itens (Figura 7A): os alunos tendem a demorar mais nos itens mais difíceis, o que faz sentido e pode ser considerado parte da capacidade de gerir o tempo de prova. Mas em que medida isso se diferencia nos dois estratos? A Figura 10 mostra que, no estrato superior, a relação entre duração e dificuldade do item é mais intensa do que no estrato inferior. A regressão linear confirma essa observação.

FIGURA 10 – Relação entre duração (em minutos) e dificuldade (na escala Provinha Brasil - Leitura) dos itens em cada estrato. Alunos de 1º e 2º ano do ensino fundamental



Fonte: Elaboração dos autores.

Nota: Linha azul estimada por regressão linear ($p < 0,001$).

Em suma, nossos resultados sugerem que os alunos mais proficientes tendem a gerir melhor o tempo nos seguintes aspectos: a) investem mais tempo na prova, de um modo geral; b) investem mais tempo nos itens mais difíceis, e menos tempo nos mais fáceis; c) investem o tempo de forma mais homogênea em cada item – o que poderia estar relacionado, já no campo da especulação, a uma capacidade de concentração mais constante.

CONCLUSÃO

Os resultados obtidos confirmam a qualidade do TAI da Provinha Brasil – Leitura e dos métodos utilizados para desenvolver o algoritmo, partindo do arcabouço conceitual da TRI, com base em simulações e *software* livre. Com efeito, este texto pode servir como referência para a construção de outros testes adaptativos em moldes semelhantes. Foram observadas diversas características esperadas em um TAI, como a redução no tamanho médio das provas e no tempo médio de resolução, o uso diferencial dos itens e a ausência de correlação entre dificuldade e média de acerto dos itens. O TAI da Provinha Brasil – Leitura se mostrou também um instrumento útil para ser aplicado a diferentes populações de alunos, podendo provavelmente incluir também o 3º ano, especialmente com a ampliação do banco com itens que pudessem atender a alunos com maior proficiência, como é de se esperar nesse ano; embora possa haver aqueles que tenham desempenho parecido com seus colegas de anos anteriores. Para os alunos do 1º ano, pelo fato de a aplicação ter ocorrido no final do ano letivo, e como realizariam o Teste 1 no início do próximo ano letivo, o banco de itens foi adequado, além do objeto de avaliação do TAI – alfabetização e letramento inicial – ser relativo a uma competência que é trabalhada desde o 1º ano e que é cumulativa. Esse caráter mais cumulativo e menos restrito a cada série ficou evidenciado no espectro das proficiências médias observadas no 1º e 2º anos, confirmando, em termos psicométricos, a qualidade da própria Provinha Brasil – Leitura, ainda que desenvolvida pelo Inep para uso prioritário dos professores do 2º ano do ensino fundamental. Vale, também, destacar que nenhum aluno do 1º ano manifestou “surpresa” com o conteúdo dos itens respondidos, considerando que todas as aplicações foram acompanhadas e registradas em relatórios.

O critério de finalização de teste proposto neste trabalho (confiabilidade do nível de proficiência) também se mostrou eficaz e potencialmente útil para testes com escala discretizada (em níveis), que priorizem a avaliação (não apenas a mensuração) e a interpretação pedagógica, seja com finalidade formativa ou somativa. Cabe notar que o grau de influência desse critério depende diretamente do número de pontos de corte e sua distribuição na escala de proficiência.

Além disso, a análise da duração do teste confirma a hipótese de que os alunos mais proficientes tendem a gerir melhor o tempo de prova. Em primeiro lugar, há uma associação positiva entre proficiência e tempo de prova, que se estabiliza na porção superior da escala de proficiência. Tal associação foi observada tanto nos testes adaptativos quanto não adaptativos. Outra tendência geral observada é de os itens mais difíceis demorarem mais para serem respondidos. Comparando os terços superior e inferior dos alunos na escala de proficiência, observou-se que: a) o terço superior demora mais tempo respondendo aos itens do teste; b) o terço superior apresenta menor variação no tempo dedicado a cada item, o que poderia estar

relacionado – especulamos – a certa constância na capacidade de concentração; c) o terço superior apresenta uma inclinação maior na relação entre dificuldade do item e duração de aplicação, ou seja, esses alunos tendem a dedicar mais tempo aos itens mais difíceis e menos tempo aos itens mais fáceis. Esse aspecto da gestão do tempo de prova foi encontrado em toda a população, mas está especialmente presente nos alunos mais proficientes.

É importante explicitar algumas limitações deste trabalho. Em primeiro lugar, não foi incluído o aspecto do engajamento pessoal do examinando nas simulações, apenas sua proficiência e os parâmetros dos itens do banco. No entanto, é esperado que haja um engajamento maior dos alunos em um teste adaptativo, especialmente dos alunos nos extremos superior e inferior da escala de proficiência. Em segundo lugar, o método de estimação utilizado (EAP) é bayesiano, produzindo resultados menos precisos quando a população não apresenta média próxima da esperada *a priori*. Nesse sentido, recomenda-se um método não bayesiano, como a verossimilhança ponderada (WL), caso não haja informação confiável sobre a população. Ou, ainda, pode-se empregar um método misto, iniciando o teste com WL e terminando com EAP. Em terceiro lugar, o banco de itens utilizado foi bastante reduzido, e, como vimos, a taxa de exposição de alguns itens se mostrou bastante alta, enquanto outros sequer foram aplicados. Para superar tal limitação, recomenda-se outro método de seleção de itens, pois a Máxima Informação de Fisher tende a gerar esse tipo de resultado, especialmente se associada a um modelo logístico com parâmetro α (inclinação) constante, como foi o caso. Outra limitação deste estudo, particularmente importante, é o tamanho reduzido do banco de itens, sendo aconselhável um banco com pelo menos 100 itens para um desempenho eficaz dos testes adaptativos.

Por fim, cabe mencionar possíveis aprimoramentos futuros, como a inclusão de métodos para otimizar a segurança do banco de itens (controlando a taxa de exposição, por exemplo), a transformação do TAI em um Teste de Múltiplos Estágios (TME), em que os itens são selecionados em blocos, ou mesmo a possibilidade de se pré-testar novos itens durante a aplicação do teste.

REFERÊNCIAS

ALAVARSE, O.; CATALANI, E.; MENEGHETTI, D.; TRAVITZKI, R. Teste Adaptativo Informatizado como recurso tecnológico para alfabetização inicial. *Revista Iberoamericana de Sistemas, Cibernética e Informática*, v. 15, n. 3, p. 68-78, 2018.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística, 2000.

BABCOCK, B.; WEISS, D. J. Termination criteria in Computerized Adaptive Tests: do variable-length CATs provide efficient and effective measurement? *Journal of Computerized Adaptive Testing*, v. 1, n. 1, p. 1-18, Dec. 2012.

BAKER, F. B. *The basics of Item Response Theory*. 2nd ed. Washington: ERIC Clearinghouse on Assessment and Evaluation, 2001.

BARRADA, J. R. A method for the comparison of Item Selection Rules in Computerized Adaptive Testing. *Applied Psychological Measurement*, v. 34, n. 6, p. 438-452, 2010.

BARRADA, J. R. Tests adaptativos informatizados: uma perspectiva general. *Anales de Psicología*, v. 28, n. 1, p. 289-302, 2012.

BIRNBAUM, A. Statistical theory for logistic mental test models with a prior distribution of ability. *Journal of Mathematical Psychology*, v. 6, p. 258-276, 1969.

BOCK, R. D.; MISLEVY, R. J. Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement*, v. 6, n. 4, p. 431-444, 1982.

BRASIL. Ministério da Educação. Portaria Normativa n. 10, de 24 de abril de 2007. *Diário Oficial da União*, Brasília, 26 abr. 2007a. Disponível em: <http://portal.mec.gov.br/arquivos/pdf/provinha.pdf>. Acesso em: 20 ago. 2020.

BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. 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. *Diário Oficial da União*, Brasília, p. 5, 25 abr. 2007b. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6094.htm. Acesso em: 20 ago. 2020.

BRASIL. Ministério da Educação. *Provinha Brasil: avaliando a alfabetização: guia de apresentação, correção e interpretação dos resultados: Teste 2*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2016.

CATALANI, Érica Maria Toledo. *Teste Adaptativo Informatizado da Provinha Brasil: a construção de um instrumento de apoio para professores(as) e gestores(as) de escolas*. 201. 282 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2019.

FERREIRA, D. F. *Estatística básica*. Lavras, MG: Ed. UFLA, 2005.

GEORGIADOU, E. G.; TRIANTAFILLOU, E.; ECONOMIDES, A. A. A Review of Item Exposure Control Strategies for Computerized Adaptive Testing developed from 1983 to 2005. *The Journal of Technology, Learning, and Assessment*, v. 5, n. 8, p. 1-38, 2007.

GONTIJO, C. M. M. Avaliação da alfabetização: Provinha Brasil. *Educação e Pesquisa*, v. 38, n. 3, p. 603-622, 2012.

LINDEN, W. J.; GLAS, C. A. W. *Elements of Adaptive Testing*. New York: Springer, 2010.

LORD, F. *Applications of Item Response Theory to practical testing problems*. Hillsdale: Lawrence Erlbaum Associates, 1980.

MAGIS, D.; RAÏCHE, G. Random generation of response patterns under Computerized Adaptive Testing with the R Package catR. *Journal of Statistical Software*, v. 48, n. 8, p. 1-31, 2012.

MORAIS, A. G.; LEAL, T. F.; ALBUQUERQUE, E. B. C. “Provinha Brasil”: monitoramento da aprendizagem e formulação de políticas educacionais. *Revista Brasileira de Política e Administração da Educação*, v. 25, n. 3, p. 301-320, maio/ago. 2009.

PARTCHEV, I. *irtoys: A Collection of Functions Related to Item Response Theory (IRT)*. S.l.: The Comprehensive R Archive Network, 2016. Disponível em: <https://cran.r-project.org/package=irtoys>. Acesso em: 19 mar. 2020.

REIF, M. *PP: Estimation of person parameters for the 1,2,3,4-PL model and the GPCM*. S.l.: GitHub, 6 abr. 2019. Disponível em: <https://github.com/manuelreif/PP>. Acesso em: 19 mar. 2020.

REVUELTA J.; PONSODA, V. A comparison of item exposure control methods in computerized adaptive testing. *Journal of Educational Measurement*, v. 35, n. 4, p. 311-327, 1998.

SOARES, M. *Alfabetização: a questão dos métodos*. São Paulo: Contexto, 2016.

WARM, T. A. Weighted likelihood estimation of ability in item response theory. *Psychometrika*, v. 54, n. 3, p. 427-450, Sept. 1989.

WEISS, D. J.; KINGSBURY, G. G. Application of Computerized Adaptive Testing to educational problems. *Journal of Educational Measurement*, v. 21, n. 4, p. 361-375, Winter 1984.

NOTA: Este artigo foi produzido colaborativamente pelos autores. Rodrigo Travitzki foi responsável pela criação do algoritmo, Ocimar Munhoz Alavarse executou a coordenação geral do projeto, Douglas De Rizzo Meneghetti criou o *software* de aplicação da prova eletrônica e Érica Maria de Toledo Catalani realizou a análise psicométrica dos resultados.

Recebido em: 19 MARÇO 2020

Aprovado para publicação em: 16 SETEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7203>

AS PROVAS DO EXAME NACIONAL DO ENSINO MÉDIO SÃO UNIDIMENSIONAIS?

 LEANDRO ARAUJO DE SOUSA^I

 JOSÉ AIRTON DE FREITAS PONTES JUNIOR^{II}

 ADRIANA EUFRÁSIO BRAGA^{III}

^I Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Canindé-CE, Brasil; leandro.sousa@ifce.edu.br

^{II} Universidade Estadual do Ceará (Uece), Fortaleza-CE, Brasil; japontesjr@gmail.com

^{III} Universidade Federal do Ceará (UFC), Fortaleza-CE, Brasil; adrianaufc@yahoo.com.br

RESUMO

O modelo de Teoria da Resposta ao Item utilizado em muitos testes educacionais no Brasil, como o Exame Nacional do Ensino Médio, exige que os itens sejam unidimensionais. Assim, esta pesquisa teve o objetivo de analisar se os itens desse exame apresentam essa suposição. Para tanto, com base em uma amostra aleatória de participantes que realizaram a prova em 2017, foi investigada a dimensionalidade das provas do exame por meio do teste de Análise Paralela e Análise Fatorial de Informação Plena. Os resultados encontrados indicam que alguns itens não são unidimensionais.

PALAVRAS-CHAVE AVALIAÇÃO DA EDUCAÇÃO • MÉTODOS DE AVALIAÇÃO • ANÁLISE FATORIAL • TEORIA DA RESPOSTA AO ITEM.

¿LAS PRUEBAS DEL EXAMEN NACIONAL DE LA SECUNDARIA SUPERIOR SON UNIDIMENSIONALES?

RESUMEN

El modelo de Teoría de la Respuesta al Ítem utilizado en muchas pruebas educativas en Brasil, como el Exame Nacional do Ensino Médio [Examen Nacional de la Secundaria Superior], exige que los ítems sean unidimensionales. De este modo, este estudio tuvo el propósito de analizar si los ítems de este examen presentan dicha suposición. Para ello, en base a una muestra aleatoria de participantes que realizaron la prueba en 2017, se investigó la dimensionalidad de las pruebas del examen por medio del test de Análisis Paralelo y Análisis Factorial de Información Plena. Los resultados encontrados indican que algunos de los ítems no son unidimensionales.

PALABRAS CLAVE EVALUACIÓN DE LA EDUCACIÓN • MÉTODOS DE EVALUACIÓN • ANÁLISIS FACTORIAL • TEORÍA DE LA RESPUESTA AL ÍTEM.

ARE THE TESTS OF THE NATIONAL EXAM OF UPPER SECONDARY EDUCATION UNIDIMENSIONAL?

ABSTRACT

The Item Response Theory model used in many educational tests in Brazil, such as the Exame Nacional do Ensino Médio [National Exam of Upper Secondary Education], requires that items be unidimensional. This research therefore aimed to analyze whether the items of this exam contain such an assumption. Based on a random sample of participants who took the exam in 2017, the dimensionality of its tests was investigated using the Parallel Analysis test and the Full Information Factor Analysis. The results show that some items are not unidimensional.

KEYWORDS EDUCATION ASSESSMENT • ASSESSMENT METHODS • FACTOR ANALYSIS • ITEM RESPONSE THEORY.

INTRODUÇÃO

O Exame Nacional do Ensino Médio (Enem) é uma prova aplicada em larga escala, tem caráter multidisciplinar e os seus resultados têm sido utilizados por muitas instituições de ensino superior (IES) brasileiras como principal critério de seleção aos cursos de graduação. Assim, é necessário que as provas sejam elaboradas de forma que forneçam resultados confiáveis e possibilitem a seleção justa dos candidatos. Nesse sentido, a psicometria auxilia na elaboração de testes que apresentem boa capacidade de realizar medidas. Entre as características necessárias de um teste educacional estão a validade e a fidedignidade, ou seja, deve apresentar boas evidências de que realiza o que se pretende e de forma precisa (TOFFOLI *et al.*, 2016).

Por muito tempo, a Teoria Clássica dos Testes (TCT) foi e continua sendo utilizada na análise da qualidade métrica de instrumentos de medida nas avaliações educacionais (SARTES; SOUSA-FORMIGONI, 2013; SOUSA; BRAGA, 2020). No entanto, nos últimos anos, tem ganhado destaque a Teoria da Resposta ao Item (TRI) em avaliações em larga escala, sob a justificativa de oferecer vantagens como estabilidade e comparabilidade dos resultados, algo não oferecido pela TCT (ANDRADE; LAROS; GOUVEIA, 2010). Ainda assim, considera-se que o modelo TRI não tem substituído totalmente a TCT, mas complementa as suas análises (SARTES; SOUSA-FORMIGONI, 2013).

A TCT, para alguns autores (ANDRADE; TAVARES; VALLE, 2000; KLEIN, 2013; SARTES; SOUSA-FORMIGONI, 2013), apresenta problemas, como a dependência da amostra, ou seja, do particular conjunto de sujeitos avaliados, dessa forma o teste apresenta escores diferentes para grupos diferentes de avaliados; a dependência do teste e dos itens, pois escores distintos são obtidos se um grupo de sujeitos é avaliado com diferentes testes sobre o mesmo conhecimento; e, em decorrência disso, os testes não permitem a comparabilidade dos resultados, sendo, dessa forma, instáveis.

Para propor alternativas a esses problemas, surge a TRI, que alega proporcionar estabilidade dos resultados, ou seja, os sujeitos terão os mesmos escores, ou notas, mesmo que sejam utilizados testes com itens diferentes. Isso se torna possível porque o parâmetro de análise é o item, em que, independentemente dos avaliados, terá sempre os mesmos parâmetros. Por apresentar essa invariabilidade, os resultados tornam-se comparáveis (PASQUALI, 2009; VALLE, 2000).

Diante disso, a partir de 2009, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) passou a utilizar a TRI na validação dos itens e na análise dos resultados do Enem. O exame utiliza o modelo logístico de três parâmetros, em que considera dificuldade, discriminação e probabilidade de acerto casual do item (BRASIL, 2011). Assim, acreditamos que a mudança de método de análise pode influenciar na confiabilidade da medida realizada, pois o modelo

mais convencional da TRI exige a unidimensionalidade dos itens, algo muito difícil de ser obtido. Algumas críticas questionam a coerência metodológica do exame (TAVARES, 2013). Entre as discussões, Tavares (2013) indaga a possibilidade de cumprir o pressuposto, já que o conhecimento humano é multideterminado, ou seja, depende de vários fatores. Questionou-se também a validade das medidas realizadas, já que um modelo unidimensional não consegue explicar uma realidade multidimensional, embora se advogue por uma “unidimensionalidade essencial” (STOUT, 1990).

Como possibilidade de um modelo de melhor ajuste a essa realidade surgem os modelos de TRI multidimensional (RECKASE, 2009). No entanto, esses modelos ainda são pouco implementados nas avaliações educacionais em larga escala.

Com base nisso, considerando que o Enem é uma prova multidisciplinar, questionamos a sua validade como um instrumento unidimensional. Embora o exame apresente áreas de avaliação bem determinadas, como Linguagens e Códigos (LC), Matemática (MT), Ciências da Natureza (CN) e Ciências Humanas (CH), cada área é composta por conhecimentos de diferentes disciplinas. Por exemplo, na prova de LC há itens referentes aos conhecimentos de Língua Portuguesa, Língua Estrangeira, Artes e Educação Física. Embora possamos compreender que todas essas áreas constituem uma linguagem, cada uma resguarda a sua especificidade que a diferencia das outras.

Em estudos anteriores, evidências desse problema têm sido apontadas. Ao analisar o Enem 2012 como prova única de 180 itens por meio da TRI multidimensional, foram identificadas quatro dimensões, no entanto, os itens não foram relacionados com as áreas de conhecimento que constituem a prova, gerando dificuldade de interpretação semântica do exame (VIEIRA, 2016). Nesse sentido, a divisão dos itens nas quatro áreas de conhecimento do exame não apresenta estrutura fatorial consistente com os conteúdos dos itens. Em outro estudo, a mesma análise foi realizada com o certame de 2016, em que foram identificados dois fatores (PICCIRILLI; SOUZA, 2018). Na interpretação dos autores, sugere-se que essas duas dimensões estão relacionadas a interpretação de texto e raciocínio lógico.

Modelos com duas dimensões para as provas de CH e MT, com três dimensões para a prova de CN com uma estrutura diferente da originalmente proposta, foram mais bem ajustados em análises fatoriais exploratórias e confirmatórias para o exame realizado em 2010 (MUNER, 2013). Nesse mesmo estudo, por meio de uma análise semântica por juízes independentes, identificou-se a presença das habilidades de inteligência cristalizada, conhecimento quantitativo, inteligência fluida e leitura e escrita, em que a sua estrutura interna apresentou validade de construto, mas de forma convergente para uma reorganização dos itens de acordo com o modelo hierárquico da inteligência de Catell-Horn-Carroll (CHC).

Presença de um segundo fator também foi identificada na prova de MT do exame realizado em 2015, utilizando análise de componentes principais dos resíduos a partir do ajuste do modelo de Rasch, em que considera apenas o parâmetro de dificuldade dos itens (PRIMI; CICCHETTO, 2018). Nesse estudo foi observado que itens difíceis dessa prova compunham uma dimensão secundária e que uma quantidade significativa de 183.450 participantes com nota geral baixa apresentava bom desempenho nesses itens. Como o modelo de TRI unidimensional de três parâmetros, que leva em conta o parâmetro de acerto casual (ou “chute”) utilizado pelo exame, exige coerência nas respostas por penalizar os candidatos que acertam itens difíceis quando erram itens fáceis, esses participantes são prejudicados em sua pontuação final quando a prova não cumpre o pressuposto da unidimensionalidade.

No entanto, em outro estudo (TRAVITZKI, 2017), ao utilizar a análise paralela para a verificação da dimensionalidade do Enem, técnica também utilizada nesta pesquisa, o autor considera que a estrutura unidimensional para as provas do Enem de 2009 e 2011 é suficiente para explicar a variabilidade das respostas aos itens. Por outro lado, o autor também identifica a existência de um segundo fator significativo, mas que considera de menor importância. Ao analisar a qualidade psicométrica da prova de MT de 2018, verificou a presença do pressuposto da unidimensionalidade a partir da análise paralela, mas pelo menos quatro itens apresentaram problemas de ajuste e foram excluídos para prosseguimento da análise (SOARES; SOARES; SANTOS, 2020). Em estudo (SOUSA; PONTES JUNIOR; BRAGA, 2020) que analisa a estrutura fatorial da prova de LC dos anos de 2009 a 2014 com base na análise de componentes principais, é identificada a presença de uma dimensão dominante nas provas, exceto para a prova do ano de 2014. No entanto, esse estudo indica que o percentual de variância explicada das provas é muito baixo, sendo a maior para a prova do ano de 2011, com 14,71% de variância explicada pelo primeiro fator.

Problemas na estrutura fatorial dos itens do Enem podem ser identificados ainda em fase de análises preliminares do exame por meio de alguns parâmetros psicométricos como fidedignidade, dificuldade e discriminação dos itens. Travitzki (2017) analisa a fidedignidade das provas do Enem como inadequado para a prova de MT e aceitável, mas próximo do limiar inferior, para as provas de CN. Problemas de fidedignidade também são encontrados em estudo que analisa o Enem 2011 por meio de um modelo bifatorial (GOMES; GOLINO; PERES, 2018). Nesse estudo, os autores, apoiados em Hogan (2013 *apud* GOMES; GOLINO; PERES, 2018), consideram que em exames de alto impacto (*high stake*) os índices de fidedignidade deveriam ser superiores a 0,95. Com base nisso, índices de fidedignidade inadequados foram encontrados para as quatro provas do exame. No entanto, em

um modelo bifatorial, o fator geral, denominado no estudo de desempenho escolar geral (DEG), apresenta fidedignidade aceitável. Diante desses dados, os autores consideram que “a operacionalização da matriz de referência por meio dos itens que compõem as provas do Enem parece não ser capaz de avaliar especificamente cada domínio isoladamente. No entanto, os quatro domínios contribuem para a formação do DEG” (p. 341). Ressaltam ainda que isso indica que o Enem consiste em um modelo multidimensional.

Além de problemas de fidedignidade identificados em provas do Enem, alguns estudos indicam problemas nos itens do exame. Travitzki (2017), com base em critérios como correlação, correlação bisserial, discriminação e o índice de ajuste, identifica 33% dos itens de CN na prova de 2009, e 29% na prova de 2011 foram considerados ruins ou duvidosos. Para a prova de MT, chegou a 49% em 2009 e 18% com essa classificação. Da mesma forma, estudo aponta qualidade desejável da prova de MT do Enem 2015 quando se considera parâmetros de TCT e TRI (TOFFOLI, 2019). Nesse estudo, a prova possuía muitos itens com índices de discriminação muito baixos ($< 0,3$), ou seja, não apresentam a capacidade de diferenciar candidatos de baixa e alta habilidade, e não apresentavam itens considerados fáceis, quando analisados pela TCT. Quando analisados pela TRI, muitos itens apresentaram baixa discriminação ou discriminação negativa, e itens com parâmetros de dificuldade muito altos, fora do intervalo entre -3 e $+3$, considerados adequados. Problemas dessa natureza podem indicar má elaboração dos itens e podem influenciar na sua correlação com a área de conhecimento e conseqüentemente em sua dimensionalidade.

Diante disso, as evidências dessas pesquisas indicam que os itens do exame não se ajustam bem a um modelo unidimensional. Dessa forma, esta pesquisa está pautada no seguinte problema e questões decorrentes: os itens das provas do Enem 2017 são unidimensionais? Diante disso, este estudo teve o objetivo de analisar a dimensionalidade das provas das quatro áreas do Enem (LC, MT, CN e CH) por meio da análise paralela e da análise fatorial de informação plena.

MÉTODO

População e amostra

A população-alvo deste estudo é constituída de 6.731.341 candidatos de todas as regiões e estados do Brasil que realizaram o Enem 2017. Foram excluídos os candidatos que não estiveram presentes no exame e os que não responderam a nenhum dos itens do exame, permanecendo 4.426.755 candidatos.

A amostra final desta pesquisa é formada por 10.000 participantes do Enem de 2017 que estiveram presentes em todas as provas do exame. Esses participantes

foram selecionados por amostragem aleatória simples. Esse número reduzido da amostra em relação ao quantitativo total dos participantes se deu pela capacidade limitada de processamento dos computadores utilizados nesta pesquisa. O volume muito grande de dados não foi suportado pela capacidade operacional, gerando instabilidade no processamento das máquinas. Considerando o tamanho populacional, o erro estimado foi menor que 1% para um intervalo de confiança de 95%.

Caracterização do exame

O Enem é um exame que busca avaliar competências e habilidades desenvolvidas pelos alunos no decorrer da educação básica, sendo orientado por uma matriz de referência especificamente construída.

A matriz de referência atualmente possui cinco eixos cognitivos comuns a todas as áreas, a saber: i) dominar linguagens; ii) compreender fenômenos; iii) enfrentar situações-problemas; iv) construir argumentação; e v) elaborar propostas. Suas competências e habilidades são divididas em quatro grandes áreas do conhecimento: Linguagens e Códigos, que contempla os conhecimentos de Português, Educação Física, Artes, Língua Estrangeira Moderna e Tecnologia da Informação e Comunicação; Ciências da Natureza, que abrange a Biologia, Química e Física; Ciências Humanas, envolvendo a História, Geografia, Sociologia e Filosofia; e Matemática.

O exame é constituído de 180 itens de múltipla escolha com cinco alternativas, sendo 45 itens para cada grande área e uma redação.

Destacamos que utilizamos apenas um dos quatro cadernos de prova de cada área. Para LC e CN, utilizou-se o caderno azul, para CH, o caderno amarelo, e para MT, o caderno cinza. A seleção da amostra foi realizada para cada caderno de prova separadamente, ou seja, a amostra de candidatos de um caderno não é a mesma de outra. Para o caderno de LC, o candidato opta por Língua Inglesa ou Espanhola. Este estudo foi realizado com os candidatos que escolheram a segunda opção de Língua Estrangeira. Todas as escolhas foram realizadas por seleção aleatória simples, ou seja, por sorteio.

Os microdados dos resultados das provas e as respostas de cada candidato, nos itens do Enem 2017, estão disponíveis no *site* do Inep e são de livre acesso ao público.

Tratamento para dados ausentes

Entre os participantes da amostra deste estudo, ocorreu ausência de respostas a determinados itens das provas. No entanto, em nenhuma delas a quantidade de valores ausentes ultrapassou 1%. Para não eliminar os participantes das análises posteriores, optou-se por substituir os valores ausentes pelo método de imputação múltipla.

O método de imputação múltipla consiste na substituição dos valores ausentes por meio de várias simulações até chegar a uma aproximação estatística ótima. Esse método é bastante robusto quando os valores ausentes são completamente aleatórios (NEWMAN, 2014) e em casos de dados binários ou dicotômicos (BÉLAND *et al.*, 2018). Para a aplicação desse método, utilizamos o *software* SPSS 20.0.

Análise estatística

Para todas as análises realizadas neste trabalho, foi utilizado o *software* R, programa livre e amplamente usado pelos pesquisadores para implementação de análises estatísticas. Esse *software* incorpora uma ampla possibilidade e flexibilidade nas análises.

Inicialmente foram obtidas a discriminação dos itens com base na TCT. Foram estimados os parâmetros de discriminação por meio da correlação ponto bisserial, uma vez que os itens foram dicotomizados em certo e errado (PASQUALI, 2009). Para essas análises, foi utilizado o pacote “ltm” (RIZOPOULOS, 2006). Itens com discriminação muito baixa ($r_{pb} < 0,15$) foram excluídos da análise, pois indicam baixa correlação do item com o escore total. Inicialmente a discriminação foi utilizada com finalidade de analisar sua adequação para a realização da análise fatorial exploratória.

Para uma análise exploratória da dimensionalidade, submeteram-se os dados a um teste de Análise Paralela. Para aplicação desse teste, utilizou-se o pacote “psych” (REVELLE, 2017). Essa análise consiste na comparação dos *eigen value* dos dados reais com os de um conjunto de dados simulados gerados aleatoriamente com igual número de variáveis e de mesmo tamanho amostral (HAYTON; ALLEN; SCARPELLO, 2004). O critério para a determinação do número de fatores a serem retidos se baseia na comparação dos *eigen value* dos dados reais e dos dados gerados. Retêm-se os fatores no momento em que o valor *eigen* dos dados reais é menor que o dos dados simulados. Para a análise da existência de uma dimensão dominante nos dados, realizou-se uma comparação do primeiro *eigen value* com o segundo (Eigen1/Eigen2). O *scree plot* também foi analisado na definição do número de fatores.

Outra análise utilizada para verificar o pressuposto da unidimensionalidade foi a Análise Fatorial de Informação Plena (FIFA – *Full Information Factor Analysis*), um modelo baseado na TRI, considerada mais adequada em situações de itens dicotômicos (BARTHOLOMEW, 1980) típicos de testes educacionais, como é o caso deste trabalho, que utiliza os dados da prova do Enem. Para essa análise, foi utilizado o pacote estatístico “mirt” (CHALMERS, 2012).

O ajuste dos modelos foi realizado com base nas medidas do *Root-Mean-Square Error of Approximation* (RMSEA), *Standardised Root Mean Square Residual* (SRMSR),

Tucker-Lewis Index (TLI) e *Comparative Fit Index* (CFI). São considerados desejáveis para um bom ajuste dos modelos valores de RMSEA e SRMSR abaixo de 0,05. Já para os valores de TLI e CFI são considerados como indicativos de bom ajuste os valores acima de 0,95. Para essas análises, foi utilizado o pacote estatístico “mirt” (CHALMERS, 2012).

Também foram analisados os valores dos índices de ajuste dos modelos com base no Critério de Informação de Akaike (AIC – *Akaike’s Information Criterion*) e o Critério de Informação Baysiano (BIC – *Bayesian Information Criterion*) (NYLUND; ASPAROUHOV; MUTHÉN, 2007). O modelo que produz menores valores de ambos os critérios é o de melhor ajuste. No entanto, o AIC tende a superestimar a quantidade de dimensões e o BIC, a subestimar. O índice considerado neste estudo foi o BIC, pois é avaliado como mais consistente que o índice AIC por meio de estudos de simulação Monte Carlo (NYLUND; ASPAROUHOV; MUTHÉN, 2007).

A análise do ajuste do modelo também foi realizada com base no Índice de Dimensionalidade (ID) considerando as recomendações de Nojosa (2002). Consiste, inicialmente, em obter os valores de ajuste do modelo com um fator (M1). Posteriormente estima-se os modelos com dois (M2), três fatores (M3) e assim sucessivamente, e então são comparados entre si.

Ao comparar os modelos, obtêm-se um valor com uma distribuição qui-quadrado (X^2). No entanto, esse valor é superestimado. Recomenda-se então dividir esse valor por dois ou por três para um ajuste mais adequado (NOJOSA, 2002). Esse valor será denominado nesta pesquisa de qui-quadrado corrigido (X^2_{corr}). Ele, então, é dividido pelos graus de liberdade (gl). Nojosa (2002) ressalta que esse valor não é interpretável diretamente, pois só a diferença do valor X^2 entre os modelos deve ser considerado. Esse valor obtido será denominado de ID. Na comparação de dois modelos, M1 x M2, um ID com valor positivo maior que 2,0 indica que o segundo modelo é melhor, se o valor for menor que 2,0, o primeiro modelo é preferível. Nesse sentido, espera-se que, se os itens forem unidimensionais, a comparação entre os modelos 1 (M1) e 2 (M2) forneça um ID positivo menor que 2,0.

RESULTADOS E DISCUSSÃO

Análise descritiva dos itens pela teoria clássica

Para compreender melhor o comportamento dos itens, foram estimados os parâmetros com base na teoria clássica. O objetivo foi analisar a discriminação dos itens por meio do coeficiente de correlação ponto bisserial de modo a identificar o quanto o item se correlaciona com o escore total do conjunto de itens. Os valores estão dispostos na Tabela 1.

TABELA 1 – Parâmetros clássicos dos itens do Enem 2017

ÁREA	LC		MT		CN		CH	
ITEM	r_{pb}	d	r_{pb}	d	r_{pb}	d	r_{pb}	d
1	0,34	0,31	0,31	0,29	0,29	0,25	0,48	0,30
2	0,30	0,46	0,33	0,64	0,22	0,25	0,47	0,55
3	0,36	0,45	0,22	0,27	0,23	0,55	0,16	0,23
4	0,15	0,29	0,27	0,17	0,35	0,49	0,37	0,53
5	0,30	0,33	0,33	0,43	0,15	0,16	0,40	0,49
6	0,36	0,65	0,09	0,11	0,29	0,16	0,39	0,59
7	0,34	0,37	0,31	0,12	0,40	0,46	0,47	0,35
8	0,24	0,41	0,22	0,22	0,16	0,19	0,12	0,23
9	0,36	0,50	0,44	0,40	0,23	0,60	0,51	0,35
10	0,27	0,26	0,25	0,24	0,13	0,12	0,26	0,49
11	0,45	0,61	0,00	0,12	0,14	0,17	0,39	0,44
12	0,26	0,20	0,31	0,30	0,19	0,32	0,22	0,23
13	0,20	0,19	0,21	0,25	0,30	0,42	0,50	0,58
14	0,44	0,37	0,30	0,25	0,18	0,16	0,21	0,36
15	0,23	0,50	0,29	0,28	0,21	0,35	0,46	0,41
16	0,27	0,33	0,29	0,31	0,18	0,19	0,29	0,54
17	0,36	0,51	0,34	0,30	0,40	0,25	0,41	0,31
18	0,26	0,23	0,14	0,16	0,21	0,28	0,28	0,31
19	0,50	0,58	0,13	0,19	0,43	0,37	0,19	0,36
20	0,23	0,13	0,29	0,27	0,33	0,37	0,39	0,36
21	0,31	0,66	0,11	0,13	0,06	0,47	0,19	0,23
22	0,33	0,37	0,22	0,29	0,38	0,23	0,37	0,43
23	0,39	0,57	0,35	0,43	0,13	0,21	-0,05	0,21
24	0,33	0,28	0,25	0,38	0,22	0,24	0,30	0,29
25	0,10	0,15	0,25	0,17	0,15	0,26	0,04	0,25
26	0,21	0,41	0,40	0,59	0,13	0,23	0,38	0,48
27	0,46	0,65	0,37	0,50	0,17	0,18	0,50	0,42
28	0,23	0,54	0,31	0,27	0,19	0,15	0,31	0,30
29	0,05	0,14	0,15	0,06	0,36	0,46	0,34	0,39
30	0,26	0,30	0,24	0,28	0,10	0,14	0,04	0,20
31	0,48	0,46	0,33	0,34	0,25	0,21	0,39	0,38
32	0,45	0,39	0,29	0,30	0,09	0,08	0,32	0,32
33	0,00	0,17	0,24	0,11	0,33	0,36	0,52	0,49
34	0,39	0,77	0,13	0,23	0,10	0,13	0,20	0,24
35	0,13	0,25	0,38	0,37	0,07	0,09	0,52	0,31
36	0,05	0,25	0,05	0,23	0,35	0,51	0,06	0,27
37	0,35	0,51	0,25	0,27	0,20	0,16	0,15	0,27
38	0,29	0,30	0,24	0,15	0,26	0,20	0,26	0,23
39	0,08	0,18	0,37	0,12	0,31	0,27	0,40	0,45
40	0,37	0,41	0,11	0,17	0,20	0,11	0,19	0,22
41	0,22	0,31	0,14	0,16	0,27	0,29	0,47	0,45
42	0,36	0,41	0,20	0,28	0,24	0,33	0,29	0,38
43	0,34	0,53	0,27	0,24	0,40	0,29	0,30	0,31

(Continua)

(Continuação)

ÁREA	LC		MT		CN		CH	
ITEM	r_{pb}	d	r_{pb}	d	r_{pb}	d	r_{pb}	d
44	0,35	0,35	0,12	0,23	0,12	0,23	0,20	0,14
45	0,26	0,19	0,15	0,23	0,31	0,25	0,15	0,11

Fonte: Elaboração dos autores (2020).

Legenda: r_{pb} : correlação ponto biserial; d : dificuldade do item.

Nota: Destaque em negrito para as correlações ponto-biserial abaixo de 0,15.

A partir disso, sete itens de LC (4, 25, 29, 33, 35, 36 e 39), dez itens de MT (6, 11, 18, 19, 21, 34, 36, 40, 41 e 44), dez itens de CN (10, 11, 21, 23, 26, 30, 32, 34, 34 e 44) e cinco itens da prova de CH (8, 23, 25, 30 e 36) apresentaram discriminação abaixo de um valor minimamente aceitável ($> 0,15$). Portanto, esses itens foram excluídos das análises seguintes. Esse procedimento é recomendado na análise inicial da qualidade dos itens (COSTA; FERRÃO, 2015). Esse valor de discriminação também foi utilizado como parâmetro para exclusão de itens da análise fatorial no estudo de Ferreira (2009).

Diante disso, observam-se problemas em muitos itens das provas do Enem de 2017. Muitos itens apresentam discriminação muito baixa. Isso indica que esses itens não têm boa correlação com o escore total da prova (MUÑIZ, 1994). Do ponto de vista prático, isso demonstra que o item não está conseguindo diferenciar candidatos de escore total mais alto daqueles com escore total mais baixo (MUÑIZ, 1994; VIANNA, 1976), ou seja, ambos tiveram probabilidade de acertar o item bem próximo (ANDRADE; LAROS; GOUVEIA, 2010). Dito de outra forma, o item não está discriminando os examinados eficazes e os ineficazes em um teste (ANDRIOLA, 1998; MUÑIZ, 1994). Itens com problema de discriminação devem ser rejeitados (VIANNA, 1976).

Esse tipo de problema é uma tarefa que pode ser resolvida pelos elaboradores de itens. Comumente, problemas de clareza e objetividade do item causam dificuldades em sua interpretação, o que pode induzir candidatos de bom desempenho ao erro. Em um estudo foi identificado que itens com formulação complexa e enunciado confuso apresentam baixa discriminação (CANÇADO; CASTRO; OLIVEIRA, 2013). Por outro lado, o mesmo estudo indicou que os itens simples, objetivos e claros apresentam boa qualidade discriminativa.

Diante disso, verifica-se um problema importante na estruturação da prova. Pois considerando um exame que objetiva, atualmente, a seleção de candidatos para os cursos de graduação oferecidos pelas IES, é desejável que apresentem boa qualidade discriminativa. Após essa avaliação inicial, foi realizada a análise da dimensionalidade com base na análise paralela e a FIFA, com o objetivo de compreender a estrutura fatorial dos itens.

Análise da dimensionalidade das provas do Enem 2017

Nesta seção foi conduzida uma análise paralela com a verificação de componentes principais com base em uma matriz de correlações tetracóricas. Na Tabela 2 estão dispostos os valores próprios (*eigen value*) para as cinco primeiras dimensões de cada prova do Enem 2017.

TABELA 2 - *Eigen value* da análise de componentes principais, Enem 2017

FATORES	LC		MT		CN		CH	
	c.p	d.s	c.p	d.s	c.p	d.s	c.p	d.s
1	6,80	1,12	5,04	1,11	4,52	1,11	8,09	1,12
2	1,52	1,10	1,81	1,10	1,61	1,1	1,48	1,11
3	1,23	1,10	1,17	1,09	1,17	1,09	1,11	1,10
4	1,11	1,09	1,15	1,08	1,14	1,08	1,09	1,09
5	1,06	1,08	1,11	1,07	1,13	1,07	1,07	1,08
6	1,05	1,07	1,07	1,07	1,10	1,07	1,05	1,08
7	1,03	1,07	1,03	1,06	1,06	1,06	1,04	1,07
8	1,00	1,06	1,02	1,05	1,05	1,05	1,03	1,06
9	0,98	1,06	1,01	1,05	1,03	1,05	0,99	1,06
10	0,98	1,05	0,99	1,04	1,02	1,04	0,96	1,05
11	0,96	1,04	0,97	1,04	1,00	1,04	0,95	1,05
12	0,94	1,04	0,95	1,03	0,98	1,03	0,93	1,04
Eigen1/Eigen2	4,47		2,78		2,80		5,47	

Fonte: Elaboração dos autores (2020).

Legenda: c.p: *eigen values* dos componentes principais; d.s: *eigen values* dos dados simulados.

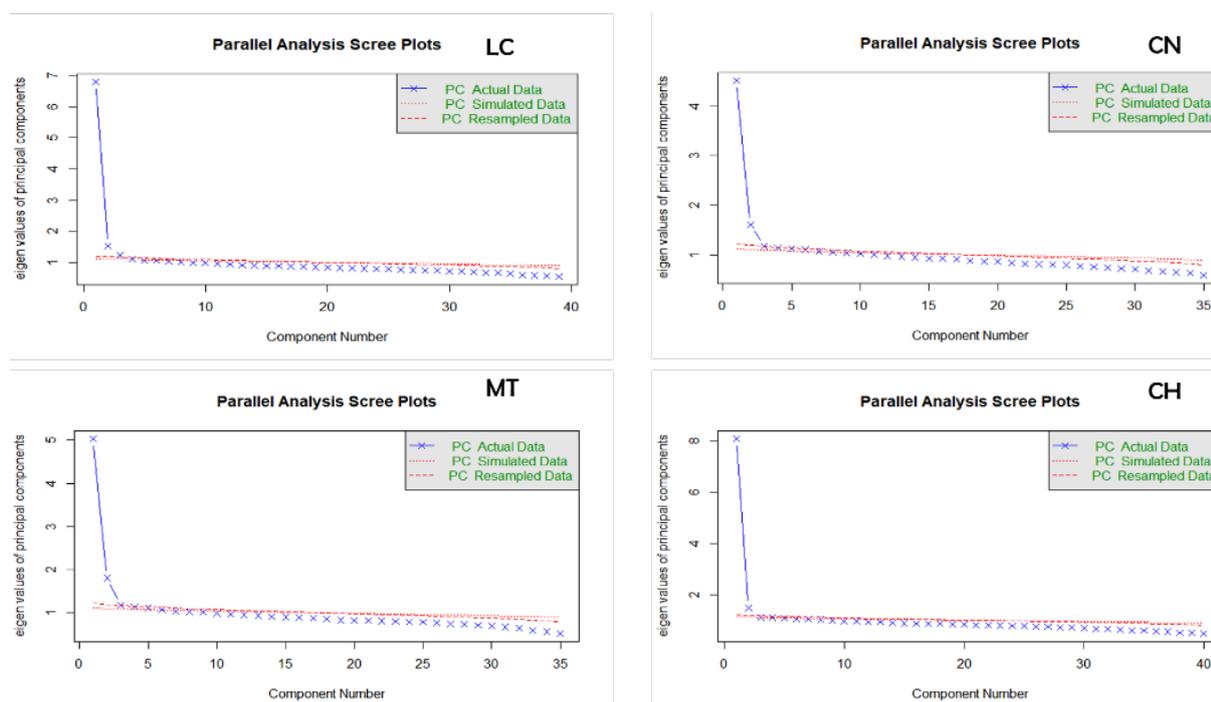
Nota: Destaque para *eigen values* dos dados reais quando é menor ou igual aos *eigen values* dos dados simulados.

Ao submeter os dados de cada prova à análise paralela, esperava-se a existência ou predominância de um fator, ou seja, que os itens fossem unidimensionais, considerando que esse é o pressuposto do modelo de TRI utilizado no exame. No entanto, nas comparações dos *eigen value* dos dados reais com os simulados, o teste indicou a existência de quatro fatores para a prova de LC, cinco fatores para as provas de MT, seis fatores para a prova de CN e três fatores para a prova de CH.

Entretanto, há um certo consenso no entendimento de que é suficiente admitir a existência de uma dimensão dominante, uma “dimensionalidade essencial” (STOUT, 1990). Diante disso, apenas duas provas parecem apresentar uma dimensão dominante. A prova de LC apresentou a razão entre o eigen1 e eigen2 de 4,47, ou seja, cerca de quatro vezes maior que o segundo, e a prova de CH com valor 5,47, ou seja, o eigen1 é cinco vezes maior que o eigen2. Para as provas de MT e CN, o eigen1 é apenas duas vezes maior que o eigen2.

Para melhor visualização da magnitude dos *eigen value* das provas, foi construído um *scree plot* dos valores de cada fator dos dados reais e dos dados simulados (Figura 1).

FIGURA 1 - Scree plot da análise paralela das provas do Enem 2017



Fonte: Elaboração dos autores (2020).

Nesses gráficos, torna-se mais evidente a existência de uma dimensão dominante nas provas de LC e CH. No entanto, nas provas de MT e CN, evidencia-se, de forma mais acentuada, a presença de uma segunda dimensão.

Após uma análise exploratória inicial da dimensionalidade dos dados por meio da correlação ponto biserial e da análise paralela, submetemos os itens a uma FIFA. Uma primeira análise foi realizada com os valores de ajuste dos modelos com base nas medidas do RMSEA, SRMSR, TLI e CFI. Os valores desses índices podem ser observados para cada prova na Tabela 3.

TABELA 3 - Índice de ajuste dos modelos estimados para as provas do Enem 2017

PROVAS	MODELOS	RMSEA	SRMSR	TLI	CFI
LC	1	0,011	0,014	0,987	0,988
	2	0,009	0,013	0,991	0,992
	3	0,008	0,012	0,993	0,994
	4	0,007	0,011	0,993	0,995
MT	1	0,009	0,014	0,983	0,985
	2	0,009	0,013	0,984	0,987
	3	0,008	0,013	0,988	0,990
	4	0,007	0,012	0,989	0,992
CN	1	0,009	0,014	0,977	0,980
	2	0,009	0,013	0,980	0,983
	3	0,008	0,012	0,984	0,988
	4	0,008	0,012	0,984	0,988

(Continua)

(Continuação)

PROVAS	MODELOS	RMSEA	SRMSR	TLI	CFI
CH	1	0,010	0,013	0,992	0,993
	2	0,009	0,012	0,994	0,995
	3	0,008	0,011	0,995	0,996
	4	0,007	0,010	0,996	0,997

Fonte: Elaboração dos autores (2020).

Como foi constatado na Tabela 3, os valores de ajuste para todas as provas estimados para os modelos de 1, 2, 3 e 4 parâmetros mostraram-se adequados. Os valores do RMSEA e SRMSR se mantiveram abaixo de 0,05 e os valores do TLI e CFI acima de 0,95. Esses índices não conseguiram identificar com precisão o modelo fatorial mais adequado para esses itens. Isso pode ter ocorrido porque esses índices são sensíveis a grandes amostras.

Também foram analisados os valores dos índices de ajuste dos modelos com base nos valores de AIC e BIC. Os dados dessa análise estão dispostos na Tabela 4.

TABELA 4 - Critério de Informação Baysiano (BIC) e Critério de Informação de Akaike (AIC) dos modelos estimados para as provas do Enem, 2017

DIMENSÕES	LC		MT		CN		CH	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1	469439,3	470282,9	380552,8	381309,9	387059,2	387816,3	468863,2	469728,4
2	469253,3	470370,9	380444,1	381446,3	387063,7	388065,9	468714,5	469861,0
3	469121,0	470505,4	380366,6	381606,8	386936,4	388176,6	468616,5	470037,0
4	469059,6	470703,5	380355,4	381826,4	386892,6	388363,5	468608,2	470295,4

Fonte: Elaboração dos autores (2020).

O índice de ajuste do AIC, como apresentado na literatura, tende a ser mais permissivo quanto ao número de fatores (NYLUND; ASPAROUHOV; MUTHÉN, 2007), o que ocorreu nesta pesquisa. Os valores decrescem à medida que são acrescentados mais fatores ao modelo. Indicando, por esse índice, que quatro ou mais fatores se ajustam bem.

Ao verificar os índices de ajuste oferecidos pelo BIC, observamos que o menor valor, portanto, o modelo mais bem ajustado, foi o dos modelos unidimensionais das provas. Notamos também que, à medida que são acrescentados mais fatores, os valores do BIC aumentam. Isso indica que se perde qualidade no ajuste ao estimar modelos com mais de uma dimensão. No entanto, embora o BIC tenha sido considerado mais consistente na indicação do número de dimensões a serem retidas na análise fatorial, o teste é considerado como conservador, ou seja, valoriza modelos com menos dimensões.

Também foi analisado o ID por meio da comparação entre os modelos. Os ID das comparações dos modelos estão dispostos na Tabela 5.

TABELA 5 - Índice de dimensionalidade das provas do Enem 2017

PROVAS	COMPARAÇÃO	X ₂	X ₂ CORR	GL	ID
LC	M1 - M2	262,002	87,334	38	2,298
	M2 - M3	206,238	68,746	37	1,858
	M3 - M4	133,464	44,488	36	1,235
MT	M1 - M2	176,698	58,899	34	1,732
	M2 - M3	143,442	47,814	33	1,499
	M3 - M4	75,188	25,063	32	0,783
CN	M1 - M2	63,559	21,186	34	0,623
	M2 - M3	193,233	64,411	33	1,952
	M3 - M4	107,810	35,937	32	1,123
CH	M1 - M2	226,643	75,547	39	1,937
	M2 - M3	173,993	57,998	38	1,526
	M3 - M4	82,315	27,438	37	0,741

Fonte: Elaboração dos autores (2020).

Em todas as comparações realizadas para as provas de MT, CN e CH, os valores de ID foram positivos menores que 2,0. Dessa forma, o ID indica que essas provas são unidimensionais. No prosseguimento das comparações, a prova de LC se ajustou melhor ao modelo de dois fatores. As demais provas tiveram melhor ajuste ao modelo unidimensional. Diante disso, esses resultados corroboram os apresentados pelo critério BIC.

No prosseguimento das análises, as cargas fatoriais de cada item também são indicativos da dimensionalidade destes, já que consistem na correlação do item com o fator. Dessa forma, espera-se que se os itens das provas, especialmente da prova de LC, são unidimensionais, cargas fatoriais altas, 0,30 ou mais, sejam identificadas. Optamos por realizar sucessivas análises das cargas fatoriais dos itens até se obter itens com cargas satisfatórias. As cargas fatoriais dos itens para o modelo unidimensional estão na Tabela 6.

Tabela 6 - Cargas fatoriais dos itens com um fator do Enem 2017

ITENS	LC			MT	CN		CH	
	1º EXTRAÇÃO	2º EXTRAÇÃO	3º EXTRAÇÃO	1º EXTRAÇÃO	1º EXTRAÇÃO	2º EXTRAÇÃO	1º EXTRAÇÃO	2º EXTRAÇÃO
i1	0,54	0,52	0,51	0,79	0,74	0,74	0,78	0,77
i2	0,69	0,69	0,69	0,63	0,38	0,36	0,68	0,67
i3	0,51	0,49	0,49	0,82	0,46	0,45	0,41	0,38
i4	0,40	0,41	0,41	0,86	0,48	0,48	0,44	0,44
i5	0,65	0,65	0,65	0,83	0,64	0,64	0,59	0,58
i6	0,45	0,45	0,45	*	0,82	0,81	0,48	0,48
i7	0,57	0,56	0,56	0,92	0,63	0,63	0,81	0,81
i8	0,24	**	**	0,57	0,87	0,87	*	*
i9	0,46	0,46	0,45	0,76	0,26	**	0,77	0,76

(Continua)

(Continuação)

ITENS	LC			MT	CN		CH	
	1º EXTRAÇÃO	2º EXTRAÇÃO	3º EXTRAÇÃO	1º EXTRAÇÃO	1º EXTRAÇÃO	2º EXTRAÇÃO	1º EXTRAÇÃO	2º EXTRAÇÃO
i10	0,50	0,49	0,48	0,88	*	*	0,28	**
i11	0,59	0,59	0,59	*	*	*	0,47	0,47
i12	0,44	0,43	0,43	0,87	0,60	0,60	0,68	0,67
i13	0,54	0,53	0,53	0,81	0,63	0,63	0,76	0,76
i14	0,77	0,78	0,77	0,78	0,95	0,95	0,25	**
i15	0,23	**	**	0,86	0,69	0,69	0,73	0,73
i16	0,44	0,44	0,44	0,79	0,86	0,86	0,32	0,32
i17	0,41	0,41	0,41	0,60	0,76	0,76	0,82	0,82
i18	0,60	0,59	0,59	*	0,67	0,67	0,77	0,76
i19	0,67	0,66	0,66	*	0,81	0,81	0,18	**
i20	0,77	0,77	0,77	0,89	0,43	0,43	0,50	0,49
i21	0,36	0,36	0,36	*	*	*	0,62	0,60
i22	0,58	0,58	0,57	0,67	0,87	0,87	0,53	0,53
i23	0,46	0,46	0,46	0,57	*	*	*	*
i24	0,59	0,58	0,58	0,34	0,78	0,78	0,35	0,35
i25	*	*	*	0,84	0,12	**	*	*
i26	0,19	**	**	0,72	*	*	0,50	0,50
i27	0,63	0,63	0,63	0,52	0,68	0,66	0,74	0,74
i28	0,23	0,23	***	0,81	0,81	0,80	0,78	0,78
i29	*	*	*	0,86	0,67	0,67	0,55	0,55
i30	0,44	0,44	0,45	0,76	*	*	*	*
i31	0,66	0,66	0,66	0,60	0,87	0,87	0,67	0,66
i32	0,71	0,71	0,70	0,85	*	*	0,81	0,81
i33	*	*	*	0,72	0,82	0,81	0,85	0,85
i34	0,59	0,59	0,61	*	*	*	0,48	0,48
i35	*	*	*	0,66	*	*	0,87	0,86
i36	*	*	*	*	0,73	0,75	*	*
i37	0,40	0,40	0,40	0,65	0,63	0,63	0,70	0,71
i38	0,59	0,59	0,58	0,91	0,86	0,85	0,55	0,54
i39	*	*	*	0,92	0,68	0,70	0,59	0,58
i40	0,67	0,68	0,68	*	0,89	0,89	0,71	0,71
i41	0,33	0,33	0,33	*	0,75	0,76	0,73	0,72
i42	0,47	0,48	0,48	0,78	0,86	0,86	0,66	0,65
i43	0,39	0,39	0,39	0,73	0,86	0,85	0,71	0,70
i44	0,74	0,74	0,73	*	*	*	0,86	0,86
i45	0,80	0,80	0,79	0,74	0,89	0,89	0,63	0,62
%Var	29,5	31,2	31,8	58,1	52,4	55,0	40,9	43,0

Fonte: Elaboração dos autores (2020).

Notas: a) *Itens excluídos por baixa discriminação (< 0,15). **Itens excluídos por baixa carga fatorial na primeira extração (< 0,30). ***Itens excluídos por baixa carga fatorial na segunda extração (< 0,30).

b) Destaque em negrito para cargas fatoriais menores que 0,30.

Ao analisar as cargas fatoriais dos itens das provas do Enem 2017, apenas os itens da prova de MT apresentaram adequação já na primeira extração, com cargas fatoriais variando entre 0,34 e 0,92. O modelo final apresentou um percentual de variância explicado de 58,1%. Durante o processo de ajuste de um modelo fatorial para essa prova, 10 de 45 itens foram excluídos por baixa discriminação, permanecendo 35 itens.

Para os itens da prova de LC, foram necessárias três extrações para a obtenção de um modelo fatorial satisfatório. No modelo final, dez itens também foram excluídos, inicialmente por baixa discriminação e ao aplicar o critério da carga fatorial durante o ajuste do modelo, permanecendo 35 itens. As cargas fatoriais desses itens variaram entre 0,33 e 0,79. O percentual de variância explicado foi de 31,8%.

Para os itens da prova de CN, foram necessárias duas extrações para a obtenção de um modelo fatorial adequado. No modelo final 12 itens foram excluídos por baixa discriminação e baixas cargas fatoriais, permanecendo 33 itens no modelo final. As cargas fatoriais variaram entre 0,36 e 0,95. O percentual de variância explicado foi de 55%.

Por último, os itens da prova de CH se ajustaram também na segunda extração. Inicialmente foram excluídos cinco itens por baixa discriminação e depois mais três itens por baixa carga fatorial na primeira extração. O modelo final resultou em 37 itens como cargas fatoriais entre 0,32 e 0,86 e percentual de variância explicado de 43%.

Na Tabela 7 estão os itens excluídos durante o processo de análise fatorial para cada prova do exame.

TABELA 7 - Itens excluídos durante a análise fatorial exploratória, Enem 2017

PROVAS	TOTAL DE ITENS	$r_{pb} < 0,15$	1ª EXTRAÇÃO >0,30	2ª EXTRAÇÃO >0,30	TOTAL DE ITENS EXCLUÍDOS
LC	45	25, 29, 33, 35, 36 e 39	8, 15 e 28	28	10
MT	45	6, 11, 18, 19, 21, 34, 36, 40, 41 e 44	-	-	10
CN	45	10, 11, 21, 23, 26, 30, 32, 34, 35, e 44	9 e 25	-	12
CH	45	8, 23, 25, 30 e 36	10, 14 e 19		8

Fonte: Elaboração dos autores (2020).

Durante a análise fatorial exploratória, foram excluídos dez itens da prova de LC e MT, 12 itens de CN e oito itens de CH. Esperávamos que as análises indicassem que todos os itens apresentassem unidimensionalidade. No entanto, isso não ocorreu, o que demonstra problemas métricos dos itens das provas. Investigações mais aprofundadas são necessárias para analisar o impacto desses problemas na estimação das habilidades dos sujeitos. Supomos, inicialmente, que esses problemas podem prejudicar os candidatos no processo de seleção para os cursos de graduação das IES.

Para todas as provas do exame, foi possível ajustar um modelo unidimensional, no entanto, após a exclusão de muitos itens (ver Tabela 8). Embora seja parcimonioso um modelo unidimensional, considerando apenas uma dimensão dominante (STOUT, 1990), essa dimensão não conseguiu explicar a variabilidade dos outros itens. Provavelmente esses itens podem ser explicados pela existência de outras dimensões. Nesse caso, modelos de TRI multidimensionais são mais adequados (RECKASE, 2009).

Os pressupostos dos modelos de TRI unidimensionais, o mesmo utilizado no Enem, têm sido fortemente criticados. Tavares (2013, p. 69) faz o seguinte questionamento sobre esse ponto: “Se pensarmos em um exame específico utilizado em larga escala em nosso país, como o Enem, podemos assegurar que a TRI baseada em um modelo logístico unidimensional de três parâmetros é a melhor opção metodológica para esse caso?”. Ainda acrescenta apontando a perspectiva interdisciplinar do exame como um de seus diferenciais, o que torna mais difícil essa ideia.

Uma das saídas apontadas para esse problema da dimensionalidade do Enem e seu caráter interdisciplinar não comportado pelo modelo de TRI unidimensional é a utilização de modelos multidimensionais (OLIVEIRA, 2015). Esse problema da estrutura fatorial do exame foi testado empiricamente em algumas pesquisas.

Alguns estudos que realizam uma análise da estrutura fatorial dos itens para a compreensão dos construtos medidos por esse exame têm sido realizados. Em uma pesquisa se analisou o Enem de 2010 por meio da análise fatorial exploratória e confirmatória (MUNER, 2013). Nesse estudo, a autora conseguiu ajustar um modelo unidimensional com explicação de 30% da variância apenas para a prova de LC, mas após a exclusão de alguns itens. Na prova de CH e MT, foi ajustado um modelo bidimensional, ambos com 40% de variância explicada, e na prova de CN, um modelo com três dimensões e 50% de variância explicada. A estrutura fatorial dessas provas foi ratificada pela análise fatorial confirmatória com bons valores de ajuste do modelo.

Resultados que corroboram os deste estudo foram encontrados em outro que analisa o Enem, mais especificamente a prova de MT de 2015 (PRIMI; CICCHETTO, 2018). Nessa pesquisa, os autores implementaram uma análise de componentes principais sobre os resíduos obtidos no ajuste do modelo de TRI de Rasch. Com essa análise foi identificada a existência de um fator secundário importante indicando sistematicidade no acerto de participantes considerados como de baixa habilidade em itens difíceis.

Em outro estudo analisou-se as quatro provas do Enem de 2012 tomadas com prova única com 180 itens sob a ótica da TRI uni e multidimensional (VIEIRA, 2016). As análises indicaram que um modelo unidimensional se ajustou bem. O modelo multidimensional com quatro fatores também apresentou bom ajuste,

mas não sugeriu que os itens estavam relacionados com suas respectivas áreas de conhecimento, o que dificulta a interpretação semântica dos fatores.

Embora Vieira (2016) tenha evidenciado que um modelo unidimensional se ajusta bem aos itens das quatro áreas, torna-se ainda mais complicada a interpretação desse fator e recai nas discussões apresentadas por Tavares (2013), de que é difícil compreender que apenas um fator latente é responsável pelas respostas aos itens de um teste educacional, ainda mais se tratando do Enem.

Por outro lado, uma outra pesquisa, ao analisar a estrutura fatorial de todos os 180 itens do Enem de 2016 por meio do método de análise fatorial de informação plena, identificou a existência de pelo menos duas dimensões (PICCIRILLI; SOUZA, 2018). Ao realizar a interpretação dos itens, os autores sugerem que os dois fatores medidos pelos itens são interpretação de texto e raciocínio lógico.

Também foram realizadas análises psicométricas da estrutura fatorial das provas do exame antes de 2009, quando o Enem ainda utilizava técnicas de psicometria clássica na análise dos resultados. Em estudo realizado com o Enem de 1999 (NOJOSA, 2002) e 2001 (COSTA, 2015), foi possível ajustar um modelo de TRI multidimensional com cinco fatores. Nojosa (2002) ressalta que esse modelo seria o mais adequado para a análise do exame considerando suas características de conteúdo.

Os resultados indicaram que alguns itens das provas do exame não atenderam ao pressuposto de unidimensionalidade, pois só foi possível ajustar um modelo unidimensional após a exclusão de alguns itens. Na prova de MT, por exemplo, um modelo unidimensional foi ajustado após a exclusão de 12 dos 45 itens da prova. Com base nessas evidências ainda não é possível realizar conclusões fortes sobre a dimensionalidade do Enem.

Não obstante, o exame apresenta algumas inadequações quanto à sua proposta inicial (ANDRADE, 2012): i) para que os resultados sejam comparados, é necessário que haja alguns itens em comum, no entanto, não é possível que se apliquem itens já utilizados em outros exames, já que é um teste de seleção e suas provas são divulgadas na íntegra; ii) para a validação dos itens, é necessário que eles sejam pré-testados em uma amostra do universo de participantes, quando isso ocorre alguns possíveis candidatos têm acesso aos itens anteriormente; iii) uma das justificativas para a utilização da TRI é que ela tem foco na análise do item, assim é possível atribuir pesos distintos para os itens, algo que também é possível e legítimo com a TCT, com a qual se solucionariam os problemas da dificuldade de interpretação dos resultados do exame pelos candidatos, de modo que estes não têm como estimar o seu resultado.

Além desses problemas apresentados, há uma outra questão que precisa ser levantada. Segundo uma nota técnica publicada pelo Ministério da Educação

(MEC),¹ a utilização da TRI no Enem tem duas finalidades principais: i) permitir a comparabilidade entre os anos; ii) permitir a aplicação do exame várias vezes ao ano.

Em relação ao primeiro objetivo apresentado, cabe uma reflexão. A comparabilidade entre os resultados é particularmente importante quando se quer acompanhar a evolução do aprendizado de um determinado grupo (ANDRADE; LAROS; GOUVEIA, 2010), com vistas à tomada de decisão ao redirecionamento de políticas e recursos. No entanto, o objetivo principal do Enem atualmente é ser parâmetro de seleção para os cursos de graduação nas universidades, institutos federais e instituições privadas de ensino superior. Dessa forma, os resultados têm como foco o desempenho individual dos candidatos para fins de classificação, que não necessariamente são estudantes. Assim, resultados comparáveis para tomada de decisão e definição de políticas públicas educacionais já são oferecidos por outras provas, como as realizadas no âmbito do Sistema de Avaliação da Educação Básica (Saeb) (KLEIN, 2009).

Quanto ao segundo objetivo alegado pelo Inep, este não tem se concretizado na prática, pois o exame ainda é aplicado uma única vez ao ano. Além desses dois principais objetivos, aponta-se também a possibilidade de se aplicarem provas distintas, ou seja, com itens diferentes, sem, contudo, alterar o nível de dificuldade da prova. Esse procedimento também não tem sido realizado. Atualmente são aplicados cadernos de prova com cores diferentes, em que há apenas a alteração da posição dos itens em cada caderno.

Com base nisso, acredita-se que os resultados desta pesquisa podem contribuir para a discussão da utilização do modelo de TRI unidimensional no Enem, considerando a falta de evidência sobre a sua estrutura latente. No entanto, ressalta-se a necessidade de mais estudos empíricos que possibilitem melhores informações sobre essa questão, sobretudo, debates e reflexões em relação à real necessidade desse modelo para os atuais objetivos do exame. Também ressaltamos a necessidade de pesquisas que analisem se a ordem dos itens interfere na dimensionalidade do exame; e o impacto da falta de ajuste ao modelo unidimensional na pontuação dos participantes. Aspectos esses que não são abordados nesta pesquisa.

Como discutido, não é apenas uma questão de modelo matemático, mas de adequação a pressupostos que, se violados, podem prejudicar milhões de candidatos que almejam uma vaga nas IES de todo o país.

1 Disponível em: http://download.inep.gov.br/educacao_basica/enem/nota_tecnica/2011/nota_tecnica_tri_enem_18012012.pdf. Acesso em: 14 dez. 2020.

REFERÊNCIAS

- ANDRADE, Dalton Francisco de; TAVARES, Heliton Ribeiro; VALLE, Raquel da Cunha. *Teoria de resposta ao item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística, 2000.
- ANDRADE, Gisele Gama. A metodologia do Enem: uma reflexão. *Série-Estudos*, Campo Grande, MS, n. 33, p. 67-76, 2012.
- ANDRADE, Josemberg Moura de; LAROS, Jacob Arie; GOUVEIA, Valdiney Veloso. O uso da teoria de resposta ao item em avaliações educacionais: diretrizes para pesquisadores. *Avaliação Psicológica*, Campinas, SP, v. 9, n. 3, p. 421-435, 2010.
- ANDRIOLA, Wagner Bandeira. Avaliação da aprendizagem: uma análise descritiva segundo a teoria de resposta ao item (TRI). *Educação em Debate*, Fortaleza, v. 20, n. 36, p. 93-102, 1998.
- BARTHOLOMEW, David J. Factor analysis for categorical data. *Journal of the Royal Statistical Society*, v. 42, n. 3, p. 293-321, 1980.
- BÉLAND, Sébastien; JOLANI, Shahab; PICHETTE, François; RENAUD, Jean-Sébastien. Impact of simple substitution methods for missing data on classical test theory difficulty and discrimination. *The Quantitative Methods for Psychology*, v. 14, n. 3, p. 180-192, 2018.
- BRASIL. Ministério da Educação. *Teoria de resposta ao item avalia habilidade e minimiza o “chute” de candidatos*. 2011. Disponível em: <http://portal.mec.gov.br/ultimas-noticias/389-ensino-medio-2092297298/17319-teoria-de-resposta-ao-item-avalia-habilidade-e-minimiza-o-chute>. Acesso em: 14 dez. 2020.
- CANÇADO, Regina; CASTRO, Maria Jose Pereira; OLIVEIRA, Isabella Fernandes de. Análise pedagógica de itens de teste por meio da teoria de resposta ao item. In: REUNIÃO DA ABAVE, 7., 2013, Brasília-DF. *Anais [...]*. Brasília-DF, 2013.
- CHALMERS, R. Philip. MIRT : A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, v. 48, n. 6, p. 1-29, 2012.
- COSTA, Carlos Eduardo Sousa. *Análise da dimensionalidade e modelagem multidimensional pela TRI no Enem (1998-2008)*. 2015. Dissertação (Mestrado em Métodos e Gestão em Avaliação) – Universidade Federal de Santa Catarina, Florianópolis, 2015.
- COSTA, Patrícia; FERRÃO, Maria Eugénia. On the complementarity of classical test theory and item response models: Item difficulty estimates and computerized adaptive testing. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 23, n. 88, p. 593-610, 2015.
- FERREIRA, Francisco Fialho Guedes. *Escala de proficiência para o Enem utilizando teoria de resposta ao item*. 2009. Dissertação (Mestrado em Matemática e Estatística) – Universidade Federal do Pará, Belém, 2009.
- GOMES, Cristiano Mauro Assis; GOLINO, Hudson Fernandes; PERES, Alexandre José de Souza. Análise da fidedignidade composta dos escores do Enem por meio da análise fatorial de itens. *European Journal of Education Studies*, v. 5, n. 8, p. 331-344, 2018. <http://dx.doi.org/10.46827/ejes.v0i0.2178>.

- HAYTON, James C.; ALLEN, David G.; SCARPELLO, Vida. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, v. 7, n. 2, p. 191-205, 2004.
- KLEIN, Ruben. Utilização da teoria de resposta ao item no Sistema Nacional de Avaliação da Educação Básica (Saeb). *Meta: Avaliação*, Rio de Janeiro, v. 1, n. 2, p. 125-140, 2009.
- KLEIN, Ruben. Alguns aspectos da Teoria de Resposta ao Item relativos à estimação das proficiências. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 21, n. 78, p. 35-56, 2013.
- MUNER, Luana Comito. *Análise fatorial exploratória e confirmatória do Enem 2010 com estudantes paulistas*. 2013. Dissertação (Mestrado em Psicologia) – Universidade São Francisco, Itatiba-SP, 2013.
- MUÑIZ, José. *Teoría clásica de los testes*. Madrid: Pirámide, 1994.
- NEWMAN, Daniel A. Missing data: five practical guidelines. *Organizational Research Methods*, v. 17, n. 4, p. 372-411, 2014.
- NOJOSA, Ronaldo Targino. Teoria da Resposta ao Item (TRI): modelos multidimensionais. *Estudos em Avaliação Educacional*, São Paulo, n. 25, p. 123-166, jan./jun. 2002.
- NYLUND, Karen L.; ASPAROUHOV, Tihomir; MUTHÉN, Bengt O. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, v. 14, n. 4, p. 535-569, 23 out. 2007.
- OLIVEIRA, Bolivar Alves. Interdisciplinaridade e dimensionalidade das provas do Enem. In: REUNIÃO DA ABAVE, 8., 2015, Florianópolis. *Anais [...]*. Florianópolis: Abave, 2015.
- PASQUALI, Luiz. *Psicometria: teoria dos testes na psicologia e na educação*. 3. ed. Petrópolis, RJ: Vozes, 2009.
- PICCIRILLI, Giovanni Pastori; SOUZA, Aparecida Donizete Pires de. Teoria da Resposta ao Item multidimensional: análise da dimensionalidade da prova do Enem 2016. In: CONGRESSO DE INICIAÇÃO CIENTÍFICA DA UNESP, 30., 2018, Presidente Prudente. *Anais [...]*. Presidente Prudente: Unesp, 2018.
- PRIMI, Ricardo; CICCHETTO, Airton A. Como os escores do Enem são atribuídos pela TRI? In: CONBRATRI: MÉTODOS PARA DETECÇÃO DE FRAUDES EM TESTES, 6., Juiz de Fora. *Anais [...]*. Juiz de Fora: Abave, 2018.
- RECKASE, Mark D. *Multidimensional item response theory*. New York, NY: Springer, 2009.
- REVELLE, W. *Psych: Procedures for personality and psychological research*, 2017. Disponível em: <https://cran.r-project.org/package=psych>. Acesso em: 14 dez. 2020.
- RIZOPOULOS, Dimitris. Ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, v. 17, n. 5, p. 1-25, 2006.
- SARTES, Laisa Marcocela Andreoli; SOUSA-FORMIGONI, Maria Lucia Oliveira de. Avanços na psicometria: da teoria clássica dos testes à teoria de resposta ao item. *Psicologia: Reflexão e Crítica*, Porto Alegre, v. 26, n. 2, p. 241-250, 2013.

SOARES, Denilson Junio Marques; SOARES, Talita Emidio Andrade; SANTOS, Wagner dos. Análise da qualidade psicométrica da prova de matemática do Exame Nacional do Ensino Médio brasileiro de 2018. *Revista Actualidades Investigativas en Educación*, v. 21, n. 1, p. 1-28, 2020. <http://dx.doi.org/10.15517/aie.v21i1.42338>.

SOUSA, Leandro Araujo de; BRAGA, Adriana Eufrásio. Teoria clássica dos testes e teoria de resposta ao item em avaliação educacional. *Revista de Instrumentos, Modelos e Políticas em Avaliação Educacional*, Itaperi, CE, v. 1, n. 1, p. e020002, 2020.

SOUSA, Leandro Araujo de; PONTES JUNIOR, José Airton de Freitas; BRAGA, Adriana Eufrásio. Educação física no Exame Nacional do Ensino Médio: análise via teoria clássica dos testes. *Revista Actualidades Investigativas en Educación*, v. 20, n. 1, p. 257-277, Abr. 2020. <http://dx.doi.org/10.15517/aie.v20i1.40126>.

STOUT, William F. A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, v. 55, n. 2, p. 293-325, 1990.

TAVARES, Cristina Zukowsky. Teoria da resposta ao item: uma análise crítica dos pressupostos epistemológicos. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 54, p. 56-76, jan./abr. 2013.

TOFFOLI, Sônia Ferreira Lopes. Análise da qualidade de uma prova de matemática do Exame Nacional do Ensino Médio. *Educação e Pesquisa*, São Paulo, v. 45, e187128, 2019. <http://dx.doi.org/10.1590/S1678-4634201945187128>.

TOFFOLI, Sônia Ferreira Lopes; ANDRADE, Dalton Francisco de; BORNIA, Antonio Cezar; QUEVEDO-CAMARGO, Gladys. Avaliação com itens abertos: validade, confiabilidade, comparabilidade e justiça. *Educação e Pesquisa*, São Paulo, v. 42, n. 2, p. 343-358, jun. 2016.

TRAVITZKI, Rodrigo. Avaliação da qualidade do Enem 2009 e 2011 com técnicas psicométricas. *Estudos em Avaliação Educacional*, São Paulo, v. 28, n. 67, p. 256-288, jan./abr. 2017. <http://dx.doi.org/10.18222/ea.v28i67.3910>.

VALLE, Raquel da Cunha. Teoria de resposta ao item. *Estudos em Avaliação Educacional*, São Paulo, n. 21, p. 7-92, jan./jun. 2000.

VIANNA, Heraldo Marelim. *Testes em educação*. São Paulo: Ibrasa, 1976.

VIEIRA, Nara Núbia. *As provas das quatro áreas do Enem vista como prova única na ótica de modelos da Teoria da Resposta ao Item uni e multidimensional*. 2016. Dissertação (Mestrado Profissional em Métodos e Gestão da Avaliação) – Universidade Federal de Santa Catarina, Florianópolis, 2016.

Recebido em: 15 MARÇO 2020

Aprovado para publicação em: 4 DEZEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v0ix.6931>

O QUE DIFERENCIA ESCOLAS COM BOM DESEMPENHO NOS ANOS FINAIS DO ENSINO FUNDAMENTAL?

 LUIZ GUILHERME SCORZAFAVE^I

 ERNESTO MARTINS FARIA^{II}

 BRUNA ALVES^{III}

^I Universidade de São Paulo (USP), *campus* de Ribeirão Preto, Ribeirão Preto-SP, Brasil; scorza@usp.br

^{II} Interdisciplinaridade e Evidências no Debate Educacional (Iede), São Paulo-SP, Brasil, ernesto.faria@gmail.com

^{III} Instituto Sonho Grande (ISG), São Paulo-SP, Brasil, bruna.alves@sonhogrande.org

RESUMO

O artigo identifica escolas de anos finais do ensino fundamental no Brasil com bons resultados escolares e que atendem alunos de baixo nível socioeconômico. Foram identificadas 329 escolas que atendiam os critérios propostos (baseados em proficiência e evasão escolar) em 2015. Por meio de análise de diferenças em diferenças, foram investigados fatores associados ao melhor desempenho dessas escolas. Os resultados mostram que elas apresentam menores índices de violência, mais crença dos professores no potencial futuro dos alunos, menor indisciplina, mais alunos que fazem lição de casa, maior correção dessas tarefas e menor absenteísmo discente. Salário, formação e experiência de professores/diretores não foram diferentes entre as escolas que se destacaram e as demais.

PALAVRAS-CHAVE AVALIAÇÃO DA EDUCAÇÃO • QUALIDADE DA EDUCAÇÃO • ANOS FINAIS DO ENSINO FUNDAMENTAL • EQUIDADE EM EDUCAÇÃO.

¿QUÉ DIFERENCIA A LAS ESCUELAS CON BUEN DESEMPEÑO EN LOS ÚLTIMOS AÑOS DE LA ENSEÑANZA PRIMARIA?

RESUMEN

El artículo identifica escuelas de los últimos años de la enseñanza primaria en Brasil con buenos resultados escolares y que atienden a alumnos de bajo nivel socioeconómico. Se identificaron 329 escuelas que respetaban los criterios propuestos (basados en proficiencia y evasión escolar) en 2015. Por medio del análisis de diferencias, se investigaron factores vinculados al mejor desempeño en esas escuelas. Los resultados muestran que ellas presentan menores índices de violencia, más creencia de los profesores en el potencial futuro de los alumnos, menos indisciplina, más alumnos que hacen sus deberes para casa, menos errores en esas tareas y menos ausentismo de alumnos. Sueldos, formación y experiencia de profesores/directores no fueron diferentes entre las escuelas que se destacaron y las otras.

PALABRAS CLAVES EVALUACIÓN DE LA EDUCACIÓN • CALIDAD DE LA EDUCACIÓN • ÚLTIMOS AÑOS DE LA ENSEÑANZA PRIMARIA • EQUIDAD EN EDUCACIÓN.

WHAT DIFFERENTIATES SCHOOLS WITH GOOD PERFORMANCE IN THE FINAL YEARS OF ELEMENTARY EDUCATION?

ABSTRACT

The article identifies schools in the final years of elementary education in Brazil, with good academic results, and that serve students from low socioeconomic levels. 329 schools were identified that met the proposed criteria (based on proficiency and dropouts) in 2015. Factors associated with the better performance of these schools were investigated through the analysis of differences-in-differences. The results show that they have lower rates of violence, more teachers belief in the students' future potential, better discipline, more students doing homework, more correction of these tasks and less student absenteeism. There were no differences of salary, training and experience of teachers and principals between the schools that stood out and the others.

KEYWORDS EVALUATION OF EDUCATION • QUALITY OF EDUCATION • FINAL YEARS OF ELEMENTARY EDUCATION • EQUITY IN EDUCATION.

INTRODUÇÃO

Os anos finais do ensino fundamental têm se mostrado uma etapa relativamente negligenciada no debate educacional brasileiro referente à compreensão dos fatores escolares associados ao desempenho dos alunos, bem como no enfrentamento dos problemas de evasão e repetência que crescem nessa etapa de ensino. Por um lado, após a consolidação do processo de inclusão das crianças no ensino fundamental (EF), boa parte da atenção da agenda de pesquisa passou a focar em políticas e ações que pudessem ser trabalhadas a fim de garantir a alfabetização e a numeração dos alunos na idade correta, ou seja, nos primeiros anos desse nível de ensino. Grande parte das redes de ensino concentrou esforços e recursos na tentativa de resolver essas questões básicas, com base na premissa de que seria condição necessária para que avanços pudessem ser obtidos em etapas subsequentes de ensino. Por outro lado, o ensino médio apresenta sérios problemas de fluxo escolar. Entre 2016 e 2017, a taxa de promoção escolar no ensino médio foi de 78%, enquanto nos anos iniciais foi de 91% (BRASIL, 2018). Nesse sentido, alguns trabalhos buscaram identificar os motivos da baixa promoção escolar no ensino médio, bem como suas consequências (FUNDAÇÃO BRAVA *et al.*, 2017; SOUZA *et al.*, 2012; NERI, 2009; GREMAUD *et al.*, 2010).

No meio dessas duas etapas, estão os anos finais do ensino fundamental, que têm complexidades próprias. Muitos alunos, ao ingressarem nessa fase de ensino, trocam de escola e sofrem a transição de um professor generalista para professores especialistas (SANTOS *et al.*, 2017). Além disso, existe o desafio de consolidar habilidades cognitivas que já deveriam estar desenvolvidas ao final dos anos iniciais (DAVIS *et al.*, 2012). Por fim, verifica-se uma baixa integração entre os currículos dos anos iniciais e finais do ensino fundamental (BATISTA *et al.*, 2015).

Embora haja algum esforço de estudo dessa etapa de ensino (PADILHA *et al.*, 2012; DAVIS *et al.*, 2012; MANSUTTI *et al.*, 2007), são poucas as iniciativas que procuram compreender de que modo características específicas da escola e da gestão educacional se associam aos resultados dos alunos. Destacam-se, como exceções, os trabalhos do Fundo das Nações Unidas para a Infância (Unicef) (2007, 2009, 2010) que se basearam em pesquisas de campo *in loco* e apontam três ações pedagógicas como sendo de grande importância para um bom resultado no Sistema de Avaliação da Educação Básica (Saeb): a implantação de sistemas de avaliação e monitoramento; a presença de coordenadores nas escolas; e a utilização das matrizes da Prova Brasil como direcionador das práticas pedagógicas em Língua Portuguesa e Matemática, juntamente com o currículo. Os estudos também ressaltam a importância de um bom ambiente de aprendizagem, de boa formação do professor e de ações de suporte e valorização dos educadores. Outra exceção é o estudo de Menezes Filho (2012), que aponta o número de horas-aula como um dos fatores escolares mais associados a bons resultados no Saeb.

Entre os aspectos pouco estudados dos anos finais do ensino fundamental está a compreensão dos fatores que distinguem escolas que, apesar de atenderem majoritariamente alunos de baixo nível socioeconômico, conseguem obter bons resultados em termos de aprendizado. Madalozzo e Faria (2014) investigaram os fatores que explicam o sucesso de 215 escolas com indicadores muito bons nos anos iniciais do ensino fundamental, apesar de receberem alunos com nível socioeconômico baixo, denominadas escolas de excelência com equidade. O presente artigo realiza uma reflexão semelhante, contudo voltada aos anos finais do ensino fundamental.

Utilizando critérios semelhantes aos de Madalozzo e Faria (2014), foram selecionadas 329 escolas com bons resultados que recebem alunos de baixo nível socioeconômico. O objetivo da análise é tentar identificar fatores que diferenciam essas escolas de outras que não tiveram o mesmo êxito, apesar de serem semelhantes em várias dimensões, até 2011. Com dados secundários da Prova Brasil, é possível mapear ações e iniciativas das escolas e das redes que podem ter levado a esse melhor desempenho nos últimos anos. Particularmente, estamos interessados em investigar se tais práticas podem ser replicáveis em outras escolas e redes de ensino.

Além da introdução, este artigo está dividido em mais cinco seções. Na segunda seção, serão abordados os critérios estabelecidos para a seleção das escolas de bom desempenho, referidas neste estudo como escolas de excelência. Na terceira, serão descritas a base de dados e algumas características das 329 escolas selecionadas. Já na quarta seção, será apresentada a metodologia empregada. Discutem-se, então, na quinta seção, os resultados da análise econométrica, que mapeou as características que distinguem as escolas de bom desempenho das demais. A ideia foi realizar uma análise de diferenças em diferenças, encontrando, para cada escola de bom desempenho, uma outra que, em 2011, fosse semelhante a ela em fatores observáveis. Por fim, na última seção, são tecidas as considerações finais do trabalho.

CRITÉRIOS DE SELEÇÃO

Os indicadores de qualidade dos anos finais do ensino fundamental têm se tornando, recentemente, foco de grande preocupação. Uma vez que, nas avaliações externas, houve um progresso razoável nos anos iniciais desse nível de ensino, havia uma expectativa de que tal avanço impactasse os resultados dos anos finais do ensino fundamental. Infelizmente, não é isso que vem ocorrendo no Brasil, pois permanecem os problemas de baixa proficiência e elevada reprovação e evasão escolar nessa etapa de ensino. Ademais, essas complicações tendem a ser mais graves

quanto pior for a condição socioeconômica dos alunos (CARLOS, 2019; SOUZA *et al.*, 2012). Além disso, é importante que a busca por um melhor desempenho ocorra em condições mais igualitárias (TODOS PELA EDUCAÇÃO, 2010; SOARES; ANDRADE, 2006), sem induzir um processo de desistência da escola por parte dos alunos com mais dificuldade e promovendo o aprendizado de todos os estudantes, especialmente aqueles com maior *deficit* de aprendizado.

Nesse sentido, ganha importância a investigação de escolas com alunos de baixo nível socioeconômico, mas que conseguem se destacar nos anos finais do ensino fundamental. Tais exemplos podem servir como norteador para que as demais escolas possam adaptar práticas que estejam se mostrando bem-sucedidas.

Nesta seção, são discutidos os critérios para identificar escolas que atendem a população de baixo nível socioeconômico e que vêm se destacando em termos de proficiência nos anos finais do ensino fundamental. Foram considerados dois aspectos que diferenciam as escolas de anos finais do ensino fundamental e que acreditamos serem importantes para nossa análise. O primeiro diz respeito à complexidade da gestão escolar em municípios muito grandes. Assim, os municípios foram separados de acordo com o número de alunos nos anos finais do ensino fundamental (em 2015) em dois grupos: aqueles com mais de 50 mil alunos, e os demais. Com esse recorte, ficaram no primeiro grupo os 15 maiores municípios brasileiros: Belo Horizonte, Belém, Brasília, Campinas, Curitiba, Fortaleza, Goiânia, Guarulhos, Manaus, Porto Alegre, Recife, Rio de Janeiro, Salvador, São Luís e São Paulo.

O segundo aspecto refere-se ao perfil dos alunos que a escola recebe no 6º ano do ensino fundamental. Idealmente, gostaríamos de saber, para cada aluno da escola, qual foi seu desempenho na Prova Brasil ao término do 5º ano do ensino fundamental. Isso nos permitiria levar em conta a contribuição que os anos finais efetivamente deram para o aumento da proficiência dos alunos, pois pode haver escolas recebendo alunos advindos de escolas de baixa proficiência no 5º ano, enquanto outras poderiam estar recebendo alunos, predominantemente, de escolas de alta proficiência. Assim, identificada, para cada aluno do 9º ano do ensino fundamental em 2015, a escola em que ele se encontrava em 2011, calculamos o Índice de Desenvolvimento da Educação Básica (Ideb) de entrada no 6º ano.

Esse indicador foi obtido a partir da informação da distribuição dos alunos do 9º ano da escola em 2015, de acordo com a escola na qual o aluno concluiu o 5º ano em 2011, por meio da seguinte expressão:

$$IDEB_{ent,i} = \sum_{j=1}^J \pi_j IDEB_{5,j} \quad (1)$$

Ou seja, o Ideb de entrada no 6º ano da escola i é calculado como sendo a média ponderada do Ideb de 5º ano das escolas de origem dos alunos de 9º ano da escola i ,

onde o peso é definido pela proporção de alunos do 9º ano que vieram de cada escola de 5º ano j . Esse indicador permite identificar um perfil aproximado da proficiência dos alunos no momento do ingresso nos anos finais do ensino fundamental. Cabe destacar que, na construção da medida, consideramos a parte do Ideb de 5º ano apenas relativa à proficiência, ignorando a medida de fluxo nos primeiros anos do ensino fundamental, pois o objetivo era obter uma medida de proficiência inicial dos anos finais dessa etapa. Com base nisso, separamos as escolas dos anos finais do fundamental em três grupos, de acordo com o Ideb de entrada no 6º ano: aquelas com Ideb de entrada até 5,0 (baixo índice de entrada), entre 5,0 e 6,0 (médio índice de entrada) e maior que 6,0 (alto índice de entrada). Assim, formamos seis *clusters* de escolas, definidos pelo tamanho do município e Ideb de entrada no 6º ano.

TABELA 1 – Distribuição das escolas dos anos finais do ensino fundamental em seis *clusters* determinados pelo tamanho do município e o Ideb de entrada no 6º ano da escola

TAMANHO DO MUNICÍPIO (EM TERMOS DO NÚMERO DE ALUNOS NOS ANOS FINAIS DO EF)	IDEB DE ENTRADA DO ALUNO NO 6º ANO DO EF DA ESCOLA		
	ALTO (IDEB > 6,0)	MÉDIO (5,0 < IDEB <= 6,0)	BAIXO (IDEB <= 5,0)
Grande (mais de 50 mil)	515	2269	1044
Médio ou pequeno (50 mil ou menos)	4082	9217	6852

Fonte: Elaboração própria com base nos dados da Prova Brasil e Censo Escolar.

Em seguida, foram desenhados os critérios que definiram a seleção das escolas de baixo nível socioeconômico que mais se destacaram nos anos finais do ensino fundamental em 2015. Utilizou-se o nível socioeconômico médio dos alunos da escola para qualificar as escolas elegíveis para a análise, calculado pelos valores do Critério Brasil (ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA – ABEP, 2016). Foram estabelecidos três critérios para determinar se a escola seria considerada elegível: seleção das escolas com nível socioeconômico médio dos alunos nos níveis C2, D e E, de acordo com o Critério Brasil; seleção das 25% piores escolas em termos de nível socioeconômico de cada Unidade Federativa (UF); e seleção, em cada um dos 15 maiores municípios já apontados anteriormente, dos 25% de escolas com nível socioeconômico mais baixo.¹ Assim, se a escola atendeu a algum desses três critérios, ela entrou no conjunto de escolas passíveis de serem classificadas como de excelência.

1 No caso desses critérios, calculamos 25% de escolas com mais baixo nível socioeconômico utilizando dois ranqueamentos de escolas: um gerado pelos valores do Critério Brasil e outro construído com o Índice Socioeconômico das Escolas (Inse), produzido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) (BRASIL, 2015). Se a escola respeitasse o critério por algum desses ordenamentos, era considerada de baixo nível socioeconômico.

Os critérios utilizados para a seleção das escolas no caminho da excelência com equidade foram os seguintes:

- *Taxa de evasão escolar da escola entre o 6º e o 8º ano do ensino fundamental*: a ideia foi selecionar escolas com menos de 8% de evasão nessas séries, que corresponde à evasão média das escolas brasileiras. Assim, pretende-se destacar escolas que estejam conseguindo melhorar a proficiência dos alunos sem “expulsar” os piores alunos, gerando evasão escolar. Apesar de não ser uma variável de completo controle da escola, entende-se que uma escola de “sucesso” não deveria admitir elevadas taxas de evasão escolar, pois a evasão poderia atuar como um mecanismo de seleção dos melhores alunos.
- *Valor adicionado da nota*: por se tratar da segunda etapa do ensino fundamental, propõe-se uma medida de valor adicionado de Língua Portuguesa e Matemática como critério de seleção de escolas de excelência. A ideia é calcular a nota média de entrada dos alunos da escola no 6º ano do ensino fundamental, que deverá ser subtraída da nota média da escola na Prova Brasil quatro anos depois.
- *Taxa de participação de pelo menos 70% dos alunos na Prova Brasil (Matemática e Língua Portuguesa)*: esse critério é importante na medida em que um dos mecanismos para inflar artificialmente o Ideb pode ser provocar a ausência dos alunos com maiores dificuldades de aprendizado no dia da Prova Brasil.
- *Porcentagem máxima de alunos no nível insuficiente² de proficiência*: definida de acordo com o *cluster* ao qual a escola pertencia, variando entre 10% e 20%.
- *Porcentagem mínima de alunos no nível adequado³ de proficiência*: definida de acordo com o *cluster* ao qual a escola pertencia, variando entre 25% e 34% (Matemática) e 25% e 53% (Língua Portuguesa).
- *Ideb da escola*: a escola deveria estar pelo menos no percentil 75 de Ideb do seu *cluster* em 2015, variando de 3,9 a 5,4.
- *Município deve ter pelo menos 75% das escolas com evolução no Ideb entre 2011 e 2015*: a ideia é punir municípios que estejam concentrando “bons” alunos em determinadas escolas, aumentando a desigualdade entre escolas no tempo. A rede como um todo tem que ter evoluído positivamente no período 2011-2015.

A Tabela 2 apresenta os valores de corte para a definição das escolas de excelência:

2 Foi considerada nível insuficiente uma pontuação inferior a 225 pontos em Língua Portuguesa e a 250 pontos em Matemática.

3 Foi considerada nível adequado uma pontuação maior ou igual a 275 pontos em Língua Portuguesa e a 300 pontos em Matemática.

TABELA 2 – Valores de corte para os critérios de seleção das escolas em cada um dos seis *clusters* definidos pelo tamanho do município e Ideb de entrada no 6º ano da escola

CRITÉRIOS DE SELEÇÃO DAS ESCOLAS	TAMANHO DO MUNICÍPIO E IDEB DE ENTRADA NO 6º ANO DA ESCOLA					
	GRANDE			MÉDIO E PEQUENO		
	ALTO	MÉDIO	BAIXO	ALTO	MÉDIO	BAIXO
Valor adicionado da nota: Matemática	46	44	53	42	49	58
Valor adicionado da nota: Português	63	62	70	60	65	71
Taxa de evasão escolar: 6º a 8º ano	8%	8%	8%	8%	8%	8%
Taxa de participação na Prova Brasil 2015	70%	70%	70%	70%	70%	70%
% Max. no nível Insuficiente: Matemática	11%	20%	20%	10%	19%	20%
% Max. no nível Insuficiente: Português	10%	14%	19%	10%	11%	20%
% Mínima no nível adequado: Matemática	34%	25%	25%	33%	25%	25%
% Mínima no nível adequado: Português	53%	34%	25%	50%	35%	25%
Ideb em 2015	5,4	4,6	4,1	5,4	4,7	3,9
% Escolas com crescimento do Ideb	75%	75%	75%	75%	75%	75%

Fonte: Elaboração própria com base nos dados da Prova Brasil e do Censo Escolar.

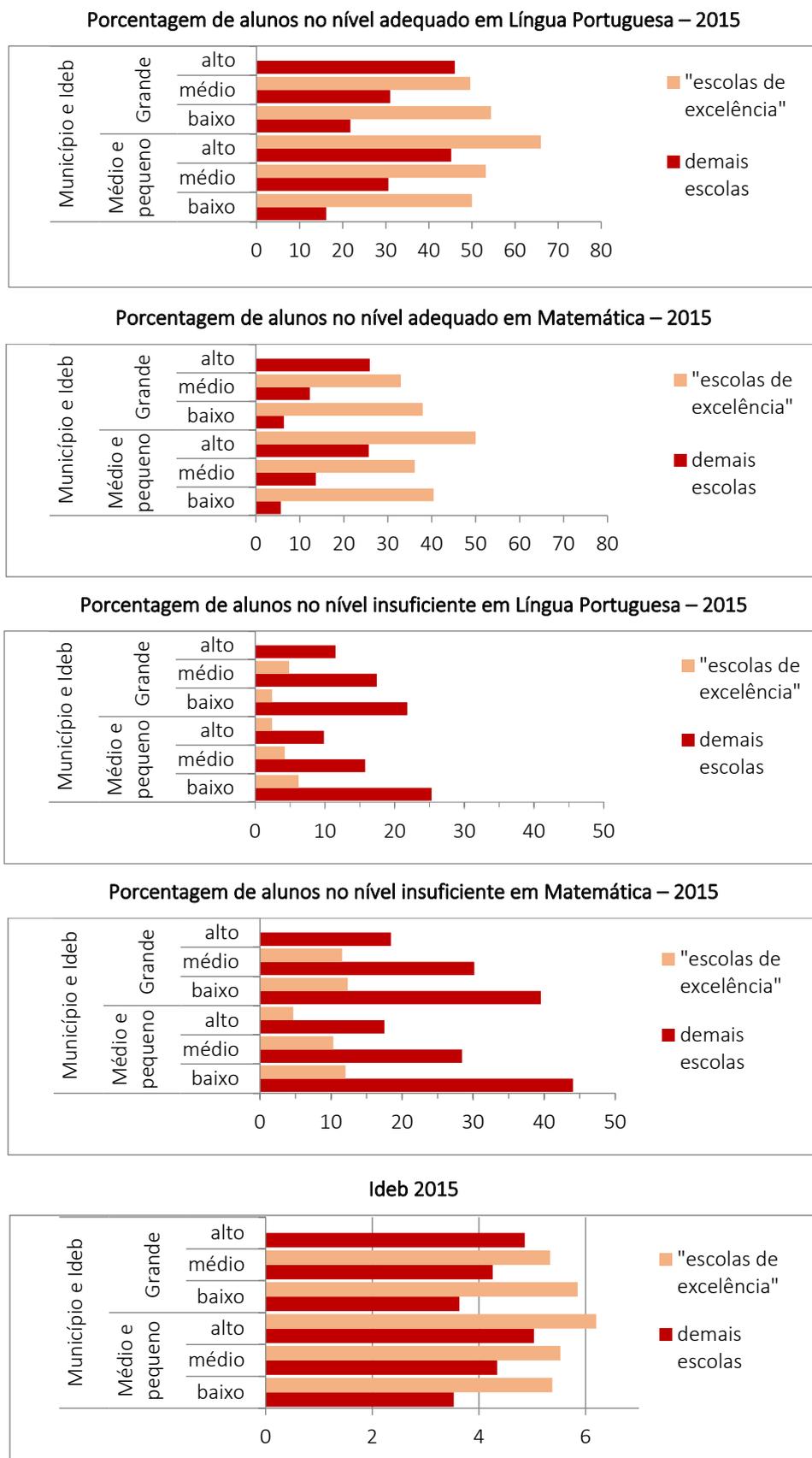
DADOS E ANÁLISE DESCRITIVA

As informações utilizadas para identificar as características e práticas das escolas de excelência são provenientes dos questionários das Provas Brasil. Nesse sentido, buscou-se identificar e analisar o maior número possível de variáveis que pudessem estar relacionadas ao fato de a escola ter atingido um *status* de excelência com equidade nos anos finais do ensino fundamental.

Os gráficos abaixo permitem uma breve análise descritiva de algumas variáveis, tanto para as escolas de excelência em 2015 quanto para as demais escolas, de acordo com o *cluster* ao qual pertencem. Para a porcentagem de alunos no nível adequado em Língua Portuguesa e em Matemática, é possível notar que as médias das escolas de excelência são sempre maiores que as médias das demais escolas. Por outro lado, para a porcentagem de alunos no nível insuficiente de proficiência em Língua Portuguesa e em Matemática, as médias das escolas de excelência são inferiores às médias das demais escolas. O que confirma um melhor desempenho das escolas de excelência.

Cabe destacar também que parte considerável das perguntas dos questionários de aluno, professor e diretor da Prova Brasil corresponde a variáveis categóricas. Ou seja, as alternativas de respostas são categóricas (como, por exemplo, “sempre”, “frequentemente”, “às vezes”, “quase nunca”, “nunca”). Nesses casos, foram criadas variáveis binárias que assumiam valor um ou zero, a depender da categoria respondida. No exemplo acima, atribuiu-se o valor 1 quando a resposta era “sempre”, e zero no caso das outras respostas.

FIGURA 1 - Comparação das escolas de excelência e demais escolas para variáveis da Prova Brasil, de acordo com o *cluster* a que pertencem



Fonte: Elaboração própria com base nos dados da Prova Brasil.

Para essas variáveis binárias, foram construídos alguns índices (0 a 1) que podem nos orientar sobre os próprios resultados para identificar os fatores que diferenciam as escolas com melhores desempenhos. A Figura 2 apresenta a evolução de alguns índices: alunos que fazem o dever de Língua Portuguesa e Matemática (questionário de aluno); expectativa dos professores sobre quantos alunos concluirão o ensino médio; agressão física ou verbal a alunos cometida por outros alunos (questionário de professores); e índice de falta dos alunos (questionário de diretores).

FIGURA 2 - Evolução dos índices da Prova Brasil para as escolas de excelência e demais escolas



Fonte: Elaboração própria com base nos dados da Prova Brasil.

Nesses casos, nota-se que nas escolas de excelência os índices são melhores (mais próximos de um, exceto no caso do índice de agressão, que deve ser mais próximo de zero) em relação às demais escolas.

METODOLOGIA

A fim de entender melhor as características e práticas que diferenciam as escolas de excelência das demais escolas, a estratégia de estimação adotada baseia-se no pareamento por score de propensão e diferenças em diferenças.⁴ A ideia do pareamento consistiu em procurar, para cada escola selecionada como de excelência em 2015, outra (de nível socioeconômico baixo) que acabou não se tornando de excelência, mas que fosse o mais semelhante possível em características observáveis à de excelência em 2011. No nosso caso, selecionamos as seguintes variáveis para tornar as escolas pareadas o mais parecidas possível: Ideb em 2011; tamanho do município em termos do número de alunos nos anos finais do ensino fundamental; Ideb de entrada no 6º ano; nível socioeconômico médio dos alunos da escola; e taxa de abandono do 8º e 9º anos em 2011. A escolha dessas variáveis visa a garantir maior probabilidade de que as escolas tivessem, inicialmente, condições semelhantes, seja em proficiência, seja em abandono e aprovação, além de estarem no mesmo nível socioeconômico e com o mesmo perfil de entrada dos alunos no 6º ano do ensino fundamental. O pareamento das escolas de excelência com as demais foi feito por meio do método do vizinho mais próximo.

Em seguida, acompanhamos esses dois grupos de escolas ao longo do tempo (até 2015) e observamos como evoluiu uma série de suas características. Por meio da aplicação da técnica de diferenças em diferenças, pretendemos saber quais características tiveram alteração ao longo do tempo de modo particular para as escolas de excelência relativamente ao outro grupo de escolas (FARIA; GUIMARÃES, 2015; GELMAN; IMBENS, 2013).

Como temos três períodos (2011, 2013 e 2015), a especificação do modelo permite estimar o que diferencia as escolas de excelência das demais nos dois últimos períodos. Assim, seguindo Galiani, Gertler e Schargrodsky (2005) e Oshiro, Scorzafave e Dorigan (2015):

$$Y_{it} = \beta_0 + \beta_1 Esc_Excel_i + \beta_2 I_{i2013} + \beta_3 I_{i2015} + \beta_4 d_{2013} + \beta_5 d_{2015} + \varepsilon_{it} \quad (2)$$

Em que Y_{it} refere-se aos diferentes indicadores sobre os quais vamos avaliar se há diferenças entre as escolas de excelência e as demais; $Esc_Excel_i = 1$, se a escola

4 Para mais detalhes sobre o método de pareamento e diferenças em diferenças, ver, por exemplo, Gertler *et al.* (2016).

i tiver sido classificada como de excelência e 0, caso contrário; I_{i2013} é uma variável que assume valor 1 para as escolas de excelência em 2013; I_{i2015} assume valor 1 para as escolas de excelência apenas em 2015; d_{2013} e d_{2015} são *dummies* de tempo para 2013 e 2015, respectivamente.

Os parâmetros β_2 e β_3 capturam o diferencial entre as escolas de excelência e as demais em termos de cada indicador analisado em 2013 e 2015, respectivamente. Assim, sendo esses coeficientes estatisticamente significativos, inferimos que tais diferenças podem revelar um dos aspectos que fizeram as escolas de excelência conseguir essa posição de destaque.

RESULTADOS

Com base nas perguntas das Provas Brasil, foram construídos índices que as agregavam de acordo com os seus conteúdos.⁵ A ideia de apresentar os índices é avaliar em quais dimensões as escolas de excelência se diferenciam das demais. Para este estudo, se apenas olhássemos perguntas da prova isoladamente, poderia não ficar evidente quais categorias de características estariam explicando dimensões que diferenciam as escolas de excelência das demais. Na Tabela 3 apresentam-se os resultados das estimações para os índices.

Diferenças que variam de 0,3 a 0,5 desvios-padrão nos índices capturam se os alunos fazem lição de casa e se os professores corrigem lição de casa, a favor das escolas de excelência, se comparadas às demais, em 2013 e 2015. Também se destacam a maior participação dos pais (0,389 d.p.) e o fato de os alunos gostarem de estudar (0,259 d.p.) em 2015. Além disso, os professores das escolas de excelência tiveram maior expectativa com relação ao futuro escolar dos alunos do que os docentes das demais escolas (0,446 d.p.), além de reconhecerem o papel de liderança dos diretores em suas escolas e identificarem menor violência nas escolas de excelência. Nos indicadores relacionados à resposta dada pelos diretores, não houve diferença entre os dois tipos de escolas. No caso específico do indicador de violência, as escolas de excelência já eram menos violentas em 2011 do que as demais, segundo a percepção dos diretores (-0.25 d.p.).⁶

5 Os índices foram construídos via teoria de resposta ao item, a partir de itens da Prova Brasil.

6 Resultado não reportado na Tabela 3.

TABELA 3 – Diferenças entre “escolas de excelência” e as demais escolas de baixo nível socioeconômico – Índices sintéticos

ÍNDICES	2013		2015	
	COEFICIENTE	E. PADRÃO	COEFICIENTE	E. PADRÃO
Alunos				
Gosta de estudar	0,180	0,122	0,259**	0,120
Faz lição de casa	0,292**	0,124	0,523***	0,128
Professor corrige lição	0,309***	0,116	0,456***	0,120
Participação dos pais	0,169	0,128	0,389***	0,130
Professor				
Liderança do diretor	0,286**	0,116	0,233**	0,116
Coesão do corpo docente	0,371***	0,128	0,199	0,134
Expectativa do professor em relação aos alunos	0,193	0,120	0,446***	0,118
Violência	-0,265**	0,117	-0,372***	0,118
Diretor				
Recursos financeiros	0,000*	0,000	0,000	0,000
Participação dos pais na escola	-0,018	0,085	-0,018	0,085
Violência	0,169	0,132	-0,002	0,126

Fonte: Elaboração própria com base nos dados da Prova Brasil.

Nota: *** significativo a 1%; ** significativo a 5%; * significativo a 10%.

Obs.: Cada linha da tabela corresponde ao resultado da estimação de um modelo de diferenças em diferenças em que a variável dependente é um indicador formado a partir de itens da Prova Brasil. Os coeficientes na primeira e segunda colunas são, respectivamente, estimativas de β_2 e β_3 da equação (2). No Apêndice, são apresentadas as variáveis utilizadas para o cálculo de cada um dos índices sintéticos.

A análise dos índices sintéticos, no entanto, não permite a identificação das práticas concretas que podem diferenciar as escolas de excelência das demais. Nesse sentido, passa-se agora a analisar a resposta a itens isolados do questionário da Prova Brasil, cujos resultados encontram-se na Tabela 4. Vale ressaltar que, no caso da Tabela 4, as variáveis dependentes são binárias e os coeficientes indicam qual a diferença entre escolas de excelência e as demais (em pontos percentuais) em cada uma das variáveis em análise.

TABELA 4 - Diferenças entre escolas de excelência e as demais escolas de baixo nível socioeconômico - Itens da Prova Brasil

VARIÁVEIS	2013		2015	
	COEFICIENTE	E. PADRÃO	COEFICIENTE	E. PADRÃO
Alunos				
Gosta de estudar Matemática	0,023	0,0164	0,044***	0,0164
Faz dever de Português	0,042**	0,0169	0,055***	0,0169
Faz dever de Matemática	0,032*	0,0171	0,063***	0,0168
Prof. corrige dever de Português	0,038***	0,0145	0,042***	0,0148
Prof. corrige dever de Matemática	0,031**	0,0144	0,047***	0,0140
Frequência que os pais participam da reunião	0,023	0,0190	0,048**	0,0188
Pais que incentivam os alunos a estudar	0,004	0,0028	0,006**	0,0025
Pais que incentivam os alunos a fazer dever	0,009*	0,0053	0,011**	0,0049
Pais que incentivam os alunos a não faltar	0,003	0,0032	0,006*	0,0030
Aluno já reprovou	0,000	0,0186	-0,056**	0,0170
Professor				
Indisciplina dos alunos em sala de aula	-0,068	0,0450	-0,074*	0,0444
Quanto dos alunos você acha que concluirão o ensino fundamental	0,047*	0,0275	0,052**	0,0237
Quanto dos alunos você acha que concluirão o ensino médio	0,050	0,0462	0,136***	0,0454
Quanto dos alunos você acha que entrarão para a universidade	0,030	0,0281	0,058**	0,0295
Agressão física ou verbal a funcionário e/ou professores cometida por alunos	-0,093*	0,0477	-0,102**	0,0477
Agressão física ou verbal a alunos cometida por outros alunos	-0,099**	0,0460	-0,182***	0,0478
Diretor				
Diretor dá atenção especial a aspectos relacionados a aprendizagem dos alunos	0,051*	0,0300	0,053*	0,0306
Critério de formação de turmas: homogeneidade de rendimento escolar	0,032	0,0409	0,070**	0,0340
Há carência de pessoal administrativo	-0,139**	0,0679	-0,075	0,0622
Alto índice de falta dos alunos	-0,048	0,0672	-0,104*	0,0611

Fonte: Elaboração própria com base nos dados da Prova Brasil.

Nota: *** significativo a 1%; ** significativo a 5%; * significativo a 10%.

Obs.: Cada linha da tabela corresponde ao resultado da estimação de um modelo de diferenças em diferenças em que a variável dependente corresponde a itens da Prova Brasil. Os coeficientes na primeira e segunda colunas são estimativas de β_2 e β_3 da equação (2), respectivamente.

A importância de fazer o dever de casa e o fato de o professor corrigi-lo aparecem, tanto para Língua Portuguesa quanto para Matemática, como fatores que diferenciam as escolas de excelência. Na percepção dos discentes, em 2015, os pais dos alunos de escolas de excelência incentivavam seus filhos a fazerem o dever de casa, a estudar e a não faltar à aula. Por fim, um dado interessante é que, nas escolas de excelência, a proporção de alunos que já haviam sido reprovados também era menor. Em geral, a diferença entre as escolas de excelência e as demais é de cerca de cinco pontos percentuais nesses itens, exceto no caso do incentivo dos pais, em que é praticamente nula a diferença.

Com relação aos itens respondidos pelo professor, destacam-se a menor indisciplina dos alunos nas escolas de excelência e uma maior expectativa quanto ao futuro escolar dos alunos. Por exemplo, há diferença de 13 pontos percentuais a favor das escolas de excelência na expectativa dos professores de que os alunos concluiriam o ensino médio. Por fim, os professores relatam menos violência na escola, merecendo destaque uma diferença de 18 pontos percentuais na ocorrência de agressão física e verbal entre alunos, quando se comparam as escolas de excelência com as demais. Ou seja, aspectos relacionados a habilidades interpessoais parecem ser mais bem trabalhados nas escolas de bom desempenho do que nas demais, sob a ótica dos professores, corroborando parte da literatura, que aponta a importância de um bom ambiente na sala de aula para garantia do aprendizado.

Por fim, os diretores relatam menos problemas com faltas de alunos e carência de pessoal administrativo. Um último ponto interessante é que há maior proporção de escolas de excelência que formam turmas separando os alunos por homogeneidade de desempenho. De fato, apesar de esse assunto ser polêmico, a literatura sobre efeitos dos pares vem demonstrando que conformações de turmas dessa natureza, em que são alocados alunos de níveis semelhantes de aprendizado, possuem efeitos benéficos (HOXBY; WEINGARTH, 2005).

Os resultados aqui encontrados também corroboram a literatura no que se refere ao fato de que as expectativas dos professores sobre os alunos afetam o aprendizado (SOARES *et al.*, 2010) e um ambiente violento dentro da escola está associado a menor nível de proficiência escolar (SEVERNINI; FIRPO, 2009; TEIXEIRA; KASSOUF, 2015). Além disso, a literatura já aponta que fazer lição de casa explica a diferença de desempenho escolar entre os alunos (SCORZAFAVE; FERREIRA, 2011; MENEZES FILHO; RIBEIRO, 2009). Por fim, a liderança do diretor, outro fator importante de distinção das escolas de excelência, também corrobora a literatura (OLIVEIRA; CARVALHO, 2018).

CONSIDERAÇÕES FINAIS

Neste artigo procurou-se caracterizar as escolas que vêm apresentando destaque nos anos finais do ensino fundamental, mesmo tendo alunos de baixo *background* socioeconômico. Após o estabelecimento dos critérios e a seleção das denominadas escolas de excelência, buscou-se identificar em que dimensões essas escolas se diferenciam das demais. De forma geral, as relações interpessoais se destacam nessas escolas, com menores índices de violência, maior crença dos professores no potencial de estudo futuro dos alunos e menor indisciplina dos alunos. Um ponto importante para futura investigação é entender como essas escolas conseguem construir um ambiente seguro e lidar com os problemas de violência, tanto do seu entorno (que, muitas vezes, se refletem no ambiente escolar) como dentro da escola. Investigar como se constrói essa situação benéfica e como se previne a violência escolar é o próximo passo nesta agenda de pesquisa.

Esses fatores acabam sendo complementados por um ambiente em que o aprendizado é levado a sério, o que se reflete no maior índice de alunos que afirmaram fazer as lições de casa, bem como no maior comprometimento dos professores com a correção dessas tarefas. Também há menos relatos de faltas dos alunos nessas escolas do que nas demais. Tais fatores, de modo geral, costumam ser apontados como associados a um melhor desempenho escolar quando são feitas análises para as escolas como um todo, sem a realização de recorte para escolas com alunos de mais baixo *background* socioeconômico.

Por sua vez, fatores normalmente considerados para atribuir qualidade a uma escola não mostraram diferenças entre as escolas analisadas, em particular os salários de professores, bem como a qualificação e experiência de docentes e diretores, que se mostraram semelhantes tanto nas escolas de excelência como nas demais. Note-se que não se pretende dizer, com isso, que tais fatores não sejam importantes para compreendermos a qualidade das escolas, mas simplesmente que não são fatores que diferenciam as escolas de melhor desempenho das demais. Portanto, não parece ter sido por meio desses canais que as escolas conseguiram se destacar em relação às demais.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA – ABEP. *Critério de classificação econômica Brasil 2015*. São Paulo: ABEP, 2016.

BATISTA, Antônio Augusto Gomes; MELLO, Hivy Damasio Araújo; FREITAS, Pâmela Félix; RIBEIRO, Vanda Mendes (coord.). Currículos para os anos finais do ensino fundamental: concepções, modos de implantação e usos. Relatório final. *Estudos & Pesquisas Educacionais*, São Paulo, v. 5, p. 15-71, 2015.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep. *Indicador de nível socioeconômico (Inse) das escolas*. Nota técnica. Brasília, DF, 2015.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep. *Indicadores educacionais*. Brasília, DF, 2018. Disponível em: <http://inep.gov.br/web/guest/indicadores-educacionais>. Acesso em: 24 maio 2020.

CARLOS, Vitor Augusto. *O papel da repetência escolar sobre variáveis de fluxo: uma análise sobre o abandono escolar e chegada ao ensino médio*. 2019. Dissertação (Mestrado em Economia Aplicada) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, 2019.

DAVIS, Claudia Leme Ferreira; TARTUCE, Gisela Lobo B. P.; NUNES, Marina M. Rossa.; ALMEIDA, Patrícia C. Albieri de; SILVA, Ana Paula F. da; COSTA, Beatriz S. D. de Olival; SOUZA, Juliana Cedro de. Anos finais do ensino fundamental: aproximando-se da configuração atual. *Estudos & Pesquisas Educacionais*, São Paulo, v. 3, p. 103-194, 2012.

FARIA, Ernesto Martins; GUIMARÃES, Raquel Rangel. Excelência com equidade: fatores escolares para o sucesso educacional em circunstâncias desfavoráveis. *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 61, p. 192-215, jan./abr. 2015.

FUNDAÇÃO BRAVA; INSTITUTO UNIBANCO; INSPER; INSTITUTO AYRTON SENNA. *Políticas públicas para redução do abandono e evasão escolar de jovens*. Relatório final. [S.l.], 2017.

FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA – UNICEF. *Aprova Brasil: o direito de aprender – Boas práticas em escolas públicas avaliadas pela Prova Brasil*. Brasília, DF, 2007.

FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA – UNICEF. *Redes de aprendizagem: boas práticas de municípios que garantem o direito de aprender*. Brasília, DF, 2009.

FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA – UNICEF. *Caminhos do direito de aprender: boas práticas de 26 municípios que melhoraram a qualidade da educação*. Brasília, DF, 2010.

GALIANI, Sebastian; GERTLER, Paul; SCHARGRODSKY, Ernesto. Water for life: the impact of the privatization of water services on child mortality. *Journal of Political Economy*, v. 113, n. 1, p. 83-120, 2005.

GREMAUD, Amaury Patrick; NICOLELLA, Alexandre C.; SCORZAFAVE, Luiz Guilherme; OLIVEIRA, Roberto Guena. *Relação entre abandono escolar no ensino médio e desempenho escolar no ensino fundamental brasileiro*. São Paulo: Instituto Unibanco, 2010. Disponível em: <https://www.slideshare.net/institutounibanco/relao-entre-abandono-escolar-no-ensino-mdio-e-desempenho-escolar-no-ensino-fundamental-brasileiro>. Acesso em: 11 out. 2019.

GELMAN, Andrew; IMBENS, Guido. *Why ask why? Forward causal inference and reverse causal questions*. Cambridge, MA: National Bureau of Economic Research, 2013. (NBER Working Papers).

GERTLER, Paul J.; MARTINEZ, Sebastian; PREMAMAND, Patrick; RAWLINGS, Laura B.; VERMEERSCH, Christel M. J. *Impact evaluation in practice*. New York: The World Bank, 2016.

HOXBY, Caroline M.; WEINGARTH, Gretchen. *Taking race out of the equation: school reassignment and the structure of peer effects*. 2005. Unpublished manuscript. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.75.4661&rep=rep1&type=pdf>. Acesso em: 14 fev. 2018.

MADALOZZO, Regina Carla; FARIA, Ernesto Martins. Excelência com equidade: as lições das escolas brasileiras que oferecem educação de qualidade a alunos de baixo nível socioeconômico. *Economia e Políticas Públicas*, Montes Claros, MG, v. 2, n. 1, p. 87-108, 2014.

MANSUTTI, Maria Amabile; ZELMANOVITS, Maria Cristina; CARVALHO, Maria do Carmo Brant de; GURIDI, Verónica. Especial: estudo Cenpec Educação na segunda etapa do ensino fundamental. *Cadernos Cenpec | Nova série*, São Paulo, v. 2, n. 4, p. 1, 2007.

MENEZES FILHO, Naercio Aquino. Os determinantes do desempenho escolar do Brasil. In: DUARTE, Pedro; SILBER, Simão; GUILHOTO, Joaquim (org.). *O Brasil e a ciência econômica em debate: o Brasil do século XXI*. São Paulo: Saraiva, 2012. v. 1, p. 231-256.

MENEZES FILHO, Naercio; RIBEIRO, Fernanda. Comparando as escolas paulistas com melhor e pior desempenho no Saresp e na Prova Brasil. *São Paulo em Perspectiva*, São Paulo, v. 23, n. 1, p. 115-134, jan./jun. 2009.

NERI, Marcelo (coord.). *Motivos da evasão escolar*. Rio de Janeiro: Ibre/FGV, 2009.

OLIVEIRA, Ana Cristina Prado de; CARVALHO, Cynthia Paes de. Gestão escolar, liderança do diretor e resultados educacionais no Brasil. *Revista Brasileira de Educação*, Rio de Janeiro, v. 23, p. 1-18, 2018.

OSHIRO, Cláudia Hiromi; SCORZAFAVE, Luiz Guilherme; DORIGAN, Tulio Anselmi. Impacto sobre o desempenho escolar do pagamento de bônus aos docentes do ensino fundamental do Estado de São Paulo. *Revista Brasileira de Economia*, Rio de Janeiro, v. 69, n. 2, p. 213-249, abr./jun. 2015.

PADILHA, Frederica; ÉRNICA, Mauricio; BATISTA, Antônio A. Gomes; PUDENZI, Luciana. As regularidades e exceções no desempenho no Ideb dos municípios. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 51, p. 58-81, jan./abr. 2012.

SANTOS, Daniel; SCORZAFAVE, Luiz; NICOLELLA, Alexandre C.; SANT'ANNA, Elder. Mais é menos? O impacto do Projeto 6º Ano Experimental – SME/RJ. *Estudos em Avaliação Educacional*, São Paulo, v. 28, n. 69, p. 718-747, set./dez. 2017.

SCORZAFAVE, Luiz Guilherme; FERREIRA, Rodrigo Araujo. Desigualdade de proficiência no ensino fundamental público brasileiro: uma análise de decomposição. *Economia*, Brasília, DF, v. 12, n. 2, p. 337-359, maio/ago. 2011.

SEVERNINI, Edson; FIRPO, Sergio. *The relationship between school violence and student proficiency*. São Paulo: Escola de Economia de São Paulo/FGV-EESP, 2009. (Texto para discussão, n. 236).

SOARES, Jose Francisco; ANDRADE, Renato Júdice. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 14, n. 50, p. 107-125, jan./mar. 2006.

SOARES, Tufi Machado; FERNANDES, Neimar da Silva; FERRAZ, Mariana Santos Botarro; RIANI, Juliana. A expectativa do professor e o desempenho dos alunos. *Psicologia: Teoria e Pesquisa*, Brasília, DF, v. 26, n. 1, p. 157-170, jan./mar. 2010.

SOUZA, André Portela; PONCZEK, Vladimir Pinheiro; OLIVA, Bruno Teodoro; TAVARES, Priscilla Albuquerque. Fatores associados ao fluxo escolar no ingresso e ao longo do Ensino Médio no Brasil. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 42, n. 1, p. 5-39, abr. 2012.

TEIXEIRA, Evandro Camargos; KASSOUF, Ana Lúcia. Impacto da violência nas escolas paulistas sobre o desempenho acadêmico dos alunos. *Economia Aplicada*, Ribeirão Preto, SP, v. 19, n. 2, p. 221-240, abr./jun. 2015.

TODOS PELA EDUCAÇÃO. *Equidade entre as escolas das redes municipais: uma análise com base nos dados de Ideb para municípios e escolas*. São Paulo, 2010.

APÊNDICE

QUADRO A1 – Variáveis utilizadas para o cálculo dos índices sintéticos – Alunos

ÍNDICE	ITENS DA PROVA BRASIL	RESPOSTA DA PROVA BRASIL
Gosta de estudar	• Você gosta de estudar Língua Portuguesa?	1, se “sim”; 0, se “não”
	• Você gosta de estudar Matemática?	1, se “sim”; 0, se “não”
Faz lição de casa	• Você faz o dever de casa de Língua Portuguesa?	1, se “sempre ou quase sempre”; 0, se “de vez em quando”, “nunca ou quase nunca” ou “o(a) professor(a) não passa dever de casa”
	• Você faz o dever de casa de Matemática?	1, se “sempre ou quase sempre”; 0, se “de vez em quando”, “nunca ou quase nunca” ou “o(a) professor(a) não passa dever de casa”
Professor corrige lição	• O(a) professor(a) corrige o dever de casa de Língua Portuguesa?	1, se “sempre ou quase sempre”; 0, se “de vez em quando”, “nunca ou quase nunca” ou “o(a) professor(a) não passa dever de casa”
	• O(a) professor(a) corrige o dever de casa de Matemática?	1, se “sempre ou quase sempre”; 0, se “de vez em quando”, “nunca ou quase nunca” ou “o(a) professor(a) não passa dever de casa”
Participação dos pais	• Com qual frequência seus pais, ou responsáveis por você, vão à reunião de pais?	1, se “sempre ou quase sempre”; 0, se “de vez em quando” ou “nunca ou quase nunca”
	• Seus pais ou responsáveis incentivam você a estudar?	1, se “sim”; 0, se “não”
	• Seus pais ou responsáveis incentivam você a fazer o dever de casa e/ou os trabalhos da escola?	1, se “sim”; 0, se “não”
	• Seus pais ou responsáveis incentivam você a ler?	1, se “sim”; 0, se “não”
	• Seus pais ou responsáveis incentivam você a ir à escola e/ou não faltar às aulas?	1, se “sim”; 0, se “não”
	• Seus pais ou responsáveis conversam com você sobre o que acontece na escola?	1, se “sim”; 0, se “não”

Fonte: Elaboração própria com base nos dados da Prova Brasil.

QUADRO A2 - Variáveis utilizadas para o cálculo dos índices sintéticos - Professores

ÍNDICE	ITENS DA PROVA BRASIL	RESPOSTA DA PROVA BRASIL
Liderança do diretor	• O(a) diretor(a) me anima e me motiva para o trabalho.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• Tenho confiança no(a) diretor(a) como profissional.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(a) diretor(a) estimula atividades inovadoras.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(a) diretor(a) dá atenção especial a aspectos relacionados com a aprendizagem dos alunos.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(a) diretor(a) dá atenção especial a aspectos relacionados com as normas administrativas.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(a) diretor(a) dá atenção especial a aspectos relacionados com a manutenção da escola.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(a) diretor(a) discute metas educacionais com os professores nas reuniões.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• O(A) diretor(a) e os professores procuram assegurar que as questões de qualidade de ensino sejam uma responsabilidade coletiva.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
Coesão do corpo docente	• Participo das decisões relacionadas com meu trabalho.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
	• A equipe de professores leva em consideração minhas ideias.	1, se “sempre ou quase sempre” ou “frequentemente”; 0, se “algumas vezes” ou “nunca”
Expectativas dos professores em relação aos alunos	• Na sua opinião, quantos dos alunos desta turma você acha que concluirão o ensino fundamental?	1, se “quase todos os alunos”; 0, se “pouco mais da metade”, “pouco menos da metade” ou “poucos”
	• Quantos dos alunos desta turma você acha que concluirão o ensino médio?	1, se “quase todos os alunos”; 0, se “pouco mais da metade”, “pouco menos da metade” ou “poucos”
	• Quantos dos alunos desta turma você acha que entrarão na universidade?	1, se “quase todos os alunos”; 0, se “pouco mais da metade”, “pouco menos da metade” ou “poucos”

(Continua)

ÍNDICE	ITENS DA PROVA BRASIL	RESPOSTA DA PROVA BRASIL
Violência	• Agressão verbal ou física de alunos a professores ou a funcionários da escola.	1, se “sim”; 0, se “não”
	• Agressão verbal ou física de alunos a outros alunos da escola.	1, se “sim”; 0, se “não”
	• Você foi vítima de atentado à vida.	1, se “sim”; 0, se “não”
	• Você foi ameaçado por algum aluno.	1, se “sim”; 0, se “não”
	• Você foi vítima de furto (sem uso de violência).	1, se “sim”; 0, se “não”
	• Você foi vítima de roubo (com uso de violência).	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas sob efeito de bebida alcoólica.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas sob efeito de drogas ilícitas.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas portando arma branca.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas portando arma de fogo.	1, se “sim”; 0, se “não”

Fonte: Elaboração própria com base nos dados da Prova Brasil.

QUADRO A3 – Variáveis utilizadas para o cálculo dos índices sintéticos – Diretores

ÍNDICE	ITENS DA PROVA BRASIL	RESPOSTA DA PROVA BRASIL
Recursos financeiros	• Esta escola recebeu neste ano apoio financeiro do governo federal?	1, se “sim”; 0, se “não”
	• Esta escola recebeu neste ano apoio financeiro do governo estadual?	1, se “sim”; 0, se “não”
	• Esta escola recebeu neste ano apoio financeiro do governo municipal?	1, se “sim”; 0, se “não”
Participação dos pais no conselho escolar	• Neste ano, quantas vezes se reuniu o Conselho Escolar?	1, se “três vezes ou mais”; 0, se “duas vezes”, “uma vez”, “nenhuma” ou “não existe conselho”
	• Além de você, quem participa do Conselho Escolar?	1, se “pais participam”; 0, se “pais não participam” ou “não existe conselho”
Clima escolar	Agressão verbal ou física de alunos a professores ou a funcionários da escola.	1, se “sim”; 0, se “não”
	Agressão verbal ou física de alunos a outros alunos da escola.	1, se “sim”; 0, se “não”
	Você foi vítima de atentado à vida.	1, se “sim”; 0, se “não”
	• Você foi vítima de furto (sem uso de violência).	1, se “sim”; 0, se “não”
	• Você foi vítima de roubo (com uso de violência).	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas sob efeito de bebida alcoólica.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas sob efeito de drogas ilícitas.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas portando arma branca.	1, se “sim”; 0, se “não”
	• Alunos frequentaram as suas aulas portando arma de fogo.	1, se “sim”; 0, se “não”

Fonte: Elaboração própria com base nos dados da Prova Brasil.

Recebido em: 23 OUTUBRO 2019

Aprovado para publicação em: 17 SETEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7452>

EVIDÊNCIAS EMPÍRICAS DO IMPACTO DA ALFABETIZAÇÃO SOBRE O CRESCIMENTO ECONÔMICO

 VINÍCIUS DE AZEVEDO COUTO FIRME¹

¹ Universidade Federal de Juiz de Fora (UFJF), *campus* GV, Governador Valadares-MG, Brasil; vinicius.firme@ufjf.edu.br

RESUMO

Esta pesquisa usou uma versão espacial do modelo proposto por Mankiw, Romer e Weil [MRW (1992)] para avaliar o impacto da alfabetização sobre a renda per capita dos municípios brasileiros no período de 1980 a 2010. Os resultados indicam que o efeito da alfabetização é maior em regiões mais homogêneas e transborda para a vizinhança. Em média, um crescimento de 1% na taxa de alfabetização faria a renda crescer 1,684%. Portanto, o produto interno bruto (PIB) brasileiro poderia ser 15,91% maior caso o Brasil atingisse a taxa de alfabetização europeia. Alternativamente, Alagoas, Piauí, Paraíba e Maranhão (e outros dez estados) poderiam aumentar suas respectivas rendas em 32,77%, 29,09%, 26,53% e 23,97%, caso atingissem a média nacional de alfabetização.

PALAVRAS-CHAVE EDUCAÇÃO • CRESCIMENTO ECONÔMICO • TAXA DE ALFABETIZAÇÃO.

EVIDENCIAS EMPÍRICAS DEL IMPACTO DE LA ALFABETIZACIÓN SOBRE EL CRECIMIENTO ECONÓMICO

RESUMEN

Esta investigación utilizó una versión espacial del modelo propuesto por Mankiw, Romer y Weil [MRW (1992)] para evaluar el impacto de la alfabetización sobre los ingresos per capita de los municipios brasileños en el periodo de 1980 a 2010. Los resultados indican que el efecto de la alfabetización es mayor en regiones más homogéneas y transborda hacia la vecindad. En promedio, un crecimiento del 1% en la tasa de alfabetización haría que los ingresos aumentaran en 1,684%. Por lo tanto, el producto interno bruto (PIB) brasileño podría ser 15,91% mayor en el caso que alcanzase el nivel de la alfabetización europea. Alternativamente, Alagoas, Piauí, Paraíba y Maranhão (así como otros 10 estados) podrían aumentar sus respectivos ingresos en un 32,77, 29,09, 26,53 e 23,97%, si alcanzasen el promedio nacional de alfabetización.

PALABRAS CLAVE EDUCACIÓN • CRECIMIENTO ECONÓMICO • TASA DE ALFABETIZACIÓN.

EMPIRICAL EVIDENCE OF THE IMPACT OF LITERACY ON ECONOMIC GROWTH

ABSTRACT

This research used a spatial version of the model developed by Mankiw, Romer and Weil [MRW (1992)] to assess the literacy impact on the per capita income of Brazilian municipalities between 1980 and 2010. The results show that the effect of literacy is greater in regions that are more homogeneous and it spillover to the neighborhood. On average, a 1% increase in literacy rate would increase income by 1.684%. Therefore, the Brazilian gross domestic product (GDP) could be 15.91% higher if it reached the level of European literacy. On the other hand, Alagoas, Piauí, Paraíba and Maranhão (and 10 other states) could increase their respective incomes by 32.77%, 29.09%, 26.53% and 23.97%, respectively, if they reached the national literacy average.

KEYWORDS EDUCATION • ECONOMIC GROWTH • LITERACY RATE.

INTRODUÇÃO

Após Solow (1956) concluir que o crescimento econômico *per capita* dependia, basicamente, do avanço tecnológico, a literatura econômica presenciou o surgimento de diversos modelos de crescimento endógeno. Tais teorias sustentam que as externalidades positivas oriundas da inovação (ROMER, 1990; GROSSMAN; HELPMAN, 1991; AGHION; HOWITT, 1992), da experiência profissional (ARROW, 1962; LUCAS, 1988; YANG; BORLAND, 1991) e do capital humano (ROMER, 1986; LUCAS, 1998; BARRO, 1991; MANKIW; ROMER; WEIL, 1992) seriam os principais responsáveis pelo crescimento de longo prazo.

O “capital humano”, difundido na década de 1960 por Schultz (1961, 1962) e Becker (1964), inclui o conjunto de conhecimentos, habilidades e atitudes que favorecem a produtividade do trabalhador e podem ser adquiridos por meio da experiência, da educação e de melhores condições de saúde (NEVES; LIMA, 2019, p. 78). Para Krueger (1968), o capital humano poderia explicar mais de 50% do crescimento econômico.

Assim, com o propósito de analisar o efeito da educação básica (*i.e.*, alfabetização) sobre o crescimento econômico dos municípios brasileiros no período de 1980 a 2010, a presente pesquisa focou na questão do capital humano e valeu-se do modelo proposto por Mankiw, Romer e Weil [MRW (1992)], que incorpora tal variável explicitamente e foi capaz de explicar, aproximadamente, 80% do crescimento de diversos países (ROMER, 1996, p. 128).

Embora bem-sucedida, a versão de Mankiw, Romer e Weil (1992) desconsiderou os possíveis efeitos da interação espacial sobre o crescimento econômico,¹ destacados a seguir:

Os efeitos espaciais, particularmente a autocorrelação e a heterogeneidade espaciais, devem ser considerados ao analisar processos de convergência em escala regional. Há uma série de fatores – comércio entre regiões, difusão de tecnologia e conhecimento e, de modo mais geral, os transbordamentos regionais – que levam a regiões geograficamente dependentes. Devido às interações espaciais, a localização geográfica é importante na determinação do desempenho econômico das regiões.² (GALLO; ERTUR, 2003, p. 176, tradução nossa)

1 Almeida (2012) e Kelejian e Piras (2017) ressaltam que ignorar a heterogeneidade e a dependência espacial poderia induzir a estimativas inconsistentes e/ou ineficientes.

2 No original: “*Spatial effects, particularly spatial autocorrelation and spatial heterogeneity, must be taken into account when analyzing convergence processes at regional scale. There are a number of factors – trade between regions, technology and knowledge diffusion and more generally regional spillovers – that lead to geographically dependent regions. Because of spatial interactions between regions, geographical location is important in accounting for the economic performances of regions.*”

Segundo Almeida (2012), as características distintas de cada região (*i.e.*, cultura, preferências, relevo, clima, entre outros) poderiam fazer com que um mesmo estímulo causasse diferentes impactos, dependendo da localidade (problema de heterogeneidade espacial). Além disso, a conexão entre uma região e suas vizinhas poderia produzir transbordamentos e autocorrelação espaciais (problema de dependência espacial), atenuando ou potencializando o efeito de um choque exógeno qualquer. Portanto, uma versão espacial do modelo MRW (1992), com dados em painel para os períodos de 1980, 1991, 2000 e 2010, foi considerada a fim de mensurar o efeito da alfabetização sobre a renda *per capita* dos municípios brasileiros.

Cabe destacar que o capital humano tem caráter multidimensional e não é composto apenas pela alfabetização. Contudo, a desagregação espacial (municípios) e o período considerados (1980-2010), apesar de suas vantagens, impõem algumas restrições à presente pesquisa. No que se refere às vantagens, tem-se que a escala geográfica desagregada favorece o controle dos efeitos espaciais baseados na proximidade regional (ALMEIDA, 2012, cap. 3) e permite incluir mais observações, favorecendo as propriedades assintóticas dos estimadores (HSIAO, 2003). Quanto ao período, ressalta-se que a teoria do crescimento econômico se baseia em um “estado estacionário” válido para o longo prazo (seção “O modelo teórico”). Portanto, os efeitos esperados tornam-se menos prováveis em curtos períodos de tempo. Apesar das vantagens, essas opções de escala e tempo estão sujeitas à ausência de informações (*missings*) em determinados períodos (ver Figuras A.1, A.2 e A.3 – Anexo) e reduzem a oferta de *proxies* para as variáveis testadas (FIRME; SIMÃO FILHO, 2014, p. 687-689), impossibilitando que outros elementos associados ao capital físico e humano fossem incluídos.

Portanto, é possível que o impacto do capital humano, medido via taxa de alfabetização, fique subestimado, pois fatores como experiência profissional e níveis mais avançados de ensino não foram considerados. Alternativamente, o efeito da alfabetização pode ficar superestimado, visto que essa variável pode captar outras características não consideradas nos modelos (*e.g.*, é possível que cidades com maiores taxas de alfabetização também tenham mais indivíduos graduados. Desse modo, parte do efeito atribuído a “graduados” ficaria embutido na alfabetização). Ainda assim, destaca-se que existem poucos trabalhos empíricos sobre os impactos da alfabetização na economia brasileira, abrindo espaço para que políticas arbitrárias, mal planejadas e, possivelmente, ineficientes sejam utilizadas para mitigar esse problema.

Ainda que o efeito da alfabetização sobre o crescimento econômico já tenha sido analisado em alguns estudos nacionais (RENZI *et al.*, 2019; ALMEIDA; VALADARES; SEDIYAMA, 2017; SOARES, 2009; VIEIRA, 2009) e internacionais (LEIJA, 2019; KHALFAOUI, 2015; MURTHY; OKUNADE, 2014; CASTELLÓ-CLIMENT;

MUKHOPADHYAY, 2013; ADELAKUN, 2011; GUSTAFSSON *et al.*, 2010; LIN, 1997; BARRO, 1991; AZARIADIS; DRAZEN, 1990; ROMER, 1989; RAUCH, 1988), inclusive com controles espaciais (RENZI *et al.*, 2019; LEIJA, 2019; VIEIRA, 2009), a presente pesquisa se destaca por: a) ser a única aplicada aos municípios de todas as regiões brasileiras;³ b) incluir variáveis de controle baseadas em modelos macroeconômicos reconhecidos pela literatura (vide “O modelo teórico”); c) buscar o controle simultâneo da dependência e heterogeneidade espaciais;⁴ d) considerar um amplo período para análises sobre o crescimento econômico que tem sua teoria voltada para o longo prazo;⁵ e) promover simulações, com base nos resultados obtidos, que permitem visualizar os possíveis efeitos benéficos da alfabetização no território brasileiro.

Desse modo, cientes das vantagens e desvantagens do procedimento adotado nesta pesquisa, os resultados indicam que o efeito da alfabetização é maior em regiões mais homogêneas e transborda para a vizinhança. Em média, um crescimento de 1% na taxa de alfabetização faria a renda crescer 1,684%. Assim, o produto interno bruto (PIB) brasileiro aumentaria 15,91% caso o Brasil atingisse o nível de alfabetização da Europa Central. Alternativamente, estados mais pobres, como Alagoas, Piauí, Paraíba e Maranhão (e outros dez), poderiam aumentar suas respectivas rendas em 32,77%, 29,09%, 26,53% e 23,97%, caso atingissem a média nacional de alfabetização. Acredita-se que o nivelamento da alfabetização e, de forma mais geral, do próprio ensino poderia estimular o crescimento brasileiro e reduzir as desigualdades regionais de forma simultânea. Além disso, verificou-se que o impacto da alfabetização sobre o crescimento brasileiro ficaria superestimado (em 22,02%) se a dependência espacial fosse ignorada. Se isso ocorresse apenas com a heterogeneidade espacial ou se ambos os efeitos espaciais fossem ignorados, o impacto dessa variável tenderia a ficar subestimado (em -30,73% e -15,48%, respectivamente).

O restante do trabalho está organizado da seguinte forma: a próxima seção contém o referencial teórico deste trabalho; na seção seguinte encontram-se a metodologia e a descrição da base de dados usada nas estimações; as seções subsequentes apresentam os resultados, considerações finais, referências e anexo, respectivamente.

3 Renzi *et al.* (2019) e Vieira (2009) consideraram apenas municípios paranaenses e paulistas, respectivamente, enquanto Soares (2009) focou sua pesquisa nos indivíduos do Nordeste. Apenas Almeida, Valadares e Sedyama (2017) analisaram o Brasil, mas optaram por uma desagregação estadual.

4 Dentre os autores citados, Castelló-Climent e Mukhopadhyay (2013), Khalfaoui (2015) e Almeida, Valadares e Sedyama (2017) usaram técnicas com dados em painel para controlar a heterogeneidade espacial, mas ignoraram a dependência espacial. Já Vieira (2009), Leija (2019) e Renzi *et al.* (2019) buscaram controlar apenas a dependência (*cross-section*), deixando a heterogeneidade espacial em segundo plano.

5 Dentre as pesquisas nacionais, Renzi *et al.* (2019), Soares (2009) e Vieira (2009) consideraram dados *cross-section* para 2010, 2001-2005 e 1980-2000, respectivamente, enquanto Almeida, Valadares e Sedyama (2017) usaram um painel para 2001-2011. Apenas Castelló-Climent e Mukhopadhyay (2013), ao analisar a Índia no período de 1961 a 2001 (dados em painel), consideraram um período superior ao aqui proposto.

REFERENCIAL TEÓRICO

Esta seção contém uma revisão dos principais trabalhos empíricos que analisaram os efeitos econômicos gerados pelo analfabetismo e a descrição do modelo teórico utilizado nas estimações.⁶

A questão do analfabetismo

Para Suresh Lal (2015), o analfabetismo é composto por indivíduos incapazes de ler, escrever ou resolver questões básicas de matemática. Logo, não apresentam o mínimo necessário para desempenhar uma função efetiva na sociedade, fato que reduz as oportunidades de emprego, compromete a geração de renda e prejudica a própria saúde desses cidadãos, tornando-os alvos fáceis da pobreza e da criminalidade e, frequentemente, dependentes de programas sociais. O autor estima que o custo do analfabetismo no mundo chegue a US\$ 1,19 trilhão.⁷

O fato é que indivíduos analfabetos se tornam aptos a desempenhar apenas tarefas braçais básicas, que não agregam muito valor à produção e contribuem pouco para a geração de riqueza (ADELAKUN, 2011, p. 30). Como a produtividade relativa desses trabalhadores tende a ser menor que a dos indivíduos mais qualificados, a teoria econômica sugere que seus salários, em termos reais, seriam os menores possíveis.⁸

Quanto ao aspecto econômico, Azariadis e Drazen (1990) afirmam que nenhum país cresceu de forma acelerada após a II Guerra sem uma população altamente alfabetizada. Rauch (1988) argumenta que a renda de diversos países, com alfabetização acima de 95% em 1960, convergiu para um alto patamar nos anos subsequentes. Para Romer (1989), a alfabetização favorece o investimento e, portanto, afetaria o crescimento de forma indireta. Segundo Barro (1991), a taxa de alfabetização impulsionou o crescimento da renda *per capita* de diversos países no período de 1960 a 1985. Khalifaoui (2015), ao analisar diversos países muçulmanos no período de 1990 a 2014, concluiu que o efeito negativo e significativo do analfabetismo é uma das principais barreiras ao crescimento.

Dentre as pesquisas para países específicos, Adalakun (2011) afirma que o analfabetismo elevado prejudicou a produtividade do trabalhador e comprometeu o

6 Apesar do foco econômico, os impactos do analfabetismo vão além dessa esfera. Segundo Rakodi e Lloyd-Jones (2002, p. 208), indivíduos analfabetos costumam ser excluídos de processos eleitorais e de outras formas democráticas de decisão, perdendo qualquer capacidade de alterar o meio no qual se encontram.

7 Nas palavras do autor: "*Illiteracy means the lack of Reading, Writing and Arithmetic (3R's) skills. It is the lack of minimum capabilities needed to function effectively in a society. [...] the cost of illiteracy to the global economy which is estimated at 1.19 trillion dollars; [...] The illiterate people trapped in a cycle of poverty with limited opportunities for employment or income generation and higher chances of poor health, turning to crime and dependence on social welfare.*" (SURESH LAL, 2015, p. 663).

8 Em uma economia concorrencial, na qual as empresas buscam maximizar o lucro, o salário real de um grupo de trabalhadores é equivalente à produtividade média desse grupo (MANKIW, 2010, p. 46-48).

crescimento da Nigéria. Leija (2019) argumenta que a alfabetização se mostrou significativa na redução da pobreza extrema do México. Segundo Gustafsson *et al.* (2010), a redução do analfabetismo (a níveis próximos aos de outros países em desenvolvimento) poderia aumentar o PIB da África do Sul de 23% a 30%, além de melhorar a autoestima e a saúde do trabalhador. Murthy e Okunade (2014), considerando 31 províncias chinesas, sugerem que altas taxas de analfabetismo reduzem a expectativa de vida do trabalhador, prejudicando sua produtividade ao longo da vida. Lin (1997), após analisar 30 províncias chinesas, revelou que a redução do analfabetismo estimula o crescimento, mas não é suficiente para que os recém-alfabetizados migrem da agricultura para a indústria. Ademais, Castelló-Climent e Mukhopadhyay (2013), considerando 16 estados da Índia no período de 1961 a 2001, afirmam que a simples alfabetização não afetou significativamente o crescimento dessas regiões. Contudo, a continuidade dos estudos (3º grau) mostrou-se significativa.

No Brasil, Almeida, Valadares e Sediyaama (2017), utilizando um painel dinâmico para os estados brasileiros no período de 2001 a 2011, constataram que o analfabetismo é típico de regiões pobres. Para Soares (2009), o analfabetismo compromete a produtividade e o salário dos trabalhadores nordestinos (um analfabeto recebe, em média, 37% menos que um trabalhador com a 8ª série completa e até 311% menos que alguém com nível superior). Por fim, Vieira (2009) revela que o analfabetismo elevado prejudicou o crescimento dos municípios do estado de São Paulo no período de 1980 a 2000. A Tabela 1 indica que, embora a taxa de analfabetismo brasileira esteja abaixo da média mundial e tenha auferido considerável redução entre 1980 e 2010, saindo de 25,4% para 9,6%, permanece, em 2010, 15,5% acima da média da América Latina e Caribe, 58% maior que a dos países de renda média-alta (grupo do qual o Brasil faz parte), 67,8% maior que a do Extremo Oriente e Pacífico, 297,7% maior que a da Europa e Ásia Central e 788,6% maior que a da Europa Central e Países Bálticos. Portanto, há espaço para reduções significativas na taxa de analfabetismo brasileira.⁹

9 Os dados do Instituto de Pesquisa Econômica Aplicada (IPEA, 2018) e do Instituto Brasileiro de Geografia e Estatística (IBGE, 2011) sugerem que a média de analfabetismo dos municípios brasileiros, considerados nesta pesquisa, é ainda maior que a do Banco Mundial (2018), ou seja, 36,74%, 30,26%, 21,35% e 15,89% em 1980, 1991, 2000 e 2010, respectivamente (ver Tabela A.1, Anexo).

TABELA 1 - Taxa de analfabetismo no Brasil e no mundo (pessoas acima de 15 anos)

	1980	1991	2000	2010
Brasil	25,41%	n. r.	13,63%	9,62%
Mundo árabe	51,81%	43,70%	34,88%	29,24%
Europa Central e Países Bálticos	n. r.	1,80%	1,49%	1,08%
Extremo Oriente e Pacífico	31,48%	18,42%	9,69%	5,73%
Europa e Ásia Central	n. r.	4,33%*	3,42%	2,42%
América Latina e Caribe	19,41%	15,01%	10,97%	8,33%
Oriente Médio e Norte da África	52,57%	40,81%	30,89%	23,51%
Sul da Ásia	61,28%	54,14%	42,31%	33,98%
África Subsaariana	n. r.	47,08%	43,34%	40,44%
Países de renda baixa	n. r.	51,76%	47,28%	43,08%
Países de renda média-baixa	48,78%	41,63%	33,31%	27,24%
Países de renda média	38,05%	28,47%	20,65%	16,33%
Países de renda média-alta	29,38%	17,76%	9,56%	6,09%
Mundo	32,86%	25,09%	19,22%	15,90%

Fonte: Banco Mundial (2018).

Nota: n. r. = não reportado.

* Como não havia informação para 1991, considerou-se o valor de 1994.

O modelo teórico

Na tentativa de aprimorar a capacidade explicativa dos modelos de crescimento econômico, Mankiw, Romer e Weil (1992) incluíram o estoque de capital humano (H) na versão proposta por Solow (1956). Formalmente:

$$Y(t) = K(t)^\alpha H(t)^\beta [A(t)L(t)]^{1-\alpha-\beta} \quad (1),$$

onde Y , K e L representam, respectivamente, a produção, o estoque de capital físico e a força de trabalho (α e β são coeficientes de sensibilidade). Logo, a renda *per capita* no estado estacionário seria:

$$\ln y_t = \ln \left[\frac{Y_t}{L_t} \right] = \ln A_0 + gt - \frac{\alpha+\beta}{1-\alpha-\beta} \ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta} \ln(s_k) + \frac{\beta}{1-\alpha-\beta} \ln(s_h) \quad (2),$$

onde $n+g+\delta$ é o somatório do crescimento populacional, avanço tecnológico e depreciação do capital físico, enquanto A_0 , s_k e s_h correspondem à dotação inicial e às parcelas da renda investidas em capital físico e humano, respectivamente.

Relaxando a hipótese de que os países já estão no estado estacionário, temos:

$$\ln y_t = (1 - e^{-\lambda t}) \ln(y^*) + e^{-\lambda t} \ln y_0 \quad (3),$$

onde $y^* = Y/AL$, y_0 é a renda inicial e λ representa a taxa de convergência, ou seja: $\lambda = (n+g+\delta)(1-\alpha-\beta)$. Substituindo y^* (na equação 2) e subtraindo em ambos os lados, obtém-se:

$$\ln\left(\frac{y_t}{y_0}\right) = (1 - e^{-\lambda t}) \left[-\ln(y_0) + \frac{\alpha}{1-\alpha-\beta} \ln(s_k) - \frac{\alpha+\beta}{1-\alpha-\beta} \ln(n + g + \delta) + \frac{\beta}{1-\alpha-\beta} \ln(s_h) + \ln A_0 + gt \right] \quad (4).$$

Operacionalmente, a equação 4 pode ser reescrita como:

$$\ln(y_t/y_0) = \beta_0 + \beta_1 \ln(y_0) + \beta_2 \ln(s_k) + \beta_3 \ln(n + g + \delta) + \beta_4 \ln(s_h) + \varepsilon \quad (5),$$

onde $\beta_0, \beta_1, \dots, \beta_4$ são coeficientes de sensibilidade e ε é um resíduo aleatório.

Mankiw, Romer e Weil (1992) concluíram que Solow (1956) havia acertado a capacidade explicativa de K e L (em torno de 50%) e os sinais dos impactos de S_k e n sobre Y . Contudo, a magnitude desses efeitos havia sido subestimada.

Embora tenham avançado na discussão sobre o crescimento, Mankiw, Romer e Weil (1992) desconsideraram a importância do espaço. Nesse sentido, Ertur e Koch (2007) fazem uma contribuição importante, ao reconhecerem que a interdependência tecnológica entre regiões distintas pode gerar transbordamentos espaciais. Formalmente, os autores assumem que:

$$A_i(t) = \Omega(t)k_t^\phi(t) \prod_{j \neq i}^N A_j^{\gamma w_{ij}}(t) \quad (6).$$

Portanto, a tecnologia $A_i(t)$ dependeria: a) de um fator exógeno e idêntico entre as regiões, conforme proposto por Solow, $\Omega(t)$; b) do crescimento de $k = K/L$ em cada região, $k_t^\phi(t)$, onde $0 \leq \phi < 1$ reflete a magnitude desse efeito; c) dos transbordamentos espaciais da tecnologia, $\prod_{j \neq i}^N A_j^{\gamma w_{ij}}(t)$.¹⁰ Desse modo, partindo-se da função de produção $y = A + \alpha k$, é possível incluir os transbordamentos espaciais oriundos da tecnologia:

$$y = \Omega + (\alpha + \varphi)k - \alpha\gamma Wk + \gamma W y \quad (7).$$

Como resultado, Ertur e Koch (2007) afirmam que a produção de uma região i qualquer cresceria com o aumento da poupança e da renda de seus j vizinhos e diminuiria com o crescimento populacional da vizinhança. Portanto, assumindo a existência desses transbordamentos espaciais e que a parcela não consumida da renda é investida em capital físico (s_k) ou humano (s_h), conforme o modelo MRW (1992), tem-se:

10 Para Ertur e Koch (2007), o impacto dos transbordamentos da tecnologia dependem da sua magnitude inicial (parâmetro $0 < \gamma < 1$) e da conectividade entre as regiões (medida pelo termo exógeno w_{ij} , para i e $j = 1, \dots, N$ e $j \neq i$). Quanto maior a proximidade entre i e seus j vizinhos, maior será o impacto.

$$\ln(y_t/y_0) = \beta_0 + \beta_1 \ln(y_0) + \beta_2 \ln(s_k) + \beta_3 \ln(n + g + \delta) + \beta_4 \ln(s_h) + \beta_5 W \ln(s_k) + \beta_6 W \ln(n + g + \delta) + \beta_7 W \ln(s_h) + \varepsilon \quad (8),$$

onde W é uma matriz de pesos espaciais que capta o efeito associado à vizinhança de uma região i qualquer. Mais detalhes sobre W estão em Almeida (2012, cap. 3).

METODOLOGIA E BASE DE DADOS

O modelo estimado (equação 9) contém quatro *dummies* para captar a heterogeneidade regional (referentes às regiões Sul, Sudeste, Nordeste e Centro-Oeste)¹¹ e uma variável de distância de cada município até sua respectiva capital estadual.¹² Visando a captar os transbordamentos espaciais (rever seção anterior), as variáveis explicativas foram defasadas espacialmente e incluídas no modelo (β_{11} a β_{13}).¹³ Além disso, a análise do I de Moran, proposta por Cliff e Ord (1981), indicou autocorrelação espacial positiva associada à variável dependente ($I = 0,303$). Logo, assim como em Easterly e Levine (1998), o crescimento de um município impulsiona o dos seus vizinhos e vice-versa. Para controlar esse efeito, inclui-se a variável dependente defasada espacialmente (β_{10}).¹⁴ Por fim, um sinal negativo associado a $[\ln(y_{t-1})]$ indica que há convergência de renda.

$$\ln\left(\frac{y_t}{y_{t-1}}\right) = [\beta_0 + \beta_1 \ln(y_{t-1}) + \beta_2 \ln(s_k) + \beta_3 \ln(n + g + \delta) + \beta_4 \ln(s_h)] + [\beta_5 D_{Sul} + \beta_6 D_{Sud} + \beta_7 D_{Nord} + \beta_8 D_{Cent_{oest}} + \beta_9 \ln(Dist_{Mun})] + \{\beta_{10}\{W[\ln(y_t/y_{t-1})]\} + \beta_{11}\{W[\ln(s_k)]\} + \beta_{12}\{W[\ln(n + g + \delta)]\} + \beta_{13}\{W[\ln(s_h)]\}\} + u \quad (9).$$

A literatura sugere que a alfabetização (s_h) de um município i estimularia seu crescimento (Y_t / Y_{t-1}), pois indivíduos analfabetos seriam menos produtivos (seção “A questão do analfabetismo”). Todavia, o crescimento de i também poderia ser afetado pela alfabetização de seus j vizinhos ($W_s s_h$), visto que indivíduos de i poderiam estudar em j (que agora tem maior taxa de alfabetização) e pessoas de j (que, em média, tornaram-se mais produtivas após a alfabetização) poderiam

11 As *dummies* captam a diferença das demais regiões em relação à Região Norte, tida como referência (excluída).

12 O comércio desigual entre as capitais estaduais e os demais municípios pode favorecer o acúmulo de recursos nas capitais (CHRISTALLER, 1966), que passariam a gerar externalidades positivas sobre o crescimento/desenvolvimento de seus vizinhos (MYRDAL, 1957).

13 Segundo Lall e Yilmaz (2001, *apud* ABREU; GROOT; FLORAX, 2005), os transbordamentos do capital humano explicam parte da desigualdade dos estados americanos. No Brasil, Oliveira (2005) argumenta que tais externalidades impulsionaram o crescimento das cidades cearenses na década de 1990.

14 O termo $W_s(y_t / y_{t-1})$ gera endogeneidade (*i.e.*, o crescimento de um município afeta seus vizinhos e vice-versa) e requer o uso dos mínimos quadrados em dois estágios (MQ2E), com $W^2 X_t$ (*i.e.*, segunda defasagem das variáveis explicativas) como instrumento (ALMEIDA, 2012).

trabalhar em i (ALMEIDA, 2012, p. 21-27). Em ambos os casos, haveria um transbordamento espacial positivo associado à alfabetização. Além disso, Lesage e Pace (2014) e Golgher e Voss (2016) afirmam que a autocorrelação espacial, mensurada via $W_{-}(Y_t/Y_{t-1})$, geraria um efeito multiplicador, pois o crescimento de i afetaria seus j vizinhos e vice-versa, gerando um processo retroalimentador. Nesse caso, o efeito total da alfabetização seria $(s_h + W_{-}s_h)/[(1 - W_{-}(Y_t/Y_{t-1})]$. A estratégia empírica desta pesquisa consistiu em analisar o impacto da taxa de alfabetização sobre o crescimento econômico dos municípios brasileiros usando dados em painel para o período 1980-2010 (cortes considerados: 1980, 1991, 2000 e 2010).¹⁵ No contexto de dados de painel, geralmente estima-se o modelo POLS (*pooled ordinary least squares*) usando o teste de Breusch-Pagan (1980) para verificar se existe algum efeito não observado, c_i ($H_0: \sigma_c^2 = 0$). Caso H_0 prevaleça, o POLS é o mais indicado. Caso contrário ($\sigma_c^2 \neq 0$), estimam-se os modelos FE (*fixed effects* – efeitos fixos) e RE (*random effects* – efeitos aleatórios) usando o teste de Hausman (1978) para definir se c_i afeta as demais variáveis explicativas ($H_0: E[c_i | x_{it}] = 0$). Se H_0 for verdadeiro, FE e RE serão consistentes, porém RE será mais eficiente. Caso contrário, apenas FE será consistente. Caso o teste de Hausman (1978) rejeite H_0 , o teste Breusch-Pagan (1980) torna-se desnecessário (WOOLDRIDGE, 2002).

Base de dados

Para compor a equação 9, foram utilizadas as seguintes variáveis:

- a) $y_t \Rightarrow$ renda municipal *per capita*. Utilizou-se o rendimento domiciliar do mês anterior ao de referência (*i.e.*, 01/08 para o Censo de 2000 e 01/09 para 1980 e 1991), expresso em reais (R\$) de 2000 (IPEA, 2018). Em 2010, usou-se o rendimento domiciliar mensal médio (IBGE, 2011) deflacionado, para R\$ de 2000, pelo Índice Geral de Preços do Mercado (IGP-M) da Fundação Getúlio Vargas (FGV) (Ipea, 2018);
- b) $s_k \Rightarrow$ parcela da renda investida em capital físico. Utilizou-se a despesa anual em capital físico de cada município,¹⁶ fornecida pelo Ministério da Fazenda e convertida em R\$ de 2000 via Índice de Preços ao Consumidor Amplo do Instituto Brasileiro de Geografia e Estatística (IPCA/IBGE) (IPEA, 2018). Os valores anuais e deflacionados foram somados ao longo dos períodos 1985-1991, 1991-2000 e 2001-2010 e divididos pela população de cada município (IPEA, 2018) e pelo número de anos desses períodos. Assim, obtivemos a

15 A estimação em painel permite: a) incluir mais observações, assegurando as propriedades assintóticas dos estimadores e das inferências; b) promover ajustamentos dinâmicos e análises intertemporais; e c) controlar os efeitos não observados constantes no tempo (*e.g.*, cultura, preferências, clima, etc.) (HSIAO, 2003).

16 Esta variável passou a ser divulgada em 1985. Como a quantidade de anos no período de 1985 a 1991 é menor (sete anos), dividiu-se o valor acumulado pelo número de anos de cada período considerado.

média anual *per capita* de despesa municipal, em capital físico, para cada período mencionado;¹⁷

- c) $s_h \Rightarrow$ parcela da renda investida em capital humano. Considerou-se o percentual de pessoas alfabetizadas, com 15 anos ou mais, de cada município (valor médio dos períodos considerados). O percentual de analfabetos é disponibilizado pelo Ipea (2018), para 1980, 1991 e 2000, e pelo IBGE (2011), para 2010. O percentual de analfabetos foi convertido em taxa de alfabetizados (*i.e.*, $100\% - \% \text{ analfabetos}$);¹⁸
- d) $(n + g + \delta) \Rightarrow$ soma das taxas de crescimento populacional (n), progresso tecnológico (g) e depreciação (δ). O crescimento populacional (n) (IPEA, 2018) foi somado ao valor de $g + \delta$, proposto por Mankiw, Romer e Weil (1992, p. 413): “assumimos que $g + \delta$ é 0.05 e que variações nessa hipótese teriam pouco impacto nas estimativas”.¹⁹ Dadas as variações decenais consideradas, a taxa de depreciação e tecnologia foi $(1,05)^{10} - 1 = 0,629$. Para possibilitar o uso do logaritmo natural sobre $(n + g + \delta)$, somou-se 1 ao resultado dessas variações;
- e) $W \Rightarrow$ matriz de pesos utilizada nas defasagens espaciais dos modelos (controle da “dependência espacial”). Como critério de conectividade, consideraram-se os $j = 5$ vizinhos mais próximos de uma região i qualquer;²⁰
- f) **Dummies de Região** \Rightarrow *dummies* regionais, como a heterogeneidade espacial pode afetar o impacto das variáveis explicativas (ALMEIDA, 2012). As Figuras A.1, A.2 e A.3 (baseadas no *I de Moran Local* – *vide Anexo*)²¹ sugerem a existência de agrupamentos municipais, com características semelhantes intragrupos e distintas entre os grupos. Incluem-se quatro variáveis binárias (*dummies*) referentes às regiões Sul, Sudeste, Nordeste e Centro-Oeste²² a fim de captar tais características regionais;
- g) **Dist_Mun** \Rightarrow distância (em quilômetros) do município em relação à sua respectiva capital estadual (IPEA, 2018).²³ Pretende-se controlar o possível

17 Dados o período e a escala municipal considerados, não foi possível incluir os gastos privados em bens de capital (KROTH; DIAS, 2012). A escassez e a falta de qualidade das *proxies* municipais para o capital físico (s_k) e humano (s_h) no Brasil são discutidas por Firme e Simão Filho (2014, p. 687-689).

18 Embora não capte a amplitude do capital humano, ressalta-se que, devido ao período e à escala geográfica considerados, não haveria outra *proxy* viável para s_h , visto que a média municipal de anos estudados (mais usual) não foi avaliada em 2010. Assim, dadas as ressalvas, assume-se que os municípios mais alfabetizados dão mais atenção à educação e, dessa forma, ao capital humano.

19 No original: “We assume that $g + \delta$ is 0.05; reasonable changes in this assumption have little effect on the estimate”.

20 Testaram-se as matrizes de contiguidade rainha e torre e as especificações $K = 5$ e $K = 10$ vizinhos mais próximos. A matriz $K = 5$ foi selecionada, conforme Baumont (2004, p. 13), por apresentar o maior *I de Moran* associado à variável dependente no período analisado (média de $I = 0,303$).

21 Proposto por Anselin (1995) e apresentado em Almeida (2012, cap. 4).

22 Os resultados dessas *dummies* são relativos à Região Norte (não incluída).

23 Nas capitais, onde a distância é 0, adotou-se o valor 1, pois $\ln(1) = 0$.

transbordamento positivo associado às capitais estaduais (MYRDAL, 1957; CHRISTALLER, 1966). Nesse caso, um sinal negativo (esperado) indicaria que cidades mais distantes das capitais cresceriam menos que as mais próximas.

O Quadro 1 revela os sinais esperados das variáveis consideradas e suas respectivas justificativas.²⁴

QUADRO 1 - Descrição e sinais esperados das variáveis

TIPO DE VARIÁVEL	SIGLA	DESCRIÇÃO	SINAL ESPERADO	HIPÓTESES PARA O SINAL ESPERADO
Dependente	$Ln(Y_t/Y_{t-1})$	Varição da renda <i>per capita</i>	n/a	n/a
Explicativas: MRW (1992)	$Ln(Y_{t-1})$	Renda <i>per capita</i> Inicial	-	Convergência da renda
	$Ln(s_k)$	Capital Físico	+	Insumo gerador de renda
	$Ln(s_h)$	Capital Humano	+	Insumo gerador de renda
	$Ln(n+g+\delta)$	Taxas de crescimento Pop.(n) + tecnol.(g) + deprec.(δ)	-	Aumento de <i>n</i> reduz renda <i>per capita</i>
Explicativas: Heterogeneidade Espacial	Sul	Municípios da Região Sul	Não definido	Dependem de questões inerentes a cada região
	Sudeste	Municípios da Região Sudeste	Não definido	
	Nordeste	Municípios da Região Nordeste	Não definido	
	Centro-Oeste	Municípios da Região Centro-Oeste	Não definido	
	$Ln(dist_est)$	Distância do município até sua capital estadual	-	<i>Spillover</i> positivo das capitais estaduais
Explicativas: Transbordamento Espacial	$W_Ln(Y_t/Y_{t-1})$	Varição da Renda Defasada Espacialmente	+	<i>I</i> de Moran positivo (vide nota n. 22)
	$W_Ln(s_k)$	Capital Físico Defasado Espacialmente	+	Efeito transbordamento (mesmo sinal de s_k)
	$W_Ln(s_h)$	Capital Humano Defasado Espacialmente	+	Efeito transbordamento (mesmo sinal de s_h)
	$W_Ln(n+g+\delta)$	Taxas de crescimento Pop.(n) + tecnol.(g) + deprec.(θ) Defasadas Espacialmente	-	Efeito transbordamento (mesmo sinal de $n+g+\theta$)

Fonte: Elaboração do autor com base na literatura consultada.

Nota: Ln = logaritmo natural; N/A = não aplicada.

24 As estatísticas descritivas associadas aos dados utilizados estão disponíveis na Tabela A.1 (Anexo).

RESULTADOS

Os resultados dos modelos sem controles espaciais (Quadro 2, Modelo A) indicaram que todas as variáveis testadas são significativas e apresentam o sinal sugerido no Quadro 1. Tais especificações sugerem que um crescimento de 10% na taxa de alfabetização (s_h) faria a renda *per capita* dos municípios brasileiros crescer de 14,21% (A.1 e A.2) a 20,49% (A.3), sendo esse último valor preferível, segundo o teste de Hausman (1978).²⁵

As *dummies* regionais (Quadro 2, modelos B, B.1 e B.2), embora significativas, não alteraram consideravelmente o impacto atribuído à alfabetização, que passou de 1,421 (A.1 e A.2) para 1,466 (B.1 e B.2). Logo, não é possível afirmar que as características específicas dessas cinco regiões brasileiras são responsáveis pela magnitude do efeito associado à educação de base. Entretanto, o modelo de FE (A.3 e B.3), que permite um controle individual da heterogeneidade espacial, revela que o impacto da alfabetização sobre a renda seria maior (*i.e.*, 2,049) em regiões mais homogêneas.

O controle da dependência espacial (Quadro 2, modelos C, C.2, C.3 e D.3) reduziu o efeito inicial da alfabetização (s_h), mas revelou um transbordamento espacial positivo e significativo associado a essa variável (W_{s_h}). Portanto, a melhora da alfabetização em uma região qualquer geraria efeitos positivos sobre a renda de seus vizinhos, aumentando o impacto total dessa variável. Lesage e Pace (2014) e Golgher e Voss (2016) reforçam que a análise dos modelos C e D (Quadro 2) requer certos cuidados, pois, nesses casos, o impacto total de uma variável qualquer se torna o somatório de seu efeito direto (inicial: s_h + transbordamento: W_{s_h}) e indireto, oriundo do efeito multiplicador gerado por $W_{(Y_t / Y_{t-1})}$. Assim, o efeito direto da alfabetização sobre a renda seria 0,345 (D.1 e D.2) e 0,630 (D.3), enquanto o total envolve o seguinte cálculo: $\{[1/(1-0,867)] * 0,345\}=2,594$ (D.1 e D.2) e $\{[1/(1-0,626)] * 0,630\}=1,684$ (D.3).

Note que o menor impacto total da alfabetização no modelo de FE (D.3: 1,684), quando comparado ao de RE (D.2: 2,594), parece contradizer o argumento anterior, de que a heterogeneidade espacial (presente no D.2 e controlada no D.3) prejudicaria o impacto dessa variável. Contudo, como o efeito direto da alfabetização é maior no FE (D.3: 0,630) quando comparado ao RE (D.2: 0,345), a hipótese de que a homogeneidade espacial favorece o impacto positivo da educação se mantém. Na realidade, o impacto reduzido em D.3 (em relação a D.2) deve-se, exclusivamente, ao menor efeito multiplicador oriundo da autocorrelação espacial (*i.e.*, coeficiente 0,626) que afeta igualmente todas as variáveis do modelo e, portanto, não tem relação direta com o objeto de estudo deste trabalho.

25 O teste de Hausman (1978) indicou que os modelos FE são preferíveis aos RE em todas as estimativas do Quadro 2 (*i.e.*, A.3, B.3, C.3 e D.3).

Comparando o efeito total da alfabetização sem controles espaciais (A.1: 1,421), com controle apenas para a heterogeneidade (A.3: 2,049) e após controlar ambos os efeitos espaciais (D.3: 1,684), nota-se que seu impacto ficaria superestimado se a dependência espacial fosse ignorada (em 22,02%) e subestimado se a heterogeneidade espacial (em -30,73%) ou ambos os efeitos forem ignorados (em -15,48%). Logo, o efeito da alfabetização seria maior em regiões mais homogêneas (diferença entre A.1 e A.3) e menor após o controle da dependência espacial (diferença entre A.3 e D.3). Como o modelo A.3 ignora os transbordamentos associados às variáveis explicativas (que se mostraram positivos em D.3), parte desses efeitos benéficos (não controlados) pode acabar embutida nos coeficientes das variáveis consideradas em A.3 (inflando-os). Com base nesses resultados, é possível inferir que o custo do analfabetismo no mundo, sugerido por Suresh Lal (2015), ao desconsiderar os efeitos espaciais, seja ainda maior que os US\$ 1,19 trilhão previstos pelo autor.

QUADRO 2 – Resultado das estimações

TIPO DE CONTROLE	MODELO A – MRW (1992) SEM CONTROLE ESPACIAL			MODELO B – MRW (1992) CONTROLE P/ HETEROGENEIDADE ESPACIAL			ESTIMAÇÃO EM 2 ESTÁGIOS UTILIZANDO INSTRUMENTOS PARA $W_LN(Y_t/Y_{t-1})$					
	MODELO C. MRW (1992) CONTROLE P/ TRANSBORDAMENTO ESPACIAL			MODELO D. MRW (1992) CONTROLE P/ TRANSBORDAMENTO E HETEROGENEIDADE ESPACIAIS			(C.1) POLS	(C.2) RE	(C.3) FE	(D.1) POLS	(D.2) RE	(D.3) FE
	(A.1) POLS	(A.2) RE	(A.3) FE	(B.1) POLS	(B.2) RE	(B.3) FE						
$Ln(Y_{t-1})$	-0.670***	-0.670***	-1.143***	-0.668***	-0.668***	-1.143***	-0.245***	-0.245***	-0.685***	-0.265***	-0.265***	-0.685***
$Ln(s_r)$	0.112***	0.112***	0.223***	0.109***	0.109***	0.223***	0.017***	0.017***	0.045***	0.018***	0.018***	0.045***
$Ln(s_r)$	1.421***	1.421***	2.049***	1.466***	1.466***	2.049***	0.401***	0.401***	0.326***	0.412***	0.412***	0.326***
$Ln(n+g+\delta)$	-0.038*	-0.038*	-0.099***	-0.054**	-0.054**	-0.099***	-0.013	-0.013	-0.017	-0.007	-0.007	-0.017
D_Sul				0.096**	0.096**	Excluído				0.071***	0.071***	Excluído
D_Sudeste				0.089***	0.089***	Excluído				0.054***	0.054***	Excluído
D_Nordeste				0.132***	0.132***	Excluído				-0.016	-0.016	Excluído
D_Centro-Oeste				0.143***	0.143***	Excluído				0.060***	0.060***	Excluído
$Ln(dist_est)$				-0.019***	-0.019***	Excluído				-0.011***	-0.011***	Excluído
$W_Ln(Y_t/Y_{t-1})$							0.861***	0.861***	0.626***	0.867***	0.867***	0.626***
$W_Ln(s_r)$							0.038***	0.038***	0.207***	0.048***	0.048***	0.207***
$W_Ln(s_r)$							0.047	0.047*	0.304***	-0.067*	-0.067**	0.304***
$W_Ln(n+g+\delta)$							0.113***	0.113***	0.039	0.126***	0.126***	0.039
Constante	-3.165***	-3.165***	-4.053***	-3.362***	-3.362***	-4.054***	-0.952***	-0.952***	-0.359***	-0.437***	-0.437***	-0.359***
N	11961	11961	11961	11961	11961	11961	11961	11961	11961	11961	11961	11961
R ²	0.47		0.72	0.48		0.72	0.78		0.78		0.78	
R ² _overall		0.47	0.45		0.48	0.45		0.78		0.78		0.67
R ² _between		0.23	0.28		0.27	0.28		0.53		0.53		0.38
R ² _within		0.71	0.72		0.71	0.72		0.80		0.80		0.86
AIC	4773.7		-3571.8	4604.00		-3571.8						
BIC	4810.6		-3534.8	4677.9		-3534.8						
Teste de Hausman (1978) associado ao Modelos D: $\chi^2 = 11274.40; Prob. > \chi^2 = 0.000$												

Fonte: Elaboração do autor utilizando o software STATA.

Notas: a) Os modelos POLS e RE foram estimados utilizando-se a matriz robusta de White (WOOLDRIDGE, 2002).

b) * p < 0.05; ** p < 0.01; *** p < 0.001.

Como o teste de Hausman (1978) indicou que os modelos FE são sempre preferíveis ao RE, pode-se concluir que um crescimento de 1% na taxa de alfabetização faria a renda *per capita* dos municípios brasileiros crescer, em média, 1,684%. Com base nesse resultado, foram efetuadas simulações sobre: a) o valor provável do PIB *per capita* brasileiro, caso o país atingisse o mesmo nível de alfabetização da Europa Central e Países Bálticos (rever Tabela 1); b) o possível ganho na renda *per capita* dos estados brasileiros, com piores taxas de analfabetismo, caso atingissem o nível médio de alfabetização do país. Os resultados estão disponíveis nas Tabelas 2 e 3 e, em ambos os casos, utilizou-se o ano de 2010 como referência.

TABELA 2 - Alteração no PIB *per capita* brasileiro oriunda da melhora na taxa de alfabetização (ano-base: 2010)

POSIÇÃO	PAÍS	PIB 2010 REAL	NOVA POSIÇÃO	PAÍS	PIB 2010 SIMULADO
1	Catar	129852,73	1	Catar	129852,73
5	Cingapura	73060,99	5	Cingapura	73060,99
10	Suíça	54365,36	10	Suíça	54365,36
30	Japão	35883,03	30	Japão	35883,03
60	Chile	19348,13	60	Chile	19348,13
61	Argentina	18911,82	61	Argentina	18911,82
62	Letônia	18225,03	62	Letônia	18225,03
63	Irã	17879,37	63	Irã	17879,37
64	Peru	17248,62	64	Peru	17248,62
65	Uruguai	16970,30	65	Brasil	17015,75
66	Venezuela	16837,10	66	Uruguai	16970,30
67	Romênia	16730,24	67	Venezuela	16837,10
68	Barbados	16633,45	68	Romênia	16730,24
69	Líbano	16451,76	69	Barbados	16633,45
70	Azerbaijão	16318,09	70	Líbano	16451,76
71	Bielorrússia	16218,12	71	Azerbaijão	16318,09
72	México	16137,34	72	Bielorrússia	16218,12
73	Bulgária	15993,41	73	México	16137,34
74	Ilhas Maurício	15938,43	74	Bulgária	15993,41
75	Gabão	15642,07	75	Ilhas Maurício	15938,43
76	Panamá	15549,15	76	Gabão	15642,07
77	Brasil	14679,29	77	Panamá	15549,15
78	Suriname	14076,51	78	Suriname	14076,51
79	Montenegro	13845,83	79	Montenegro	13845,83
80	Tailândia	13750,23	80	Tailândia	13750,23
100	Paraguai	9712,92	100	Paraguai	9712,92
192	Congo	568,92	192	Congo	568,92

Fonte: Elaboração do autor com base nos dados da World Economic Outlook da International Monetary Fund (WEO/IMF, 2018).

Nota: Valores a preços constantes de 2010 avaliados em US\$ de paridade de poder de compra.

Retornando à Tabela 1, nota-se que a taxa de alfabetização brasileira em 2010 (*i.e.*, $100 - 9,62 = 90,38\%$) precisaria crescer, aproximadamente, 9,45% para atingir o mesmo nível da Europa Central e Países Bálticos (*i.e.*, $100 - 1,08 = 98,92\%$). Nesse caso, o crescimento econômico brasileiro aumentaria em $9,45 * 1,684 = 15,91\%$. Como a relação $(\text{PIB}_{2010}/\text{PIB}_{1980}) = (14.679,29/11.372,17) = 1,2908$ no Brasil,²⁶ a melhora na alfabetização impulsionaria o crescimento para $1,2908 * 1,1591 = 1,4962$. Logo, o PIB *per capita* brasileiro de 2010 atingiria $1,4962 * 11372,17 = 17015,75$,²⁷ ou seja, 15,91% a mais que o valor original do mesmo período. Tal avanço na alfabetização faria o PIB *per capita* brasileiro saltar da 77ª posição para a 65ª (entre 192 países), ultrapassando países como Bulgária, México, Venezuela e Uruguai e se aproximando do Peru, Argentina e Chile (Tabela 2).

As simulações associadas aos estados brasileiros (Tabela 3) seguiram o mesmo procedimento da Tabela 2.²⁸ Contudo, visando a apurar qual seria o ganho em termos de renda *per capita*, caso os estados atingissem a média de alfabetização nacional em 2010 (*i.e.*, 90,38%), consideraram-se apenas aqueles com taxas inferiores à média brasileira.

O estado de Alagoas, por ter a taxa mais alta de analfabetismo do país (24,33%), seria o mais beneficiado com o esforço em alcançar a média nacional. Tal política geraria um crescimento extra na renda *per capita* desse estado de, aproximadamente, 32,77%, fazendo-o subir da 24ª colocação para a 15ª no *ranking* de maiores rendas *per capita* estaduais do Brasil. Os efeitos positivos desse nivelamento da alfabetização seriam sentidos em 14 dos 26 estados, com destaque (além de Alagoas, já mencionado) para Piauí, Paraíba e Maranhão, onde o crescimento extra da renda *per capita*, oriundo exclusivamente da política de alfabetização, alcançaria 29,09%, 26,53% e 23,97%, respectivamente.

Note que os efeitos benéficos sobre a renda desses quatro estados, caso atingissem a média nacional de alfabetização, seriam equivalentes (ou superiores) aos descritos por Gustafsson *et al.* (2010) para a África do Sul, caso o país atingisse os níveis de alfabetização dos principais países desenvolvidos (de 23% a 30% de crescimento).

No que se refere à classificação das maiores rendas do país, destaca-se a melhora verificada no Piauí (do 25º lugar para o 17º) e na Paraíba (do 21º para o 13º) (Tabela 3). Como os estados mais pobres seriam os mais privilegiados nessa simulação, acredita-se que tal política estimule o crescimento e reduza as desigualdades simultaneamente.

26 Os valores do PIB, a preços constantes de 2010 e cotados em US\$ de paridade de poder de compra, estão disponíveis na base de dados da WEO/IMF (2018).

27 Devido à supressão de casas decimais (no texto), este valor diverge ligeiramente do exposto na Tabela 2.

28 Os cálculos foram feitos com base no crescimento econômico no período de 1991 a 2010.

TABELA 3 - Alteração na renda *per capita* dos estados brasileiros oriunda da melhora na taxa de alfabetização (ano-base: 2010)

N.	ESTADO	TAXA DE ALFABETIZAÇÃO (A)	RENDA 1991 (B)	RENDA 2010 (B)	NOVA POSIÇÃO	RENDA 2010 SIMULADA	VARIAÇÃO ORIUNDA DA ALFABETIZAÇÃO
1	São Paulo	95,68	746,22	1084,46	n.c.	983,46	n.c.
2	Rio de Janeiro	95,73	608,8	1039,30	n.c.	941,64	n.c.
3	Santa Catarina	95,87	449,78	983,90	n.c.	889,16	n.c.
4	Rio Grande do Sul	95,48	507,61	959,24	n.c.	873,10	n.c.
5	Paraná	93,72	439,09	890,89	n.c.	837,57	n.c.
6	Espírito Santo	91,88	377,38	815,43	n.c.	793,15	n.c.
7	Goiás	92,06	410,55	810,97	n.c.	786,19	n.c.
8	Mato Grosso do Sul	92,33	433,21	799,34	n.c.	771,05	n.c.
9	Mato Grosso	91,52	395,34	762,52	n.c.	746,66	n.c.
10	Minas Gerais	91,69	373,85	749,69	n.c.	731,79	n.c.
11	Rondônia	91,27	304,9	670,82	n. c.	659,92	n.c.
12	Roraima	89,67	437,24	605,59	12	613,78	1,35%
13	Amapá	91,6	378,57	598,98	n.c.	585,65	n.c.
14	Tocantins	86,91	243,58	586,62	12	626,19	6,74%
15	Rio Grande do Norte	81,46	240,33	545,42	12	646,14	18,47%
16	Amazonas	90,16	345,82	539,80	16	542,12	0,43%
17	Pernambuco	82	275,49	525,64	12	616,24	17,24%
18	Sergipe	81,6	247,78	523,53	12	618,53	18,15%
19	Acre	83,52	284,96	522,15	14	594,50	13,86%
20	Bahia	83,42	234,57	496,73	15	566,64	14,07%
21	Paraíba	78,09	196,59	474,94	13	600,95	26,53%
22	Ceará	81,26	219,83	460,63	15	547,81	18,93%
23	Pará	88,26	273,22	446,76	22	464,92	4,07%
24	Alagoas	75,67	211,98	432,56	15	574,30	32,77%
25	Piauí	77,08	167,03	416,93	17	538,20	29,09%
26	Maranhão	79,13	156,47	360,34	24	446,71	23,97%
	Brasil	90,38	447,56	793,87	-	-	-

Fonte: Elaboração do autor com base no Atlas-Brasil do Programa das Nações Unidas para o Desenvolvimento (PNUD, 2018).

Nota: A = percentual de pessoas com 15 anos ou mais que não sabem ler nem escrever; B = somatório da renda dos indivíduos residentes em domicílios particulares dividido pelo número total desses indivíduos. Valores em R\$ de 1 de agosto de 2010; n. c. = não calculado.

CONCLUSÃO

Esta pesquisa considerou uma versão espacial do modelo MRW (1992) com dados em painel, a fim de avaliar o impacto da taxa de alfabetização sobre o crescimento da renda *per capita* dos municípios brasileiros no período de 1980 a 2010. Os resultados sugerem que o impacto da alfabetização é maior em regiões mais homogêneas. Acredita-se que a heterogeneidade espacial gere a necessidade de

políticas educacionais mais específicas e mão de obra mais especializada do que ocorreria em regiões com maior semelhança em termos de cultura, clima, relevo e preferência locais. A título de exemplo, considere a dificuldade em alfabetizar povoados com dialetos distintos.

Ademais, constatou-se que melhorias na alfabetização geram transbordamentos positivos sobre a renda da vizinhança, aumentando o impacto total dessa variável. Os resultados ainda sugerem que o impacto total da alfabetização sobre o crescimento econômico ficaria superestimado se a dependência espacial fosse ignorada. Se isso ocorresse com a heterogeneidade espacial ou se ambos os efeitos espaciais fossem ignorados, o impacto dessa variável tenderia a ficar subestimado. Logo, é possível que o custo do analfabetismo no mundo, sugerido por Suresh Lal (2015), seja ainda maior que os US\$ 1,19 trilhão previstos pelo autor.

Realizados os devidos controles espaciais, estimou-se que um crescimento de 1% na taxa de alfabetização aumentaria a renda *per capita* em, aproximadamente, 1,684%. Tal resultado é coerente com a afirmação de Mankiw, Romer e Weil (1992), de que pequenas alterações no capital humano poderiam causar grandes alterações na renda. Com base nesse impacto, verificou-se que o PIB *per capita* brasileiro poderia ser, aproximadamente, 15,91% maior caso a taxa de alfabetização do país (90,38%) atingisse o nível da Europa Central (98,92%). Tal avanço faria o PIB brasileiro saltar da 77^a posição (no *ranking* dos 192 países mais ricos) para a 65^a, ultrapassando Bulgária, México, Venezuela e Uruguai e se aproximando do Peru, da Argentina e do Chile.

Alternativamente, simulou-se qual seria a renda *per capita* dos estados brasileiros caso eles atingissem a taxa média de alfabetização nacional. Os resultados indicam que haveria melhora na renda de 14 dos 26 estados, com destaque para Alagoas, Piauí, Paraíba e Maranhão, onde o crescimento oriundo exclusivamente da alfabetização atingiria 32,77%, 29,09%, 26,53% e 23,97%, respectivamente. Esses valores estão próximos aos estimados por Gustafsson *et al.* (2010) para a África do Sul, caso o país atingisse os níveis de alfabetização dos principais países desenvolvidos (onde o crescimento ficaria entre 23% e 30%). Se apenas Alagoas, com a pior taxa de alfabetização (75,67%), adotasse tal medida, sua renda subiria da 24^a para a 15^a colocação no *ranking* de maiores rendas *per capita* estaduais do Brasil.

Como não foram incluídas outras variáveis associadas ao capital humano (devido ao período e à escala considerados), é possível que seu efeito esteja subdimensionado. Além disso, caso a taxa de alfabetização também capte outras características municipais associadas a níveis mais elevados de educação, seu impacto ficaria supervalorizado. Por fim, como a exclusão dos municípios com informações faltantes

(*missings*) parece ter elevado a taxa média de analfabetismo,²⁹ é possível que o impacto da alfabetização seja ainda maior no Brasil (caso todos os municípios pudessem ter sido incluídos). Apesar das limitações mencionadas, os resultados desta pesquisa sugerem o uso mais racional e menos arbitrário dos recursos destinados à educação, ao indicar quais regiões seriam mais sensíveis às melhorias nessa área, e reforçam a importância desse tipo de investimento como instrumento de combate à pobreza e redução das desigualdades.

AGRADECIMENTOS

O autor agradece a colaboração do Prof. Dr. Ricardo da Silva Freguglia por suas importantes contribuições, principalmente no que se refere à metodologia adotada nesta pesquisa.

REFERÊNCIAS

ABREU, M.; GROOT, H. L. F.; FLORAX, R. J. G. M. Space and growth: a survey of empirical evidence and methods. *Région et Développement*, Paris, v. 21, p. 13-44, 2005.

ADELAKUN, O. J. Human capital development and economic growth in Nigeria. *European Journal of Business and Management*, Bruxelles, v. 3, n. 9, p. 29-38, 2011.

AGHION, P.; HOWITT, P. A model of growth through creative destruction. *Econometrica*, New York, v. 60, n. 2, p. 323-351, Mar. 1992.

ALMEIDA, E. *Econometria espacial aplicada*. Campinas: Alínea, 2012.

ALMEIDA, F. M.; VALADARES, J. L.; SEDIYAMA, G. A. S. A contribuição do empreendedorismo para o crescimento econômico dos estados brasileiros. *Revista de Empreendedorismo e Gestão de Pequenas Empresas*, São Paulo, v. 6, n. 3, p. 466-494, 2017.

ANSELIN, L. Local indicators of spatial association – LISA. *Geographical Analysis*, Columbus, v. 27, n. 1, p. 93-115, 1995.

ARROW, K. J. The economic implications of learning by doing. *Review of Economic Studies*, Oxford, v. 29, n. 3, p. 155-173, June 1962.

AZARIADIS, C.; DRAZEN, A. Threshold externalities in economic development. *The Quarterly Journal of Economics*, Oxford, v. 105, n. 2, p. 501-526, 1990.

BANCO MUNDIAL. *Indicadores de educação*. 2018. Disponível em: <https://data.worldbank.org/indicator>. Acesso em: 15 out. 2018.

29 A taxa de analfabetismo brasileira em 1980 era 25,41% e caiu para 9,62% em 2010 (Tabela 1). Todavia, ao excluir os *missings*, essas taxas subiram para 36,74% e 15,89%, respectivamente.

BARRO, R. J. Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, Oxford, v. 106, n. 2, p. 407-443, May 1991.

BAUMONT, C. *Spatial effects in housing price models: do house prices capitalize urban development policies in the agglomeration of Dijon*. Dijon: Université de Bourgogne; Pôle d'Economie et de Gestion, 1999. (Research Report).

BECKER, G. S. *Human capital: a theoretical and empirical analysis, with special reference to education*. New York: Columbia University, 1964.

BREUSCH, T. S.; PAGAN, A. R. The LM test and its applications to model specification in econometrics. *Review of Economic Studies*, London, v. 47, n. 1, p. 239-254, 1980.

CASTELLÓ-CLIMENT, A.; MUKHOPADHYAY, A. Mass education or a minority well educated elite in the process of growth: the case of India. *Journal of Development Economics*, Amsterdam, v. 105, p. 303-320, 2013.

CHRISTALLER, W. *Central places in Southern Germany*. New Jersey: Prentice-Hall, 1966.

CLIFF, A. D.; Ord, J. K. *Spatial processes: models and applications*. London: Pion, 1981.

EASTERLY, W.; LEVINE, R. Troubles with the neighbours: Africa's problem, Africa's opportunity. *Journal of African Economies*, Oxford, v. 7, n. 1, p. 120-142, 1998.

ERTUR, C.; KOCH, W. Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, Hamburg, v. 22, n. 6, p. 1033-1062, 2007.

FIRME, V. A. C.; SIMÃO FILHO, J. Análise do crescimento econômico dos municípios de Minas Gerais via modelo MRW (1992) com capital humano, condições de saúde e fatores espaciais, 1991-2000. *Revista de Economia Aplicada*, São Paulo, v. 18, n. 4, p. 679-716, 2014.

GALLO, J. L.; ERTUR, C. Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995. *Papers in Regional Science*, Azores, v. 82, n. 2, p.175-201, 2003.

GOLGHER, A. B.; VOSS, P. R. How to interpret the coefficients of spatial models: spillovers, direct and indirect effects. *Spatial Demography*, Sacramento, v. 4, n. 3, p. 175-205, 2016.

GROSSMAN G. M.; HELPMAN, E. *Innovation and growth in the global economy*. Cambridge: Massachusetts Institute of Technology, 1991.

GUSTAFSSON, M. A.; BERG, S. V. D.; SHEPHERD, D. L.; BURGER, C. *The costs of illiteracy in South Africa*. Stellenbosch: University of Stellenbosch, 2010. (Working Paper Economic, n. 14/10).

HAUSMAN, J. A. Specification tests in econometrics. *Econometrica*, New York, v. 46, n. 6, p. 1251-1271, Nov. 1978.

HSIAO, C. *Analysis of panel data*. 2 ed. Cambridge: Cambridge University, 2003.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Dados do CENSO 2011*. Disponível em: <https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>. Acesso em: 15 out. 2018.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – IPEA. *Ipeadata*. 2018. Disponível em: <http://www.ipeadata.gov.br/>. Acesso em: 15 out. 2018.

KELEJIAN, H.; PIRAS, G. *Spatial econometrics*. Cambridge: Academic Press, 2017.

KHALFAOUI, H. Islam and economic growth. *International Journal of Social Science Studies*, Ontario, v. 3, n. 6, p. 62-69, Nov. 2015.

KROTH, D.; DIAS, J. Os efeitos dos investimentos público e privado em capitais físico e humano sobre o produto per capita dos municípios da região sul: uma análise em painéis de dados dinâmicos. *Nova Economia*, Belo Horizonte, v. 22, n. 3, p. 621-650, set./dez. 2012.

KRUEGER, A. O. Factor endowments and per capita income differences among countries. *The Economic Journal*, Oxford, v. 78, n. 311, p. 641-659, Sept. 1968.

LEIJA, H. C. Efectos sociodemográficos en la pobreza municipal en México. Un estudio de econometría espacial. *Revista de Economía*, Mérida, v. 36, n. 92, p. 129-163, 2019.

LESAGE, J. P.; PACE, R. K. Interpreting spatial econometric models. In: FISCHER, M.; NIJKAMP, P. (ed.). *Handbook of regional science*. Berlin; Heidelberg: Springer, 2014. p. 1535-1552.

LIN, S. Education and economic development: evidence from China. *Comparative Economic Studies*, London, v. 39, n. 1, p. 66-85, 1997.

LUCAS JUNIOR, R. E. On the mechanics of economic development. *Journal of Monetary Economics*, Rochester, v. 22, n. 1, p. 3-42, 1988.

MANKIWI, N. G. *Macroeconomia*. 7. ed. Rio de Janeiro: LTC, 2010.

MANKIWI, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, Oxford, v. 107, n. 2, p. 407-437, May 1992.

MURTHY, V. N. R.; OKUNADE, A. A. Population health status and economic growth in Chinese provinces: some policy implications. *Applied Economic Letters*, London, v. 21, n. 6, p. 377-382, 2014.

MYRDAL, G. *Economic theory and under-developed regions*. London: Duckworth, 1957.

NEVES, M. F.; LIMA, A. C. C. Investimento em capital humano e retornos da educação nos mercados de trabalho brasileiros, 1991/2010. *Revista de Desenvolvimento Econômico*, Salvador, v. 1, n. 42, p. 76-107, abr. 2019.

OLIVEIRA, C. A. Externalidades espaciais e o crescimento econômico das cidades do estado do Ceará. *Revista Econômica do Nordeste*, Fortaleza, v. 36, n. 3, p. 319-337, jul. 2005.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO – PNUD. *Atlas do desenvolvimento humano no Brasil*, 2018. Disponível em: <http://atlasbrasil.org.br>. Acesso em: 15 out. 2018.

RAKODI, C. LLOYD-JONES, T. *Urban livelihoods: a people-centred approach to reducing poverty*. London: Earthscan Publications, 2002.

RAUCH, J. E. *The question of international convergence of per capita consumption: an Euler equation approach* San Diego: University of California, 1988. Mimeografado.

RENZI, A.; PAULA JÚNIOR, A.; FERREIRA, G. L. M.; PARRE, J. L. Capital humano: uma análise dos municípios paranaenses com econometria espacial. In: ENCONTRO DE ECONOMIA DA REGIÃO SUL – ANPEC SUL, 22., 2019, Maringá-PR. *Anais [...]*. Maringá-PR: Anpec, 2019.

ROMER, D. *Advanced macroeconomics*. New York: McGraw-Hill, 1996. (Advanced Series in Economics).

ROMER, P. M. Increasing returns and long-run growth. *Journal of Political Economy*, Chicago, v. 94, n. 5, p. 1002-1037, Oct. 1986.

ROMER, P. M. *Human capital and growth: theory and evidence*. Cambridge: National Bureau of Economic Research, 1989. (Working Paper, 3173).

ROMER, P. M. Endogenous technological change. *Journal of Political Economy*, Chicago, v. 98, n. 5, p. 71-102, Oct. 1990.

SCHULTZ, T. W. Investment in human capital. *The American Economic Review*, v. 51, n. 1, p. 1-17, Mar. 1961.

SCHULTZ, T. W. Investment in human beings. *Journal of Political Economy*, Chicago, v. 70, n. 5, p. 9-17, 1962.

SOARES, W. R. Diferenças salariais no Nordeste: uma análise via regressão quantílica. In: INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ (org.). *Economia do Ceará em debate*. Fortaleza: Ipece, 2009. p. 89-105.

SOLOW, R. M. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, Oxford, v. 70, n. 1, p. 65-94, Feb. 1956.

SURESH LAL, B. The economic and social cost of illiteracy: an overview. *International Journal of Advance Research and Innovative Ideas in Education*, Ahmedabad, v. 1, n. 5, p. 663-670, 2015.

VIEIRA, R. S. *Crescimento econômico no estado de São Paulo: uma análise espacial*. São Paulo: Editora Unesp; Cultura Acadêmica, 2009.

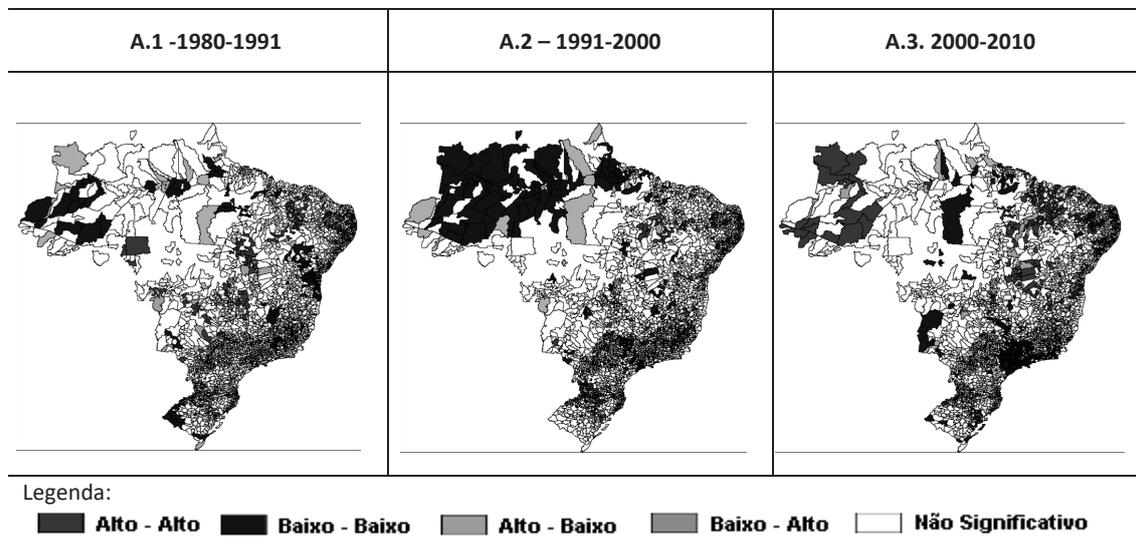
WOOLDRIDGE, J. M. *Econometric analysis of cross section and panel data*. Cambridge: Massachusetts Institute of Technology, 2002.

WORLD ECONOMIC OUTLOOK/INTERNATIONAL MONETARY FUND – WEO/IMF. *World Economic Outlook Reports*. 2018. Disponível em: <https://www.imf.org/en/publications/weo>. Acesso em: 15 out. 2018.

YANG, X.; BORLAND, J. A microeconomic mechanism for economic growth. *Journal of Political Economy*, Chicago, v. 99, n. 3, p. 460-482, June 1991.

ANEXO

FIGURA A - Mapa de *cluster* para o crescimento de renda *per capita* dos municípios brasileiros



Fonte: Elaboração do autor com base no *software* GEODA.

Nota: Este artigo considerou um painel balanceado, o que fez com que algumas informações fossem perdidas. No Censo do IBGE de 2010 (IBGE, 2011), o Brasil tinha 5.592 municípios. Todavia, só existem informações completas para 3.987 municípios, ou seja, 71,3% do total.

TABELA A.1 - Estatísticas descritivas das variáveis utilizadas

VARIÁVEL	DESCRIÇÃO	PERÍODO	MÉDIA	DESVIO PADRÃO	MÁXIMO	MÍNIMO
$[(Y_t - Y_{t-1}) / Y_{t-1}]$	Variação na renda domiciliar (%)	1980-1991	-14,04	28,42	1012,71	-75,98
		1991-2000	91,59	46,42	500,15	-82,45
		2000-2010	20,05	24,56	145,90	-62,99
Y_{t-1}	Renda domiciliar inicial (R\$ de 2000 - valor mensal)	1980	114,64	65,33	502,03	9,10
		1991	95,86	57,95	515,04	17,28
		2000	177,00	101,10	883,24	34,97
s_k	Gasto médio anual em capital fixo (R\$ de 2000 - <i>per capita</i>)	1985-1991	31,72	41,14	1473,00	2,19
		1992-2000	46,84	44,00	937,52	0,40
		2001-2010	75,60	51,17	989,44	3,09
s_h	Variação na taxa de alfabetização (%)	1980-1991	12,92	15,05	189,17	-40,49
		1991-2000	15,58	13,43	205,31	-2,97
		2000-2010	7,76	5,99	45,57	-13,12
$(n + g + \delta)$	Crescimento populacional (%), assumindo $g + \delta$ fixo	1980-1991	11,84	31,10	510,39	-86,76
		1991-2000	4,47	21,77	346,47	-75,30
		2000-2010	7,58	13,91	181,95	-45,63

Fonte: Elaboração do autor com base nos dados apresentados na seção “Base de dados”.

Nota: Estatísticas baseadas nos 3.987 municípios considerados nesta pesquisa; a renda domiciliar de 2010 foi R\$ 195,74 (valores de 2000); as taxas de alfabetização (ou analfabetos) de 1980, 1991, 2000 e 2010 foram, em sequência, 63,26% (36,74%), 69,74% (30,26%), 78,65 (21,35%) e 84,11% (15,89%).

Recebido em: 29 MAIO 2020

Aprovado para publicação em: 6 OUTUBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7705>

HETEROGENEIDADE DO IMPACTO DO TRABALHO INFANTIL NO RENDIMENTO ESCOLAR DO PARANÁ

 CARLOS ALBERTO RAMOS TORRES^I

 MARCOS DE OLIVEIRA GARCIAS^{II}

^I Universidade Federal do Paraná (UFPR), Curitiba-PR, Brasil; carlos.torres@ufpr.br

^{II} Universidade Federal da Integração Latino-Americana (Unila), Foz do Iguaçu-PR, Brasil; marcos.garcias@unila.edu.br

RESUMO

Neste estudo, estimou-se o impacto do trabalho infantil sobre o desempenho escolar em discentes do ensino fundamental público no estado do Paraná, avaliando a existência da heterogeneidade na magnitude do impacto. Para tanto, foram utilizados dados da Prova Brasil de 2007 a 2015, considerando os discentes da 5ª e 9ª séries. O impacto foi mensurado com o método Propensity Score Matching (PSM). Os resultados mostram que não houve mudanças temporais significativas nos níveis de trabalho infantil no estado nos anos avaliados, embora haja uma leve tendência de redução. Discentes que trabalham em casa e no mercado tendem a ter um pior desempenho acadêmico. Tal efeito negativo é menor em algumas mesorregiões, evidenciando que há heterogeneidade nesse impacto.

PALAVRAS-CHAVE TRABALHO INFANTIL • HETEROGENEIDADE • RENDIMENTO ESCOLAR • AVALIAÇÃO DO DESEMPENHO.

HETEROGENEIDAD DEL IMPACTO DEL TRABAJO INFANTIL EN EL RENDIMIENTO ESCOLAR DE PARANÁ

RESUMEN

En este estudio se estimó el impacto del trabajo infantil sobre el desempeño escolar en estudiantes de la educación básica pública en el estado de Paraná, evaluando la existencia de heterogeneidad en la magnitud del impacto. Para ello se utilizaron datos de la Prueba Brasil del 2007 al 2015, considerando los alumnos del 5° al 9° año. El impacto fue medido con el método Propensity Score Matching (PSM). Los resultados muestran que no hubo cambios temporales significativos en los niveles de trabajo infantil en el estado para los años evaluados, aunque se verificó una leve tendencia a la reducción. Estudiantes que trabajan en casa y en el mercado tienden a tener un peor desempeño académico. Tal efecto negativo es menor en algunas mesorregiones, lo que pone de manifiesto que hay heterogeneidad en este impacto.

PALABRAS CLAVE TRABAJO INFANTIL • HETEROGENEIDAD • RENDIMIENTO ESCOLAR • EVALUACIÓN DEL DESEMPEÑO.

HETEROGENEITY OF THE IMPACT OF CHILD LABOR ON SCHOOL PERFORMANCE IN PARANÁ

ABSTRACT

In this study we estimated the impact of child labor on school performance among elementary public-school students in the state of Paraná, assessing the role of heterogeneity in the magnitude of the impact. For this purpose, we used data from Prova Brasil from 2007 to 2015, concerning 5th and 9th grade students. The impact was measured with the Propensity Score Matching (PSM) method. Results show that there is no evidence of significant temporal changes in child labor levels in the state, although a slight tendency to its reduction is observed. Students who work at home and in the market tend to have a worse academic performance. Such negative effect is less clear in some mesoregions, proving that this impact is heterogeneous.

KEYWORDS CHILD LABOR • HETEROGENEITY • SCHOOL PERFORMANCE • SCHOOL PERFORMANCE ASSESSMENT.

INTRODUÇÃO

Há um evidente consenso de que o trabalho infantil deve ser considerado um problema social persistente, principalmente nas economias em desenvolvimento. Com base nas principais convenções internacionais,¹ o trabalho infantil pode ser definido como a participação em atividades econômicas, com ou sem remuneração, de crianças e adolescentes abaixo da idade mínima de admissão ao emprego/trabalho estabelecida no país. Especificamente no caso brasileiro, considera-se trabalho infantil quando a atividade é realizada por crianças ou adolescentes com idade inferior a 16 anos, exceto para a condição de aprendiz, em que a idade mínima permitida é de 14 anos.

No mundo, segundo dados da OIT (2013), estima-se que, em 2013, havia 168 milhões de crianças e adolescentes ocupados economicamente. Na região da América Latina e Caribe (ALC), evidenciou-se um avanço significativo na redução e combate ao trabalho infantil nas últimas duas décadas. Os dados da OIT (2013) mostram que a taxa de crianças e adolescentes ocupados em atividades econômicas foi reduzida de 14,9%, em 2002, a 10,2%, em 2012.

Isso pode ser um reflexo do último ciclo político na primeira década do século XXI e das diversas iniciativas de política social na região. No entanto, a mesma OIT (2015) reporta que, na ALC, no ano de 2015, ainda havia aproximadamente 12,5 milhões de crianças e adolescentes em condição de trabalho infantil. Especificamente no Brasil, segundo dados da OIT (2016) considerando a Pesquisa Nacional de Amostras e Domicílios (Pnad), entre 1992 e 2015, aproximadamente 5,7 milhões de crianças e adolescentes de 5 a 17 anos deixaram de trabalhar, o que significou uma redução de 68%. Entretanto, ainda há 2,7 milhões de crianças e adolescentes em situação de trabalho infantil no país, dos quais aproximadamente 160 mil estão situados no estado do Paraná.

As causas possíveis do fenômeno radicam-se em fatores como: a pobreza estrutural das famílias, determinantes culturais, nível de escolaridade dos pais, nível de desigualdade econômica, grau de informalidade da economia, falta de legislação, entre outros (BASU; TZANNATOS, 2003; KASSOUF, 2007). No que diz respeito às possíveis consequências, podem-se destacar: o risco de efeitos negativos na saúde física e mental do infante (AMBADEKAR *et al.*, 1999; ARANSIOLA; JUSTUS, 2018); a tendência a salários menores no futuro (KASSOUF; SANTOS, 2010) e consequências negativas no processo de escolarização (GALLI, 2001). Esta última consequência representa a principal preocupação desta pesquisa.

1 Particularmente as celebradas pela Organização Internacional do Trabalho (OIT): a Convenção 138 “Sobre a Idade Mínima para Admissão” (OIT, 1973), e a Convenção 182, “Sobre Proibição das Piores Formas de Trabalho Infantil e Ação Imediata para sua Eliminação” (OIT, 1999). No entanto, como todo tratado de Direito Internacional, sua adoção é facultativa para cada país conforme sua legislação interna.

Diversos estudos mostram que a educação básica é fundamental no processo de formação de capital humano, fator relevante para o crescimento e sofisticação de uma economia (HANUSHEK, 2013; GLEWWE; MURALIDHARAN, 2016; MERCAN; SEZER, 2014). Os efeitos negativos do trabalho infantil sobre a escolaridade podem interromper e afetar esse processo; no entanto, é necessário avaliar com cautela o impacto, levando em conta diversas especificidades sobre o tipo de trabalho ou sua intensidade. Normalmente os estudos na área são realizados no âmbito nacional (BEZERRA; KASSOUF; ARENDS-KUENNING, 2009; DELPRATO; AKYEAMPONG, 2019; EMERSON; PONCZEK; SOUZA, 2017; KASSOUF; TIBERTI; GARCIAS, 2020) e a respeito das causas (BASU, 2000; FORS, 2012), e não das consequências do trabalho infantil.

Esse fato justifica a relevância da presente pesquisa; ademais, o método adotado tem o potencial de representar uma inovação teórica. Isso porque a desagregação dos dados no âmbito regional permite observar especificidades e detectar diferenças de resultados obtidos em análises no contexto de países. O controle das estimações por mesorregiões poderia revelar a existência de heterogeneidade na magnitude do impacto do trabalho infantil sobre a educação.

Dito isso, o objetivo geral do estudo é estimar o impacto do trabalho infantil sobre o rendimento escolar dos estudantes do ensino básico do estado do Paraná, no Brasil, no ano de 2015, considerando os tipos de trabalho: doméstico, no mercado ou ambos.

Pretende-se, adicionalmente, identificar a evolução temporal do número de crianças trabalhando entre 2007 e 2015, nas diferentes modalidades, e detectar a existência de heterogeneidade na magnitude do impacto do trabalho infantil, controlando as estimações por cada uma das 10 mesorregiões paranaenses, levando em conta as diferenças estruturais referentes à composição setorial e o tipo de atividade econômica predominante.

TRABALHO INFANTIL E ESCOLARIDADE

Existem alguns estudos que analisam o efeito do trabalho realizado pelas crianças sobre a sua taxa de frequência ou sua matrícula escolar, e não sobre o seu desempenho na escola, como Ravallion e Wodon (2000), Canals-Cerdá e Ridao-Cano (2004), Beegle *et al.* (2008) e Putnick e Bornstein (2015). Entretanto, no Brasil, é comum ver crianças combinando trabalho e escola; nesse sentido, observar o efeito do trabalho infantil na aprendizagem e no rendimento escolar pode oferecer uma análise mais robusta e detalhada.

Autores como Patrinos e Psacharopoulos (1997), Akabayashi e Psacharopoulos (1999), Heady (2003), Stinebrickner e Stinebrickner (2003), Gunnarsson, Orazem e

Sánchez (2006), Bezerra, Kassouf e Arends-Kuenning (2009), Dumas (2012), Delprato e Akyeampong (2019), Kassouf, Tiberti e Garcias (2020), entre outros, estudaram o efeito do trabalho infantil sobre resultados dos testes de desempenho dos alunos em diferentes países.

O impacto do trabalho infantil na aprendizagem pode ser negativo se as crianças e os adolescentes dividirem seu tempo entre estudar e trabalhar muitas horas em empregos que exijam muito esforço físico e mental. Por outro lado, o impacto do trabalho infantil na aprendizagem pode ser positivo se o trabalho envolver tarefas que resultem em aprendizado e melhoria de habilidades cognitivas. Portanto, a direção esperada do impacto do trabalho infantil no desempenho escolar não é clara.

No campo teórico, existe controvérsia sobre as formas e tipos de trabalho infantil, assim como as causas e consequências do fenômeno. Por exemplo, no caso do trabalho doméstico – considerado uma das piores formas de trabalho infantil pela Convenção 182 da OIT (1999)² –, algumas pesquisas mostram que pode ser prejudicial para o infante, e, em alguns casos, as condições de trabalho podem ser análogas à escravidão (BLAGBROUGH, 2008; BOURDILLON, 2009; KASSOUF, 2007). Outras pesquisas, por sua vez, apontam que o trabalho infantil doméstico, além de ser heterogêneo, sob certas condições pode ser visto como “seguro”, e o fato de ser classificado como “pior forma de trabalho infantil” pode trazer consequências negativas para as crianças (GAMLIN *et al.*, 2015; KLOCKER, 2011).

No âmbito do trabalho infantil rural, é bastante frequente a presença de crianças e adolescentes que trabalham em atividades agropecuárias de caráter doméstico sem remuneração. No estudo de Admassie (2003), é evidenciado que, no setor rural, muitas famílias precisam da mão de obra infantil para o sustento do núcleo, e, muitas vezes, existe a compatibilidade entre trabalho e estudo sem consequências maiores para o desempenho escolar. No entanto, é necessário avaliar diversos fatores que influenciam esse processo, tais como o tipo de trabalho, a temporada na qual é desenvolvido (pelas safras e aumento de demanda de mão de obra), a composição familiar etc.

Por outro lado, Kassouf, Tiberti e Garcias (2020) avaliaram o impacto do trabalho infantil (doméstico e no mercado) sobre o rendimento escolar usando os dados da Prova Brasil dos anos de 2007 a 2017. Os resultados encontrados pelos autores mostram que, no Brasil, tanto o trabalho realizado pelas crianças em casa como o realizado no mercado afetam negativamente o rendimento escolar, sendo piores os efeitos naquelas crianças que trabalham apenas no mercado.

2 No Brasil, com a ratificação da Convenção 182 da OIT em 2008, o trabalho doméstico passou a ser considerado uma das piores formas de trabalho infantil.

Na revisão da literatura, é possível perceber que a maioria dos estudos sobre trabalho infantil é realizada no âmbito nacional, e uma boa parte é sobre as causas, e não as consequências do fenômeno. Também no contexto da avaliação de impacto do trabalho infantil sobre o rendimento escolar, as estimativas são, normalmente, baseadas em dados agregados no âmbito nacional. Dessa maneira, os resultados estimados podem omitir algumas especificidades regionais que, por diversas razões, podem ter influência na magnitude e direção dos efeitos analisados.

A implicação prática disso é que podem existir importantes diferenças entre o impacto estimado do trabalho infantil no desempenho escolar de um país e o impacto estimado em alguma das suas regiões específicas. De igual forma, dentro da mesma região, podem existir importantes diferenças na magnitude do impacto, configurando-se, assim, uma possível heterogeneidade do impacto do trabalho infantil sobre o rendimento escolar.

O TRABALHO INFANTIL NO PARANÁ

Há uma importante lacuna de trabalhos qualitativo-descritivos e/ou empírico-quantitativos que visam a medir o impacto do trabalho infantil sobre o rendimento escolar, considerando, especificamente, o caso do estado do Paraná. Contudo, existem alguns estudos relacionados, como o de Lopes, Souza e Pontili (2008), que, mediante o uso de dados da Pnad, procurou mostrar os efeitos do trabalho infantil sobre o nível de escolaridade e rendimentos futuros dos indivíduos no Paraná. O resultado foi um efeito negativo, evidenciando o fato de que pessoas inseridas no mercado de trabalho paranaense depois dos seus 18 anos de idade (seja no âmbito rural ou urbano) tendem a apresentar maiores níveis de escolaridade e, consequentemente, maiores salários do que aqueles indivíduos inseridos no trabalho durante a infância.

Por outro lado, Mota (2003) discute a evolução da problemática do trabalho infantil no estado do Paraná, destacando os precedentes políticos, econômicos e sociais, bem como as principais iniciativas de combate ao trabalho infantil no período 1992-2001. A análise mostra que, no Paraná, algumas famílias necessitam economicamente da incorporação precoce dos jovens no mercado de trabalho. No entanto, evidencia-se também uma redução significativa nos índices de trabalho infantil nos anos considerados.

Outrossim, o Instituto Paranaense de Desenvolvimento Econômico e Social (Ipardes) desenvolveu o *Mapa do trabalho infanto-juvenil no Paraná* (IPARDES, 2007), uma análise especializada no âmbito de municípios e microrregiões, apresentando uma série de indicadores que caracterizam o trabalho infantil no estado de

forma específica. Com base nessa metodologia e incorporando dados da Pnad para o ano de 2010, a Secretaria da Família e Desenvolvimento Social (Sedes) do estado do Paraná realizou um estudo mostrando que os casos de trabalho infantil no Paraná estão espalhados entre regiões, não havendo um único foco, mas pequenos contingentes que apresentam alto porcentual de infantes ocupados (SEDES, 2017).

Navarro, de Souza e Evarini (2013), a partir de microdados da Pnad, puderam analisar o perfil das crianças trabalhadoras e as condições de sua inserção no mercado de trabalho no Paraná no período 2002-2009. Os resultados mostram que as crianças ocupadas, em sua maioria, são do sexo masculino, brancas, alfabetizadas e moradoras da área urbana. A prática de tarefas domésticas é comum, a escolaridade da mãe tende a ser baixa, e a ausência da figura materna nos domicílios é crescente no período analisado. Os infantes começam a trabalhar ao redor dos 10 anos de idade, e o rendimento mensal do trabalho e a renda domiciliar *per capita* tendem a ser baixos.

Por último, Souza (2018) realizou uma pesquisa que almejou analisar o comportamento do trabalho infantil em relação às medidas institucionais para seu combate em toda a região Sul do Brasil. No tocante ao trabalho precoce, entre as principais consequências está a sua associação negativa com a escolaridade dos indivíduos. Constatou-se que o abandono escolar na região Sul foi mais elevado entre as crianças de 5 a 6 anos e entre os adolescentes de 15 a 17 anos, especialmente entre os mais velhos, acima de 16 anos. No total, 5,9% das crianças e adolescentes da região, de 5 a 17 anos, estavam em condição de abandono escolar (destes, 40% trabalhavam). Após políticas de combate ao trabalho de menores nessa região, a taxa de trabalho infantil caiu significativamente, e, em contrapartida, a taxa de matrícula escolar aumentou.

Diante da escassez de estudos no contexto regional que versem, especificamente, sobre os impactos do trabalho infantil no desempenho escolar, almejou-se avaliar o possível impacto no rendimento escolar dos alunos do estado do Paraná. Para isso, foram utilizados os dados disponíveis na Prova Brasil³ para a 5ª e 9ª séries das edições de 2007 a 2015 (BRASIL, 2007, 2009, 2011, 2013, 2015), considerando os diferentes tipos de trabalho para os gêneros masculino e feminino.

De igual forma, além de comparar os resultados com os obtidos em estudos nacionais, pretende-se avaliar nas diferentes mesorregiões a possibilidade de existir heterogeneidade na magnitude do impacto do trabalho infantil sobre o rendimento escolar paranaense. Os resultados encontrados, além de contribuir com a

3 A Prova Brasil é um conjunto de dados do censo para discentes da 5ª e 9ª séries das escolas públicas urbanas, sendo aplicada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), do Ministério da Educação, a cada dois anos. Cabe salientar que, desde 2009, a Prova Brasil inclui escolas rurais na amostra.

literatura específica, poderão auxiliar no desenho de projetos sociais, bem como nas tomadas de decisões políticas, sejam elas educacionais ou de combate ao trabalho infantil.

MATERIAIS E MÉTODOS

Para mensurar o impacto do trabalho infantil sobre o rendimento escolar paranaense, objetivo deste trabalho, foi utilizado o método Propensity Score Matching (PSM), considerando os tipos de trabalho infantil: doméstico, no mercado e ambos.⁴ Para detectar a existência de heterogeneidade na magnitude do impacto, as estimações do PSM serão controladas por mesorregiões do Paraná nos resultados do ano de 2015.

O PSM é um método de pareamento que permite comparar dois grupos: um chamado de grupo de tratamento (alunos que trabalham, nas diferentes formas), e o outro, chamado grupo de controle (alunos que não trabalham). Assim, pode-se avaliar o impacto do trabalho infantil no rendimento escolar para as crianças/jovens que trabalham e qual seria o resultado caso elas não trabalhassem. Por meio dessa hipótese, é possível avaliar um indivíduo no grupo de tratamento com as mesmas características “ χ ” que um indivíduo no grupo de controle. Por exemplo: compara-se a média na Prova Brasil de João, que tem olhos verdes, dois irmãos e trabalha, com a de Joaquim, que também tem olhos verdes, dois irmãos, porém não trabalha. Dessa maneira, avalia-se um aluno no grupo tratado que tenha as mesmas características de um aluno no grupo de controle, sendo a única diferença entre os estudantes o fato de trabalhar ou não (DEHEJIA; WAHBA, 1999).

Dada a dificuldade de encontrar pares com as mesmas características (por exemplo, ter olhos verdes e dois irmãos), tem-se o problema da multidimensionalidade.⁵ Para resolver esse problema, calcula-se o escore de propensão (PS) utilizando-se um modelo lógite (GREENE, 2003), que, por meio de variáveis selecionadas, indica a probabilidade de a criança trabalhar ou não; logo, são comparados indivíduos com probabilidades similares. Para o cálculo, foi utilizado o seguinte modelo:

$$Y = \alpha + \beta_1sex + \beta_2mulatto + \beta_3black + \beta_4asian + \beta_5moth_read + \beta_6fath_read + \beta_7hmember + \beta_8RENDA + \beta_9motivate_study + u_i$$

O Quadro 1 mostra a descrição das variáveis do modelo. É importante salientar que a maior parte dos discentes do Paraná é branca; por isso, a decisão metodológica de usar tal critério como categoria base para as variáveis *dummies* da raça.

4 “Trabalho doméstico” refere-se à atividade desenvolvida por mais de duas horas em casa. “Trabalho no mercado” refere-se às atividades laborais desenvolvidas fora de casa. “Ambos” representa a combinação das duas formas.

5 Conforme aumenta o número de características, diminuem as chances de achar um indivíduo igual a outro.

Outras variáveis, referentes às características da escola, foram omitidas na especificação do modelo por excesso de *missing* nas observações.

QUADRO 1 - Variáveis utilizadas para a construção do modelo lógite

VARIÁVEL	DESCRIÇÃO
<i>Y</i>	Variável dependente: assume valor de 1 se o estudante trabalha (em casa, no mercado ou em ambos), e de 0, caso contrário. Os β_n são os parâmetros estimados para cada variável explicativa, sendo u_i o termo de erro.
<i>sex</i>	Identifica o sexo do aluno: masculino recebe valor de 1, e feminino, de 0.
<i>mulatto</i>	Recebe valor de 1 se o aluno é de raça parda, e de 0 se o aluno é de raça branca.
<i>black</i>	Recebe valor de 1 se o aluno é de raça negra, e de 0 se o aluno é de raça branca.
<i>asian</i>	Recebe valor de 1 se o aluno é de raça amarela, e de 0 se o aluno é de raça branca.
<i>moth_read</i>	É uma variável <i>proxy</i> para escolaridade da mãe; recebe valor de 1 se a mãe do aluno sabe ler, e de 0, caso contrário.
<i>fath_read</i>	É uma variável <i>proxy</i> para escolaridade do pai; recebe valor de 1 se o pai do aluno sabe ler, e de 0, caso contrário.
<i>hhmember</i>	Representa o número de pessoas que residem com o aluno, incluindo ele mesmo.
<i>RENDA</i>	Representa uma <i>proxy</i> do <i>background</i> socioeconômico do discente. Foi construída a partir do “Critério de Classificação Econômica do Brasil” ⁶ com base nos dados do questionário da Prova Brasil.
<i>motivate_study</i>	Recebe o valor de 1 se os pais motivam o discente a estudar e fazer suas tarefas, e de 0, caso contrário. Os fatores motivacionais são importantes, porque sua ausência na análise poderia superestimar os efeitos do trabalho infantil.

Fonte: Elaboração própria com base em dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Para o cálculo de diferença de médias de quem trabalha com quem não trabalha, chamado de Average Treatment Effect on the Treated (ATT), é necessário que haja os pares entre os indivíduos dos grupos. Na literatura, são propostos diferentes métodos que têm por finalidade efetuar o pareamento do escore de propensão (CALIENDO; KOPEINIG, 2008; MAFFIOLI, 2011), ou seja, a seleção e comparação dos indivíduos presentes nos grupos de tratamento e de controle. Para isso, cada método atribui determinado peso a cada grupo. Neste trabalho, os resultados estimados foram obtidos pelos seguintes métodos de pareamento: sem restrição (*noreplacement*), 5 vizinhos e kernel.

Para a análise da heterogeneidade do impacto do trabalho infantil, será estimado o impacto mediante o método PSM para cada uma das dez mesorregiões do estado do Paraná. Da mesma forma, será utilizado o mapa do trabalho infantil do Ipardes (2007) para classificar o tipo de atividade econômica que predomina em cada mesorregião e suas respectivas microrregiões e, assim, poder comparar os efeitos segundo o tipo de atividade. A classificação regional utilizada será a descrita no Quadro 2.

6 Tal critério classifica em classes econômicas a partir de uma pontuação obtida considerando os bens domésticos reportados no questionário da prova pelo discente.

QUADRO 2 – Mesorregiões e microrregiões do estado do Paraná

NÚMERO	MESORREGIÃO	MICRORREGIÕES
1	Noroeste Paranaense	Paranavaí, Umuarama, Cianorte
2	Centro-Occidental Paranaense	Goioerê, Campo Mourão
3	Central Paranaense	Astorga, Porecatu, Florai, Maringá, Apucarana, Londrina, Faxinal, Ivaiporã
4	Norte Pioneiro	Assaí, Cornélio Procópio, Jacarezinho, Ibaiti, Wenceslau Braz
5	Centro-Oriental	Telêmaco Borba, Jaguariaíva, Ponta Grossa
6	Oeste Paranaense	Toledo, Cascavel, Foz do Iguaçu
7	Sudoeste Paranaense	Capanema, Francisco Beltrão, Pato Branco
8	Centro-Sul Paranaense	Pitanga, Guarapuava, Palmas
9	Sudeste Paranaense	Prudentópolis, Irati, União da Vitória, São Mateus do Sul
10	Metropolitana de Curitiba	Cerro Azul, Lapa, Curitiba, Paranaguá, Rio Negro

Fonte: Elaboração própria com dados do Instituto Brasileiro de Geografia e Estatística (IBGE, 2017).

A base de dados utilizada é a da Prova Brasil, censo que mensura o rendimento dos alunos das escolas públicas urbanas brasileiras nas disciplinas de Língua Portuguesa e Matemática. Para a análise descritiva, foram considerados os anos de 2007, 2009, 2011, 2013 e 2015 para discentes da 5ª e 9ª séries. Para a mensuração do impacto sobre o rendimento, foram utilizados apenas os dados de 2015.

RESULTADOS

Os resultados aqui apresentados estão divididos em duas partes. Na primeira, são discutidas as estatísticas descritivas dos resultados do Paraná na Prova Brasil, utilizando os dados de 2007, 2009, 2011, 2013 e 2015. Na segunda parte, são apresentados os impactos do trabalho infantil, mensurados utilizando-se o PSM, nas diferentes mesorregiões para o ano de 2015.

Estatística descritiva

Na Tabela 1, apresenta-se um panorama geral da evolução na situação laboral dos discentes, considerando como critérios o sexo e a série. Para cada ano da Prova Brasil, foi tabulada a porcentagem de estudantes em relação ao total.

No caso dos meninos da 5ª série, observa-se que a porcentagem média dos ocupados (em casa, no mercado ou em ambos) entre 2007 e 2015 foi de 40,97%. No caso do trabalho doméstico, observa-se um incremento significativo de 26,20%, em 2007, para 38,48%, em 2009. Apesar de uma redução em 2011 (para 23,67%), essa modalidade de trabalho infantil situou-se praticamente nos níveis iniciais até 2015 (26,66%). Por outro lado, em 2007, os ocupados no mercado de trabalho foram

9,36%, diminuindo para 3,07%, em 2009, e voltando para 9,11%, em 2011. Ou seja, não houve mudanças temporais muito significativas, mas é importante destacar que a média de alunos ocupados em trabalho doméstico se situa acima dos níveis nacionais detectados por Kassouf, Tiberti e Garcias (2020).

Quando observados os meninos da 9ª série, percebe-se que a porcentagem média dos ocupados entre 2007 e 2015 foi de 52,02%. Em 2007, a porcentagem de alunos no trabalho doméstico foi de 21,6%; em 2009, ocorreu uma leve queda para 20,23%, que persistiu em 2011 (20,42%), aumentando para 27,46%, em 2013, e 30,96%, em 2015. No caso dos que trabalham no mercado, a porcentagem foi de 21,26%, em 2007, registrando uma tendência de queda a partir de 2011, com 21,59%, chegando a 13,86%, em 2015. A redução observada nos níveis de trabalho infantil no mercado coincide com diversas políticas nacionais e regionais de combate e fiscalização no Brasil nas últimas décadas (DE BARROS; MENDONÇA, 2010; KASSOUF, 2015; SOUZA, 2018).

No caso das meninas da 5ª série, a porcentagem média das ocupadas entre 2007 e 2015 foi de 46,73%. No mesmo intervalo, os dados do trabalho doméstico apresentam uma porcentagem média de 40,31%, evidenciando-se um aumento significativo de 2007 para 2013. Já em relação às que trabalham no mercado, as porcentagens médias não apresentaram mudanças temporais significativas, passando de 3%, em 2007, para 2,43%, em 2015. Novamente, observa-se que o aumento, ao longo do tempo, do trabalho em casa é tendência, e a taxa de crianças do sexo feminino ocupadas é maior se comparada à dos discentes do sexo masculino da mesma série.

A porcentagem média das meninas ocupadas da 9ª série, entre 2007 e 2015, foi de 68,69%. O trabalho doméstico observado teve porcentagem média de 53,98%, em 2007, aumentando para 61,17%, em 2013, e 63,35%, em 2015, o que representa os maiores níveis em relação às demais categorias analisadas. Considerando as ocupadas apenas no mercado de trabalho, a porcentagem foi reduzida de 4,42%, em 2007, para 3,46%, em 2013, e 2,64%, em 2015, sem importantes mudanças temporais. Quando analisados os estudantes que trabalham em casa e no mercado, sem considerar o gênero, a porcentagem é mais elevada para discentes da 9ª série do que para os da 5ª série.

De modo geral, parece haver uma tendência de que as meninas trabalhem mais em casa, e os meninos, no mercado. Particularmente no caso das alunas da 9ª série, a porcentagem média das que trabalham em casa de 2007 a 2015 foi de 56,04%, mais do que o dobro se comparada com a média dos meninos da 9ª série, que foi de 24,13% para o mesmo período. Esse fato coincide com os achados dos trabalhos de Alberto *et al.* (2009), Alvi e Dendir (2011), Haile e Haile (2012) e Frasco-Zuker (2016), em que a “especialização” do trabalho é marcada pela propensão das

meninas ao trabalho doméstico e dos meninos a tarefas de mercado. Outro detalhe importante é que a porcentagem média de estudantes ocupados, entre 2007 e 2015, foi mais alta no caso das meninas em comparação à dos meninos.

TABELA 1 – Situação laboral dos alunos da 5ª e 9ª séries no período 2007-2015, por gênero

SITUAÇÃO	MASCULINO 5ª SÉRIE (%)					MASCULINO 9ª SÉRIE (%)				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não trabalha	57,89	54,84	61,39	59,39	61,63	49,39	48,5	49,77	45,33	46,87
Trabalha em casa	26,2	38,48	23,67	25,14	26,66	21,6	20,23	20,42	27,46	30,96
Trabalha no mercado	9,36	3,07	9,11	9,42	7,11	21,26	22,27	21,59	17,74	13,86
Trabalha em ambos	6,55	3,61	5,83	6,04	4,6	7,75	9	8,23	9,47	8,32
Total	100	100	100	100	100	100	100	100	100	100
SITUAÇÃO	FEMININO 5ª SÉRIE (%)					FEMININO 9ª SÉRIE (%)				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não trabalha	52,73	54,84	54,2	50,89	53,69	32,9	34,76	35,33	26,48	27,07
Trabalha em casa	40,46	38,48	39,4	42,29	40,94	53,98	52,18	51,84	61,17	63,35
Trabalha no mercado	3	3,07	2,92	3,16	2,43	4,42	4,78	4,71	3,46	2,64
Trabalha em ambos	3,81	3,61	3,48	3,66	2,93	8,71	8,28	8,12	8,9	6,94
Total	100	100	100	100	100	100	100	100	100	100

Fonte: Elaboração própria com base em dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Em relação ao desempenho escolar, pode ser útil começar mapeando alguns dados sobre reprovações e abandono da escola. Segundo a evidência, em média, a porcentagem de estudantes que reprovaram, pelo menos um grau, no período 2007-2015, foi de 24,67% para a 5ª série e de 33,39% para a 9ª série. Já a porcentagem média de estudantes que abandonaram pelo menos uma vez a escola foi de 4,62% para a 5ª série e de 5,85% para a 9ª série. A tendência sugere que é mais comum observar problemas de desempenho nas crianças de maior idade, no caso, da 9ª série; uma tendência parecida também foi observada em outros casos da literatura internacional (ALVI; DENDIR, 2011; DELPRATO; AKYEAMPONG, 2019). A porcentagem de reprovação e abandono dos discentes paranaenses da 5ª série apresentou resultado inferior ao nível observado no Brasil por Kassouf, Tiberti e Garcias (2020) e Kassouf *et al.* (2016); no caso dos discentes da 9ª série, a porcentagem foi similar e levemente superior em alguns anos. A Tabela 2 mostra detalhadamente a evolução dessas porcentagens em cada ano da prova.

TABELA 2 – Porcentagem de alunos da 5ª e 9ª séries que reprovaram pelo menos uma série no passado ou abandonaram a escola

REPROVADO	5ª SÉRIE (%)					9ª SÉRIE (%)				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não	73,66	72,77	73,26	79,21	77,73	65,81	65,99	66,72	69,25	65,24
Sim	26,34	27,23	26,74	20,79	22,27	34,19	34,01	33,28	30,75	34,76
Total	100	100	100	100	100	100	100	100	100	100
ABANDONOU	5ª SÉRIE (%)					9ª SÉRIE (%)				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não	95,04	95,49	94,97	95,67	95,7	93,22	93,97	94,09	94,88	94,59
Sim	4,96	4,51	5,03	4,33	4,30	6,78	6,03	5,91	5,12	5,41
Total	100	100	100	100	100	100	100	100	100	100

Fonte: Elaboração própria com base em dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Alguns estudos apontam o fato de que delegar tarefas domésticas às crianças pode fomentar responsabilidade e confiança, impactando positivamente na idade adulta (FRENCH, 2002; ROSSMAN, 2002). No entanto, sobrecarregar a criança com tais tarefas pode ocasionar prejuízos no seu desempenho acadêmico (ROSSMAN, 2002). Nesse sentido, para fins da pesquisa, considera-se que a criança trabalha em casa quando realiza atividades domésticas pelo menos duas horas ou mais por dia. Assim, na Tabela 3, apresenta-se o *score* médio, ou o rendimento médio, dos alunos da 5ª e 9ª séries nas suas provas de Português e Matemática, considerando-se, para isso, sua situação laboral.

A priori, os resultados parecem mostrar que os discentes que trabalham em casa e no mercado tendem a ter um pior desempenho acadêmico em comparação àqueles que não trabalham ou trabalham somente em casa. O mesmo pode-se dizer sobre aqueles que trabalham só no mercado. Particularmente, os discentes da 9ª série que trabalham em casa apresentaram um rendimento levemente mais alto em Português em comparação àqueles que não trabalham; isso se constata para os anos de 2009, 2011, 2013 e 2015. Esse achado sugere uma primeira divergência dos resultados obtidos no âmbito nacional nos estudos de Salvador (2016) e de Kassouf, Tiberti e Garcias (2020) e reforça a ideia de que em contextos específicos o trabalho infantil doméstico não é necessariamente prejudicial para o desempenho escolar (ADMASSIE, 2003; GAMLIN *et al.*, 2015; KLOCKER, 2011).

TABELA 3 – Média do rendimento de alunos da 5ª e 9ª séries, nas provas de Português e Matemática, considerando sua situação laboral

SITUAÇÃO	SCORE MÉDIO EM MATEMÁTICA, 5ª SÉRIE					SCORE MÉDIO EM MATEMÁTICA, 9ª SÉRIE				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não trabalha	190,04	199,11	203,67	213,91	221,68	238,9	249,61	248,01	244,7	250,56
Trabalha em casa	184,57	191,28	195,47	211,11	218,9	238,87	251,11	250,11	251,73	258,09
Trabalha no mercado	170,07	176,48	179,71	183,26	195,22	223,86	232,77	229,86	226,83	232,98
Trabalha em ambos	168,42	173,21	175,59	182,01	194,81	226,06	236,17	234,09	236,03	242,26
SITUAÇÃO	SCORE MÉDIO EM PORTUGUÊS, 5ª SÉRIE					SCORE MÉDIO EM PORTUGUÊS, 9ª SÉRIE				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Não trabalha	210,67	224,45	228,13	234,37	236,66	256,74	254,89	257,4	252,62	257,38
Trabalha em casa	203,7	215,46	217,68	229,51	233,13	250,76	249,97	253,96	253,8	257,55
Trabalha no mercado	195,6	206,47	210,32	210,23	217,52	249,78	248,21	248,94	243,93	250,53
Trabalha em ambos	192,01	202,06	202,7	205,03	215,02	242,58	242,78	245,17	245,05	250,73

Fonte: Elaboração própria com base em dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Seguidamente, na Tabela 4, são mostrados os dados correspondentes ao desempenho dos alunos considerando o gênero.

TABELA 4 – Média do rendimento de alunos da 5ª e 9ª séries, nas provas de Português e Matemática, considerando o gênero

SEXO	SCORE MÉDIO EM PORTUGUÊS, 5ª SÉRIE					SCORE MÉDIO EM MATEMÁTICA, 5ª SÉRIE				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Feminino	190,24	198,38	204,03	214,55	222,73	204,56	217,55	219,93	227,55	231,08
Masculino	180,95	189,23	192,15	204,2	214,27	207,65	220,7	224,23	231,66	236,37
SEXO	SCORE MÉDIO EM PORTUGUÊS, 9ª SÉRIE					SCORE MÉDIO EM MATEMÁTICA, 9ª SÉRIE				
	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015
Feminino	241,94	253,24	252,99	254,11	259,38	248,1	246,29	251,06	250,11	251,77
Masculino	229,16	238,8	236,91	235,85	244,29	257,17	255,7	256,96	252,99	260,6

Fonte: Elaboração própria com base em dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

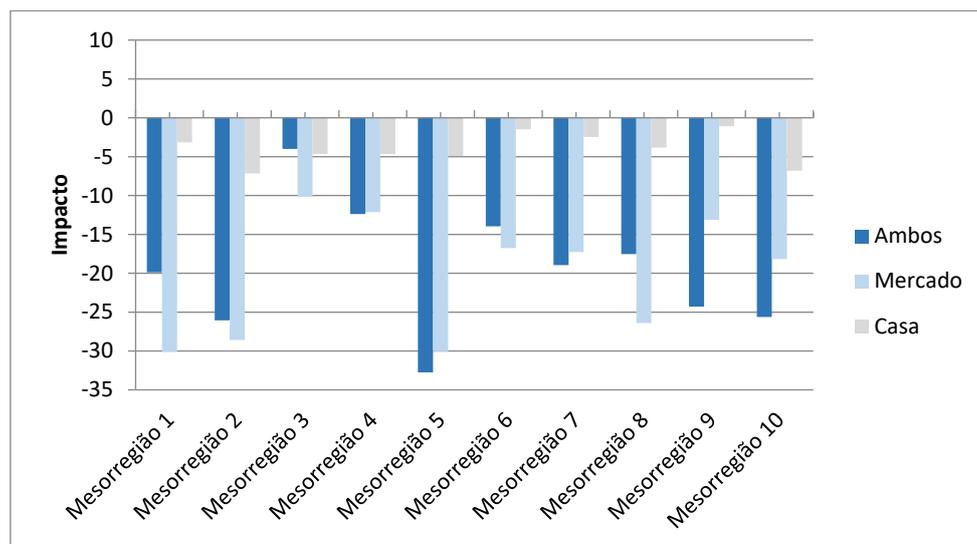
Para as meninas da 9ª série, a evidência mostra que seu rendimento nas provas de Português é melhor do que o rendimento dos meninos da mesma série; o contrário ocorre nas provas de Matemática, em que os meninos apresentam melhor desempenho. No caso dos estudantes masculinos da 5ª série, repete-se a tendência anterior: melhor desempenho nas provas de Matemática e pior desempenho nas provas de Português em comparação com as estudantes do sexo feminino.

Impactos e heterogeneidade do trabalho infantil no Paraná

Mediante os resultados obtidos por meio do método PSM,⁷ controlando as variáveis por mesorregiões do estado do Paraná, e com auxílio do mapa da composição setorial do trabalho infantil, podem-se observar algumas divergências em relação à magnitude do efeito do trabalho infantil sobre o rendimento escolar. As estimativas, nos diferentes tipos de pareamento, atenderam à hipótese de suporte comum, ou seja, foi possível comparar, no grupo de tratamento e no grupo de controle, indivíduos com probabilidades similares de trabalhar.

A Figura 1 apresenta o efeito médio do trabalho infantil sobre a proficiência em Língua Portuguesa para discentes da 5ª série nas dez mesorregiões do Paraná, considerando também o *status* ocupacional do indivíduo para o ano de 2015. Já a Figura 2, considerando os mesmos critérios, apresenta o efeito sobre o rendimento dos discentes em Matemática.

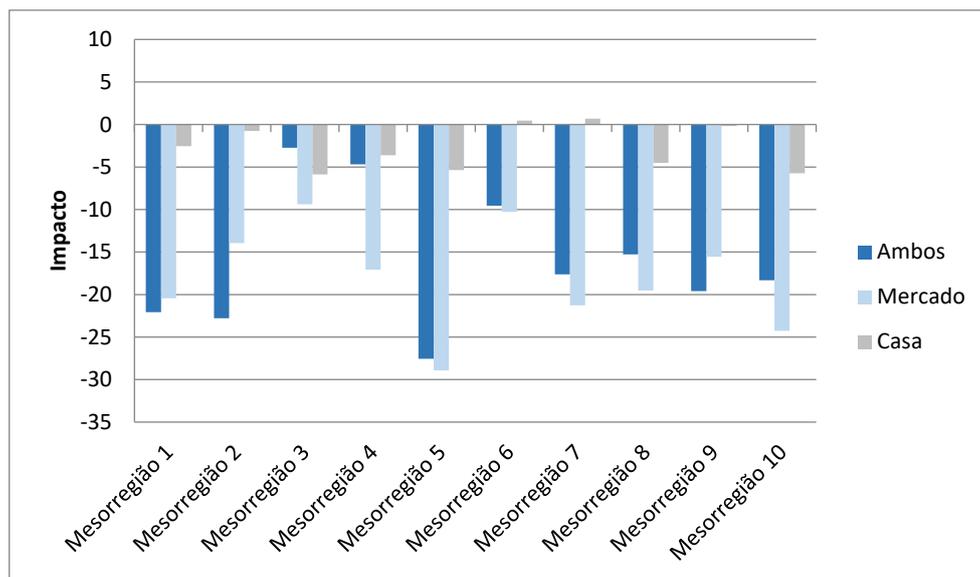
FIGURA 1 - Efeito médio do trabalho infantil na proficiência de Português, com PSM. Discentes da 5ª série, ano de 2015, por mesorregiões



Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

7 Para cada mesorregião (dez, no total) e em cada tipo de ocupação (no mercado, em casa e em ambos), foram feitos três tipos de pareamento, como dito na seção de metodologia. Em cada uma dessas estimativas, foi realizado um teste t, após o pareamento entre cada uma das variáveis utilizadas no modelo lógite. As estimativas do modelo foram obtidas em três regressões por mesorregião, totalizando 30. Em sua maioria, os resultados foram estatisticamente significativos, ou seja, não havia diferença estatística entre o grupo de controle e tratamento ao se avaliarem as variáveis explicativas, o que é o objetivo da metodologia (comparar pessoas, com características semelhantes, no grupo de controle e tratamento). Por consequência, as estimativas de impacto apresentadas nas figuras 1, 2, 3 e 4 também são estatisticamente significativas. Contudo, devido à expressiva quantidade de estimativas, a apresentação de todos os resultados dos testes e regressões se mostra inviável.

FIGURA 2 - Efeito médio do trabalho infantil na proficiência de Matemática. Discentes da 5ª série, ano de 2015, por mesorregiões



Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Os resultados mostram que, em todas as mesorregiões, o efeito do trabalho sobre alunos que trabalham em casa e no mercado ou só no mercado é negativo. No entanto, os efeitos se mostram maiores em algumas mesorregiões, principalmente na 5 (Centro-Oriental), na 1 (Noroeste) e na 2 (Centro-Occidental). A composição do trabalho infantil nas microrregiões que constituem a mesorregião 5, por exemplo, é predominantemente agrossilvopastoril e comercial; no caso da mesorregião 1, há predominância das atividades agrossilvopastoris e industriais (SEDES, 2017). Já no caso da mesorregião 3, o efeito negativo foi bem reduzido em comparação ao resto. A composição do trabalho infantil nessa mesorregião está marcada por menor predominância do setor agrossilvopastoril e maior presença da indústria e comércio. Em todos os casos, os efeitos negativos do trabalho doméstico foram relativamente pequenos.

A existência de heterogeneidade na composição do trabalho infantil pode implicar que os efeitos sejam diferenciados entre microrregiões do Paraná. Por exemplo, tomando o caso da mesorregião 3, se considerarmos as duas microrregiões mais intensivas no setor agrossilvopastoril, as duas mais intensivas em indústria e as duas mais intensivas em serviços e comércio, separadamente,⁸ observam-se algumas diferenças. No primeiro caso, os efeitos médios do trabalho infantil sobre o rendimento para alunos da 5ª série que trabalham em casa e no mercado são positivos (5,02) no caso do *score* de Português e negativos (-3,68) para Matemática; no

8 Respectivamente: Microrregião Faxinal e Ivaiporã; Microrregião Apucarana e Astorga; Microrregião Londrina e Maringá.

segundo caso, os efeitos se mostram negativos para Português (-13,33) e Matemática (-17,81); no terceiro caso, os efeitos também são negativos para Português (-26,91) e Matemática (-23,63).

Os resultados apontam que, especificamente na mesorregião 3, quando considerados os alunos da 5ª série, existe heterogeneidade do trabalho associada ao tipo de atividade praticada. Os infantes ocupados em setores com vocação agrária parecem ter efeitos negativos menores do que os ocupados em setores como indústria, comércio e serviços.

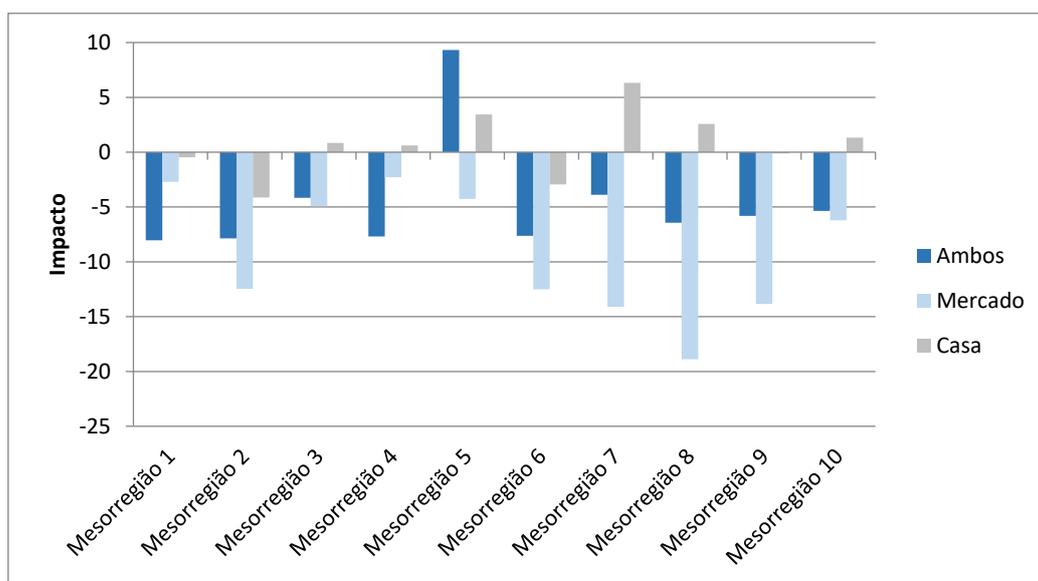
Tal resultado é contrastado com a situação observada na mesorregião 5. Quando consideradas suas duas microrregiões mais intensivas em trabalho infantil agrossilvopastoril, observa-se que o efeito médio sobre o rendimento escolar em alunos da 5ª série que trabalham no mercado e em casa é negativo no *score* de Português (-30,07) e de Matemática (-20,58), sendo bastante superior ao efeito observado nas microrregiões anteriores com trabalho infantil intensivo no mesmo setor. Já considerando sua microrregião menos intensiva no setor agrossilvopastoril e mais intensiva em comércio e serviços, e considerando também ocupados no mercado e em casa, os efeitos continuam a ser negativos em Português (-32,56) e Matemática (-33,28). O resultado indica que não necessariamente as regiões intensivas em agro terão menores efeitos negativos sobre o rendimento escolar dos infantes e sugere novamente a existência de heterogeneidade na composição do trabalho infantil no Paraná.

Nas Figuras 3 e 4, são considerados os efeitos médios do trabalho infantil sobre a proficiência em Português e Matemática, respectivamente, para alunos da 9ª série. A comparação é feita por mesorregiões do estado do Paraná levando em conta o *status* ocupacional do discente. *A priori*, é possível perceber que os efeitos do trabalho doméstico, além de serem menos prejudiciais do que para alunos da 5ª série, mostram-se positivos em algumas mesorregiões. Esse cenário tem sido observado em outros estudos (FRENCH, 2002; ROSSMAN, 2002), justificado pelo fato de que o trabalho doméstico pode gerar um impacto no sentido de responsabilidade do aluno e porque normalmente o tempo de trabalho pode ser controlado pelos pais ou responsáveis. No entanto, como assinalam Haile e Haile (2012), é necessário olhar outros fatores que podem incidir nas formas de trabalho doméstico infantil na zona rural, a exemplo de variáveis como o tamanho da família e a forma de propriedade da terra e do gado. Além disso, muitas famílias demandam estruturalmente a incorporação da mão de obra dos infantes (ADMASSIE, 2003).

Do mesmo modo, observa-se que os efeitos negativos são levemente menores em algumas mesorregiões, tanto no *score* de Português quanto no de Matemática. Inclusive, no caso da região 5, existe um impacto positivo do trabalho no rendimento quando as crianças trabalham em casa e no mercado. Tais evidências

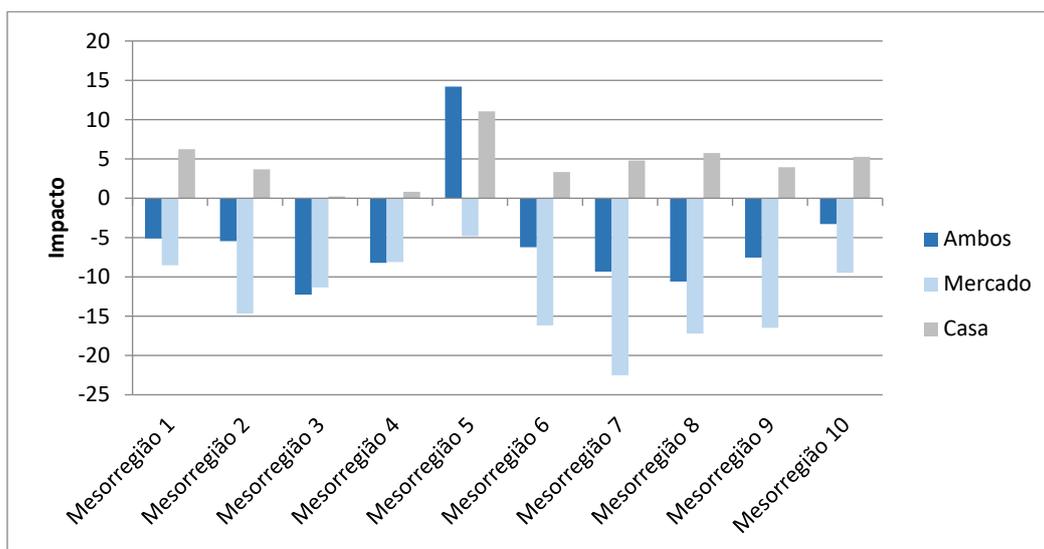
oferecem suporte à ideia de existência de heterogeneidade na magnitude do impacto do trabalho infantil, além de mostrar diferenças claras em relação aos resultados que foram estimados em outras pesquisas, já mencionadas anteriormente, em relação ao contexto nacional.

FIGURA 3 – Efeito médio do trabalho infantil na proficiência de Português. Discentes da 9ª série, ano de 2015, por mesorregiões



Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

FIGURA 4 – Efeito médio do trabalho infantil na proficiência de Matemática. Discentes da 9ª série, ano de 2015, por mesorregiões



Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015).

Levando em conta, novamente, a mesorregião 5, quando mensurados os efeitos nas suas microrregiões com predominância de trabalho infantil agrossilvo-pastoril, nota-se que o efeito médio sobre o rendimento dos alunos da 9ª série é totalmente diferente do que o observado em crianças da 5ª série. Em se tratando dos que trabalham no mercado e em casa, o efeito é positivo no caso de Português (10,36) e negativo (-4,37) no caso de Matemática. Isso não sugere que o trabalho infantil nos setores com predominância agrária seja *a priori* “positivo” para o desempenho escolar, mas que, no caso de discentes da 9ª série, o impacto pode ser menor ou menos prejudicial do que nos discentes da 5ª série. Isso segue a linha dos achados das pesquisas de Admassie (2003), Alvi e Dendir (2011), Haile e Haile (2012), Klocker (2011) e Gamlin *et al.* (2015). No entanto, é provável que a magnitude do impacto do trabalho infantil em setores agrários dependa mais da forma e esfera em que seja executado e de fatores que possam exercer influência (aumento da demanda por safras, composição familiar, fatores culturais etc.), o que representa maior dificuldade de observação utilizando unicamente uma análise quantitativa do fenômeno.

Certamente, na mesorregião 8, de predominância agro em todas as suas microrregiões, observa-se que o efeito médio quando os alunos da 9ª série trabalham só no mercado tende a ser negativo (-17,21 em Português e -18,88 em Matemática). Isso poderia ser explicado pelo fato de que o custo de trabalhar no mercado é mais elevado (principalmente pelo deslocamento até o local de trabalho, por exemplo) (BOURDILLON, 2009) e o tempo de trabalho dos alunos não é controlado pelos pais ou responsáveis. É importante assinalar que o resultado dessa mesorregião pode representar uma importante evidência a favor da heterogeneidade do impacto do trabalho infantil no rendimento escolar paranaense.

Por outro lado, no caso das microrregiões pertencentes à mesorregião 3, intensivas em trabalho infantil voltado para serviços e comércio, é possível observar que os efeitos médios negativos sobre o rendimento escolar em alunos da 9ª série são inferiores aos observados nos alunos da 5ª série, sendo negativos no *score* de Português e Matemática (-13,63 e -13,16) para quem trabalha no mercado e em casa. Os efeitos são também negativos em Português (-14,90) e Matemática (-9,51) para quem trabalha só no mercado. Isso confirma tendências relacionadas a menores prejuízos em infantes de maior idade, como apontado por Alvi e Dendir (2011).

Na mesma mesorregião 3, as microrregiões predominantes em indústria mostraram efeitos médios positivos em Português (0,11) e Matemática (2,22) para quem trabalha em ambos, e ligeiramente negativos (-5,02 em Português e -2,54 em Matemática) para alunos que trabalham só no mercado. Observa-se novamente que existe uma possível heterogeneidade do trabalho infantil no estado do Paraná, pois as diferenças nos efeitos são observados entre mesorregiões, microrregiões e

variam de acordo com a intensidade setorial da região e com a condição ocupacional das crianças.

De modo geral, os resultados mostram que (com algumas exceções, como o aumento do trabalho doméstico em alguns anos) não há mudanças temporais significativas no panorama de trabalho infantil do estado do Paraná no período analisado. Há tendências aparentes no estado que merecem atenção e outras abordagens, como o fato de que as meninas tendem a trabalhar mais em casa, e os meninos, no mercado. Apesar de existir uma tendência apresentada na literatura especializada, a direção do impacto do trabalho infantil não é totalmente clara. Ficou evidente que o trabalho doméstico nas zonas agrárias do Paraná parece ser menos prejudicial do que o trabalho em outros setores econômicos. Por último, foram obtidos resultados que sugerem que há significativa heterogeneidade na magnitude do impacto do trabalho infantil no estado.

CONCLUSÕES

O estudo procurou mostrar o impacto do trabalho infantil sobre o rendimento escolar no estado do Paraná e testar a possibilidade de heterogeneidade na magnitude do efeito. Para tal propósito, foram utilizados os dados do Inep (BRASIL, 2015) para a Prova Brasil de 2015, considerando os *scores* dos resultados das disciplinas de Português e Matemática para discentes da 5ª e 9ª séries. Nesse sentido, a partir do questionário da Prova Brasil, foram selecionadas variáveis específicas sobre o *background* socioeconômico do estudante e que podem incidir na probabilidade condicional de que ele trabalhe ou não.

A estatística descritiva, realizada com dados do Inep (BRASIL, 2007, 2009, 2011, 2013, 2015) para a Prova Brasil de 2007 a 2015, mostrou que não há mudanças temporais significativas nos níveis de trabalho infantil no estado do Paraná, embora haja uma leve tendência para sua gradual redução. Nesse sentido, os níveis de trabalho infantil nas três formas descritas no estudo são mais elevados para discentes da 9ª série, havendo uma tendência geral de que os meninos trabalhem no mercado, e as meninas, em casa. Outra evidência apontou que os discentes que trabalham em casa e no mercado tendem a ter um pior desempenho acadêmico em comparação àqueles que não trabalham ou que trabalham somente em casa.

A avaliação do impacto revelou que, em todas as mesorregiões, o efeito do trabalho infantil sobre alunos da 5ª série que trabalham em casa e no mercado ou só no mercado é negativo. No caso do trabalho doméstico, tal efeito negativo é inferior. Já nos alunos da 9ª série, os efeitos negativos do trabalho infantil são, no geral, inferiores aos observados nos alunos da 5ª série. Considerando o trabalho doméstico, foi possível observar que, além de ser menos prejudicial do que em

alunos da 5ª série, mostra-se positivo em algumas mesorregiões. Esse fato também foi constatado em outros estudos pela literatura consultada.

Já no caso dos alunos da 9ª série que trabalham só no mercado, o impacto negativo tende a ser maior, principalmente porque o custo de trabalhar no mercado é mais elevado (devido ao deslocamento até o local de trabalho, por exemplo) e porque o tempo de trabalho dos alunos não seria controlado pelos pais ou responsáveis. Por fim, foi possível detectar a heterogeneidade do impacto do trabalho infantil pela evidência de magnitudes diferenciadas, verificadas por meio do controle por mesorregiões, microrregiões e levando-se em conta as particularidades setoriais do tipo de trabalho e a forma de ocupação do estudante.

Recomenda-se fortemente que novas pesquisas de impacto do trabalho infantil na educação ponderem a existência de heterogeneidade não só no âmbito estadual, como é realizado na maioria dos trabalhos, mas que fatores intraestaduais e determinantes locais sejam levados em consideração (fatores associados a cultura, infraestrutura, saúde etc.). Tais particularidades podem mudar o perfil das políticas públicas regionais, fundamentais para o combate efetivo do trabalho infantil.

REFERÊNCIAS

- ADMASSIE, A. Child labour and schooling in the context of a subsistence rural economy: can they be compatible? *International Journal of Educational Development*, v. 23, n. 2, p. 167-185, 2003. [https://doi.org/10.1016/S0738-0593\(02\)00012-3](https://doi.org/10.1016/S0738-0593(02)00012-3).
- AKABAYASHI, H.; PSACHAROPOULOS, G. The trade-off between child labour and human capital formation: A Tanzanian case study. *Journal of Development Studies*, v. 35, n. 5, p. 120-140, 1999. <https://doi.org/10.1080/00220389908422594>.
- ALBERTO, M.; SANTOS, D.; LEITE, F.; LIMA, J.; PAIXÃO, G.; SILVA, S. Trabalho infantil doméstico: perfil bio-sócio-econômico e configuração da atividade no município de João Pessoa, PB. *Cadernos de Psicologia Social do Trabalho*, São Paulo, v. 12, n. 1, p. 57-73, 2009. <https://doi.org/10.11606/issn.1981-0490.v12i1p57-73>.
- ALVI, E.; DENDIR, S. Sibling differences in school attendance and child labour in Ethiopia. *Oxford Development Studies*, Oxford, v. 39, n. 3, p. 285-313, 2011. <https://doi.org/10.1080/13600818.2011.598923>.
- AMBADEKAR, N.; WAHAB, S.; ZODPEY, S.; KHANDAIT, D. Effect of child labour on growth of children. *Public Health*, v. 113, n. 6, p. 303-306, 1999. [https://doi.org/10.1016/S0033-3506\(99\)00185-7](https://doi.org/10.1016/S0033-3506(99)00185-7).
- ARANSIOLA, T. J.; JUSTUS, M. Child labor hazard on mental health: Evidence from Brazil. *Journal of Mental Health Policy and Economics*, v. 21, n. 2, p. 49-58, 2018.
- DE BARROS, R. P.; MENDONÇA, R. S. P. Trabalho infantil no Brasil: rumo à erradicação. *Sinais Sociais*, Rio de Janeiro, v. 5, n. 17, p. 142-173, set./dez. 2010.

- BASU, K. The intriguing relation between adult minimum wage and child labour. *Economic Journal*, v. 110, n. 462, p. 50-61, 2000. <https://doi.org/10.1111/1468-0297.00520>.
- BASU, K.; TZANNATOS, Z. The Global Child Labor Problem: What do we know and what can we do? *The World Bank Economic Review*, Oxford, v. 17, n. 2, p. 143-173, May 2003.
- BEEGLE, K.; DEHEJIA, R.; GATTI, R.; KRUTIKOVA, S. *The consequences of child labor: evidence from longitudinal data in rural Tanzania*. Washington: The World Bank, 2008.
- BEZERRA, M.; KASSOUF, A. L.; ARENDS-KUENNING, M. *The impact of child labor and school quality on academic achievement in Brazil*. Document 4062. Bonn: The Institute for the Study of Labor (IZA), 2009.
- BLAGBROUGH, J. Child domestic labour: A modern form of slavery. *Children & Society*, v. 22, n. 3, p. 179-190, 2008. <https://doi.org/10.1111/j.1099-0860.2008.00149.x>.
- BOURDILLON, M. Children as domestic employees: problems and promises. *Journal of Children and Poverty*, v. 15, n. 1, p. 1-18, 2009. <https://doi.org/10.1080/10796120802677586>.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil 2007*. Brasília: MEC/Inep, 2007.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil 2009*. Brasília: MEC/Inep, 2009.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil 2011*. Brasília: MEC/Inep, 2011.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil 2013*. Brasília: MEC/Inep, 2013.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Prova Brasil 2015*. Brasília: MEC/Inep, 2015.
- CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, v. 22, n. 1, p. 31-72, 2008. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>.
- CANALS-CERDÁ, J.; RIDAO-CANO, C. *The dynamics of school and work in rural Bangladesh*. Washington: The World Bank, 2004.
- DEHEJIA, R.; WAHBA, S. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, v. 94, n. 448, p. 1053-1062, 1999. <https://doi.org/10.1080/01621459.1999.10473858>.
- DELPRATO, M.; AKYEAMPONG, K. The effect of working on students' learning in Latin America: evidence from the learning survey TERCE. *International Journal of Educational Development*, v. 70, p. 102086, 2019. <https://doi.org/10.1016/j.ijedudev.2019.102086>.

- DUMAS, C. Does work impede child learning? The case of Senegal. *Economic Development and Cultural Change*, Chicago, v. 60, n. 4, p. 773-793, 2012. <https://doi.org/10.1086/665603>.
- EMERSON, P.; PONCZEK, V.; SOUZA, A. Child labor and learning. *Economic Development and Cultural Change*, Chicago, v. 65, n. 2, p. 265-296, 2017. <https://doi.org/10.1086/688895>.
- FORS, H. Child labour: a review of recent theory and evidence with policy implications. *Journal of Economic Surveys*, v. 26, n. 4, p. 570-593, 2012. <https://doi.org/10.1111/j.1467-6419.2010.00663.x>.
- FRASCO-ZUKER, L. Investigación etnográfica sobre experiencias de trabajo infantil en el noreste argentino. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, v. 14, n. 2, p. 1205-1216, 2016.
- FRENCH, J. L. Adolescent workers in third world export industries: attitudes of young Brazilian shoemakers. *ILR Review*, v. 55, n. 2, p. 308-323, 2002. <https://doi.org/10.1177/001979390205500206>.
- GALLI, R. *The economic impact of child labour*. Geneva: ILO Decent Work Research Programme, 2001. 26 p.
- GAMLIN, J.; CAMACHO, A.; ONG, M.; HESKETH, T. Is domestic work a worst form of child labour? The findings of a six-country study of the psychosocial effects of child domestic work. *Children's Geographies*, v. 13, n. 2, p. 212-225, 2015. <https://doi.org/10.1080/14733285.2013.829660>.
- GLEWWE, P.; MURALIDHARAN, K. Improving education outcomes in developing countries. *Handbook of the Economics of Education*, v. 5, p. 653-743, 2016. <https://doi.org/10.1016/B978-0-444-63459-7.00010-5>.
- GREENE, W. *Econometric analysis*. 5. ed. New Jersey: Prentice Hall, 2003.
- GUNNARSSON, V.; ORAZEM, P.; SÁNCHEZ, M. Child labor and school achievement in Latin America. *The World Bank Economic Review*, Oxford, v. 20, n. 1, p. 31-54, 2006. <https://doi.org/10.1093/wber/lhj003>.
- HAILE, G.; HAILE, B. Child labour and child schooling in rural Ethiopia: nature and trade-off. *Education Economics*, v. 20, n. 4, p. 365-385, 2012. <https://doi.org/10.1080/09645292.2011.623376>.
- HANUSHEK, E. Economic growth in developing countries: The role of human capital. *Economics of Education Review*, v. 37, p. 204-212, Dec. 2013. <https://doi.org/10.1016/j.econedurev.2013.04.005>.
- HEADY, C. The effect of child labor on learning achievement. *World Development*, v. 31, n. 2, p. 385-398, 2003. [https://doi.org/10.1016/S0305-750X\(02\)00186-9](https://doi.org/10.1016/S0305-750X(02)00186-9).
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Divisão regional do Brasil em regiões geográficas imediatas e regiões geográficas intermediárias*: 2017. Rio de Janeiro: IBGE, 2017. 82 p.
- INSTITUTO PARANAENSE DE DESENVOLVIMENTO ECONÔMICO E SOCIAL – IPARDES. *Mapa do trabalho infanto-juvenil no Paraná*. Curitiba: IparDES, 2007.
- KASSOUF, A. L. O que conhecemos sobre o trabalho infantil? *Nova economia*, Belo Horizonte, v. 17, n. 2, p. 323-350, maio/ago. 2007.
- KASSOUF, A. L. Evolução do trabalho infantil no Brasil. *Sinais Sociais*, Rio de Janeiro, v. 9, n. 27, p. 9-45, 2015.

KASSOUF, A. L.; SANTOS, M. J. Trabalho infantil no meio rural brasileiro: evidências sobre o “paradoxo da riqueza”. *Economia Aplicada*, v. 14, n. 3, p. 339-353, 2010.

KASSOUF, A. L.; TIBERTI, L.; GARCIAS, M. Evidence of the impact of children’s household chores and market labour on learning from School Census data in Brazil. *Journal of Development Studies*, v. 56, n. 11, p. 2097-2112, 2020. <https://doi.org/10.1080/00220388.2020.1736284>.

KASSOUF, A. L.; TIBERTI, L.; GARCIAS, M.; ONO, I. Evidence of the impact of children’s domestic and market labor on learning from School Census data in Brazil. *SSRN Electronic Journal*, p. 1-34, 2016. <https://doi.org/10.2139/ssrn.3164357>.

KLOCKER, N. Negotiating change: working with children and their employers to transform child domestic work in Iringa, Tanzania. *Children’s Geographies*, v. 9, n. 2, p. 205-220, 2011. <https://doi.org/10.1080/14733285.2011.562381>.

LOPES, J.; SOUZA, E.; PONTILI, R. Trabalho infantil e sua influência sobre a renda e a escolaridade da população trabalhadora do Paraná. In: SEMINÁRIO DO TRABALHO: TRABALHO, ECONOMIA E EDUCAÇÃO NO SÉCULO XXI, 6., 2008, São Paulo. *Anais [...]*. Marília: Editora Gráfica Massoni, 2008.

MAFFIOLI, A. *Non-experimental methods: Propensity score matching and difference in difference*. Cuernavaca: Inter-American Development Bank, 2011.

MERCAN, M.; SEZER, S. The effect of education expenditure on economic growth: The case of Turkey. *Procedia – Social and Behavioral Sciences*, v. 109, p. 925-930, 2014. <https://doi.org/10.1016/j.sbspro.2013.12.565>.

MOTA, M. *A luta pela erradicação do trabalho infantil no Brasil, enfoque no Paraná: 1992-2001*. 2003. 44 f. Monografia (Especialização em Economia do Trabalho) – Universidade Federal do Paraná, Curitiba, 2003.

NAVARRO, M.; DE SOUZA, S.; EVARINI, A. Crianças trabalhadoras e as condições de sua inserção no mercado de trabalho no Paraná. *Economia & Região*, Londrina, PR, v. 1, n. 1, p. 48-68, 2013.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO – OIT. *Convenção n. 138, de 1973 – Convenção sobre a idade mínima para admissão*. Genebra: OIT, 1973. Disponível em: https://www.ilo.org/brasilia/convencoes/WCMS_235872/lang-pt/index.htm. Acesso em: dez. 2020.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO – OIT. *Convenção n. 182 – Convenção sobre proibição das piores formas de trabalho infantil e ação imediata para sua eliminação*. Genebra: OIT, 1999. Disponível em: https://www.ilo.org/brasilia/convencoes/WCMS_236696/lang-pt/index.htm. Acesso em: dez. 2020.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO – OIT. *Medir o progresso na luta contra o trabalho infantil*. Estimativas e tendências mundiais 2000-2012. Bureau Internacional do Trabalho; Programa Internacional para a Eliminação do Trabalho Infantil (IPEC). Genebra: OIT, 2013.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO – OIT. *Invertir en todos los niños*. Estudio económico de los costos y beneficios de erradicar el trabajo infantil. Lima: OIT, 2015.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO – OIT. *Trabalho infantil: fatos e números no Brasil*. 2016. Disponível em: <https://www.ilo.org/brasilia/temas/trabalho-infantil/lang-pt/index.htm>. Acesso em: dez. 2020.

PARANÁ. Secretaria da Família e de Desenvolvimento Social –SEDES. *Caracterização do trabalho infantil no Paraná*. Curitiba: Governo do Estado do Paraná, 2017. 52 p.

PATRINOS, H.; PSACHAROPOULOS, G. Family size, schooling and child labor in Peru: an empirical analysis. *Journal of Population Economics*, v. 10, n. 4, p. 387-405, 1997. <https://doi.org/10.1007/s001480050050>.

PUTNICK, D.; BORNSTEIN, M. Is child labor a barrier to school enrollment in low- and middle-income countries? *International Journal of Educational Development*, v. 41, p. 112-120, 2015. <https://doi.org/10.1016/j.ijedudev.2015.02.001>.

RAVALLION, M.; WODON, Q. Does child labour displace schooling? Evidence on behavioural responses to an enrollment subsidy. *Economic Journal*, v. 110, n. 462, p. 158-175, 2000. <https://doi.org/10.1111/1468-0297.00527>.

ROSSMAN, M. Involving children in household tasks: is it worth the effort?. *ResearchWorks*, v. 21, p.1-2, 2002.

SALVADOR, P. Trabalho Infantil, um mal necessário? Uma análise com dados do SAEB. *Revista ESPACIOS*, Caracas, v. 37, n. 13, 2016.

SOUZA, E. *Do perfil ao comportamento do trabalho infantil na região Sul do Brasil pós-implementação de medidas institucionais de combate*. 2018. 257 f. Tese (Doutorado em Desenvolvimento Regional e Agronegócio) – Universidade Estadual do Oeste do Paraná, Toledo, 2018.

STINEBRICKNER, R.; STINEBRICKNER, T. R. Working during school and academic performance. *Journal of Labor Economics*, Chicago, v. 21, n. 2, p. 473-491, 2003. <https://doi.org/10.1086/345565>.

Recebido em: 12 AGOSTO 2020

Aprovado para publicação em: 09 DEZEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7624>

AVALIAÇÃO E POLÍTICAS DE RESULTADOS: CONDICIONANTES DA ATUAÇÃO DO COORDENADOR

 JEANNY MEIRY SOMBRA SILVA^I

 LAURINDA RAMALHO DE ALMEIDA^{II}

^I Pontifícia Universidade Católica de São Paulo (PUC-SP), São Paulo-SP, Brasil; jeanny.sombra@hotmail.com

^{II} Pontifícia Universidade Católica de São Paulo (PUC-SP), São Paulo-SP, Brasil; laurinda@pucsp.com.br

RESUMO

Este artigo traz recortes de uma pesquisa de doutorado cujo objetivo foi conhecer e analisar as práticas formativas empregadas por coordenadores pedagógicos das escolas públicas da rede estadual de ensino na cidade de São Paulo. Os sujeitos participantes foram 402 coordenadores dos ensinos fundamental e médio e foram utilizados como procedimentos metodológicos os métodos mistos (survey, observação e entrevista narrativa), que permitiram combinar olhares quantitativos e qualitativos à pesquisa. Dentre outros aspectos, o estudo revelou que as demandas políticas para alcance de metas e resultados das avaliações exercem influência na atuação do coordenador pedagógico na escola e condicionam as ações de formação docente.

PALAVRAS-CHAVE AVALIAÇÃO DO RENDIMENTO ESCOLAR • POLÍTICA EDUCACIONAL • COORDENADOR PEDAGÓGICO • FORMAÇÃO DOCENTE.

EVALUACIÓN Y POLÍTICAS DE RESULTADOS: CONDICIONANTES DE LA ACTUACIÓN DEL COORDINADOR

RESUMEN

Este artículo contiene recortes de una investigación de doctorado cuyo objetivo fue el de conocer y analizar las prácticas formativas empleadas por coordinadores pedagógicos de las escuelas públicas de la red estadual de enseñanza en la ciudad de São Paulo. Los participantes fueron 402 coordinadores de la educación básica y media y se utilizaron como procedimientos metodológicos los métodos mixtos (survey, observación y entrevista narrativa), que permitieron combinar miradas cuantitativas y cualitativas al estudio. Entre otros aspectos, el estudio reveló que las demandas políticas para alcanzar metas y resultados de las evaluaciones ejercen influencia en la actuación del coordinador pedagógico en la escuela y condicionan las acciones de formación docente.

PALABRAS CLAVE EVALUACIÓN DEL RENDIMIENTO ESCOLAR • POLÍTICA EDUCACIONAL • COORDINADOR PEDAGÓGICO • FORMACIÓN DOCENTE.

EVALUATION AND RESULTS POLICIES: CONDITIONING OF THE COORDINATOR'S PERFORMANCE

ABSTRACT

This article provides pieces of a doctoral research aimed to know and analyze the formative practices employed by pedagogical coordinators of public schools of the state school system in the city of São Paulo. The study used 402 coordinators from elementary and high schools as subjects. Mixed methods (survey, observation and narrative interview) were used as methodological procedures allowing the combination of quantitative and qualitative methods in the research. Among other aspects, the study revealed that the political demands to achieve goals and results in the evaluations influence the pedagogical coordinator performance in the school and condition the teacher training actions.

KEYWORDS EVALUATION OF SCHOOL PERFORMANCE • EDUCATIONAL POLICY • PEDAGOGICAL COORDINATOR • TEACHER TRAINING.

INTRODUÇÃO

Conforme Gatti (2002), nas últimas décadas, a avaliação educacional abrangeu um campo considerável de estudos e intervenção prática, contemplando subáreas com características distintas: avaliação de sistemas educacionais; de desempenho escolar em sala de aula; de rendimento escolar com objetivo de macroanálises; avaliação institucional; e autoavaliação. Um enfoque que tem ganhado cada vez mais dimensão é o uso da avaliação com objetivo de fornecer diagnóstico e subsídios para implementação de políticas educacionais. Gimenes *et al.* (2013, p. 16) explica que a partir dos anos 1990 houve uma introdução maciça desse tipo de avaliação no país, acompanhando “uma agenda e tendências internacionais já consolidadas que consideram nesse tipo de ação uma forma de se aferir a qualidade e a efetividade dos sistemas educacionais”. É o caso, por exemplo, do Sistema de Avaliação da Educação Básica (Saeb) que desde sua primeira edição, em 1990, teve dentre seus objetivos fazer uso dos dados dessa avaliação para a proposição de programas de políticas públicas com enfoque nos resultados (PESTANA, 2013; BONAMINO; SOUSA, 2012). A esse respeito, Vianna (2003, p. 43) afirma que o Saeb foi implementado “como tentativa de encontrar um caminho para a solução de alguns problemas educacionais mais prementes, esperando, possivelmente, que os processos avaliati- vos determinariam, entre outros aspectos, a elevação dos padrões de desempenho”.

Estudos também indicam que nos últimos anos houve um acentuado crescimento em sistemas de avaliações próprios de secretarias estaduais e/ou municipais em todo o país (BAUER *et al.*, 2015; BROOKE; CUNHA, 2011; GIMENES *et al.*, 2013). Tais autores apontam que o crescimento está associado não apenas ao desejo de melhoria de qualidade do ensino e à necessidade de acompanhamento das aprendizagens, mas também a aspectos relacionados aos usos das avaliações no âmbito de programas educacionais, mobilização da escola, monitoramento de redes de ensino, entre outros.

Ainda que no campo educacional o termo *qualidade* possua diversos significados e comporte diferentes dimensões, “o fato é que tais avaliações rapidamente se associaram à ideia de qualidade, limitando-a aos resultados obtidos pelos alunos” (GIMENES *et al.*, 2013, p. 16)

No âmbito da rede pública de ensino de São Paulo, o foco nos resultados dos alunos consubstanciou-se em programas de política pública da Secretaria de Educação do Estado de São Paulo (SEE-SP) (FINI, 2014). Nesse contexto, a avaliação em larga escala ganhou centralidade, como é o caso do Sistema de Avaliação e Rendimento Escolar do Estado de São Paulo (Saresp).

Entretanto, a leitura dos dados do Saresp por parte da equipe pedagógica da escola tem sido constantemente apontada como um complicador (SZEWCZUK, 2013; SOUZA, 2013). Cada escola precisa estudar seus resultados para compreender o diagnóstico aferido pela prova, o que envolve entender a matriz de referência do instrumento, as

escalas de proficiência de aprendizagem e os termos técnicos dos relatórios do Saresp. Mas a dificuldade em entendê-los muitas vezes afasta a escola do uso pedagógico desta avaliação. Na tentativa de entender como as escolas se apropriam dos dados do Saresp, Bauer (2006) fez uma pesquisa em cinco Diretorias Regionais de Ensino (DRE) e entrevistou 14 profissionais (dirigentes regionais, supervisores de ensino, coordenadores de Núcleo Pedagógico e coordenadores de escola), dentre outros de órgãos centrais da SEE-SP, e constatou que essa dificuldade se repete em diferentes instâncias.

Ler o relatório do Saresp deixa os professores enlouquecidos. Eles não conseguem entender, mudar a concepção deles de prova, de avaliação, é uma coisa muito demorada. A questão de pensar em habilidade, pensar em competência. Habilidade virou nome feio, ninguém entende nada! A gente está até evitando falar nisso para ver se eles vão assimilando algumas coisas, para depois tratar disso (depoimento de uma Dirigente Regional de Ensino). (BAUER, 2006, p. 149)

Num levantamento feito na literatura que aborda usos de resultados das avaliações externas por equipes gestoras. Se trata de uma coordenada restritiva, e não explicativa em que se encontra o coordenador pedagógico foi possível perceber que:

Não se tem ideia do impacto das avaliações externas nas ações dos gestores locais e dos professores. Por alguns estudos sabemos que escolas se apropriaram de fato dos dados, eles foram discutidos e geraram ações; outras muitas deixaram os resultados na gaveta, sendo os dados de conhecimento de poucos professores. (GATTI, 2007, p. 9)

Na mesma direção, Sousa e Oliveira (2010, p. 813) afirmam que “o uso dos resultados das avaliações externas por parte dos gestores é escasso ou inexistente”. Também Souza (2013), em seu estudo sobre a utilização dos resultados da avaliação externa da escola pela equipe gestora, constatou:

Hoje, e sempre, a grande preocupação quando se fala em avaliação de escola/sistema/redes de ensino é não somente fazer chegar até a escola os resultados em termos estatístico-quantitativos, mas como fazer para que toda a escola entenda esses resultados e possa fazer uso deles, transformando-os em informações importantes para melhorar a qualidade do seu trabalho. (SOUZA, 2013, p. 163)

Isso não significa que ações com base nas avaliações não aconteçam. Em sua pesquisa sobre o impacto pedagógico do Saresp na rede de ensino, Bauer (2006)

observou nas DREs investigadas que há reuniões por polos, oficinas dos gestores com professores, elaboração de projetos, entre outros. Mas, segundo a autora, essas ações ainda estão muito voltadas às divulgações gerais dos dados, com tradução pedagógica precária e de pouco sentido didático para as escolas, raramente dedicando-se tempo à reflexão e ao planejamento com base nesses resultados. Bauer (2006, p. 156) conclui: “enquanto a lógica gerencial e política ocupar o lado mais pesado da balança, pouco efetiva será a utilização dos resultados”.

Com vistas a aproximar os professores dos resultados produzidos pelos alunos de forma individualizada em escala menor, implementou-se, em 2012, na rede estadual de São Paulo, a Avaliação da Aprendizagem em Processo (AAP). Trata-se de uma prova de Língua Portuguesa e Matemática com perguntas de múltipla escolha e produção textual, composta a partir de habilidades operacionais da matriz de referência do Saesp e do currículo dos ensinos fundamental II e médio da SEE-SP. A prova é aplicada no início de cada semestre e, conforme documentos oficiais da SEE-SP, um dos objetivos desse instrumento é subsidiar o planejamento de ensino e redimensionar objetivos e metas para a realização do trabalho pedagógico, modificando o baixo desempenho discente. Para isso, equipe gestora e professores

[...] devem verificar a maior incidência de erros nas diferentes turmas de alunos relacionada aos temas/conteúdos/objetos de ensino testados em cada questão, possibilitando ao professor a ação necessária para que seu aluno tenha a possibilidade de avançar. (SÃO PAULO, 2012, p. 4)

Conforme a SEE-SP, para que tal intenção se concretizasse, as escolas deveriam reservar os horários institucionais, tais como as Aulas de Trabalho Pedagógico Coletivo (ATPC), para ponderar os resultados gerais de cada sala de aula e de cada aluno, objetivando transformá-los em propostas de intervenção (SÃO PAULO, 2014a).

Com o passar do tempo, tais propostas interventivas deixam de ocorrer apenas no bojo da escola e se centralizam na SEE-SP por meio de ações prescritivas advindas de programas e projetos de política de avaliação, tais como o Programa Gestão em Foco. Implantado em 2017, tal programa utiliza o Método de Melhoria de Resultados (MMR) para que, segundo informações da SEE-SP, as escolas conquistem avanços educacionais, pedagógicos e de gestão. No portal institucional do Programa, encontra-se a seguinte informação:

Em conjunto com a associação Parceiros da Educação, que conta com empresas, empresários e organizações da sociedade civil para auxiliar a Secretaria em projetos educacionais [...] o objetivo do Gestão em Foco é melhorar o aprendizado de 1 milhão de estudantes do Ensino Fundamental e Médio com

a formulação de planos de trabalho personalizados e monitorados pela própria comunidade escolar. O método é utilizado desde o planejamento estratégico para o ano letivo e passa por etapas como identificar os desafios, planejar formas de superá-los e implantar as soluções elaboradas. (SÃO PAULO, 2017)

Também nesse portal encontram-se orientações do tipo “passo a passo” sobre como as escolas podem proceder com os resultados dos desempenhos dos alunos, envolvendo áreas avaliadas, processos avaliativos formais, metodologia para construção de provas, relação dos resultados da AAP com o Saesp, entre outras informações. “Em todas as escolas que fazem parte do programa, existem painéis visíveis para pais, alunos e professores que apresentam as metas estipuladas e quais as formas para atingi-las” (SÃO PAULO, 2017).

A maneira como as informações e orientações são apresentadas no Programa MMR transmite uma visão pragmática e gerencialista na qual o objetivo principal é a melhoria do desempenho dos alunos. Tal perspectiva fortalece o que Ferreira (2007, p. 148) denomina de “cultura do resultado”, quando “as instituições são estimuladas a tornarem-se obsessivas em relação ao seu desempenho, à responsabilidade de verificar os resultados educacionais e divulgá-los, como forma de prestação de contas à sociedade”.

Resta saber como os profissionais da escola percebem essa dinâmica. Conforme Gimenes *et al.* (2013, p. 17), são poucos os estudos que discutem como a escola se apropria dos resultados e se estes, de fato, induzem a “transformação e maior qualificação do trabalho escolar e da própria prática docente”. A esse respeito, o presente artigo traz dados de uma pesquisa de doutorado desenvolvida no Programa de Psicologia da Educação da PUC-SP com coordenadores pedagógicos de escolas públicas da SEE-SP. Para este artigo, fizemos um recorte do estudo original com vistas a contemplar os seguintes objetivos: discutir, com base nos dados, como avaliação e políticas de resultados condicionam a atuação do coordenador pedagógico na escola; e identificar nos depoimentos dos coordenadores suas percepções sobre essa dinâmica.

A PESQUISA: ASPECTOS METODOLÓGICOS

O estudo utilizou o que Moscoso (2017, p. 633) caracteriza como “pesquisa com métodos mistos”, na qual dados quantitativos e qualitativos são usados de forma complementar, visando a atribuir sentidos e responder às questões da pesquisa de forma consistente e confiável. Assim, a pesquisa associou duas dimensões: uma que cobriu um maior volume de dados por meio de *survey* realizado com 402 coordenadores; e outra mais delimitada que permitiu um olhar aprofundado para o

objeto de estudo. Esta última consistiu na observação *in loco* em três escolas da capital paulista e em entrevistas com seus coordenadores. O objetivo foi coletar informações sobre as estratégias de formação docente utilizadas por esses profissionais nas reuniões pedagógicas semanais que ocorrem nas escolas.

Neste artigo, nos deteremos a um recorte do *survey* e utilizaremos, para complementar a discussão, trechos de depoimentos dos coordenadores para entrevista.

A ação formadora, ou seja, a condução da formação realizada pelo coordenador durante a reunião coletiva foi o objeto da pesquisa. Na rede pública de ensino de São Paulo, o momento de reunião coletiva ocorre por meio da chamada Aula de Trabalho Pedagógico Coletivo (ATPC). O objetivo primordial dessas reuniões é criar um espaço de discussão e formação para fortalecer o projeto político-pedagógico da escola, o que significa articular as disciplinas, estudar, atender a problemas enfrentados pelos docentes, trocar experiências, discutir planejamento e avaliação e estimular a reflexão sobre a prática docente. Para que isso aconteça, é importante haver uma boa estrutura para discussão coletiva, com objetivos e cronogramas.

O sujeito de nosso estudo, o coordenador pedagógico (CP), desempenha nesse contexto o papel fundamental de planejar, organizar e conduzir o encontro, possibilitando boas situações de aprendizagem entre todos. Temos clareza de que a dinâmica do encontro não se deve unicamente a esse profissional, no entanto, cabe a ele elaborar a pauta, conduzir as discussões, selecionar as estratégias e controlar o tempo. Seu modo de agir como mediador pode ditar o ritmo da reunião, favorecendo ou não a formação. De fato, espera-se que o CP assuma um papel de líder no encontro.

A transformação das reuniões que acontecem na escola em espaços de reflexão e produção de saberes sobre a docência exige uma metodologia proposta e dirigida pelo coordenador pedagógico, cuja liderança é essencial para que tais reuniões não assumam a condição de Horário de Trabalho Perdido. (BRUNO; CHRISTOV, 2003, p. 61)

Assim, destacamos como foco de observação do objeto da pesquisa as ações do CP durante a reunião: ele faz formação docente? Como são suas intervenções? Utiliza estratégias formativas; se sim, quais? De que maneira esse sujeito faz as mediações entre as estratégias propostas e o grupo docente?

A rede pública da SEE-SP conta com aproximadamente dois mil coordenadores na capital¹ e nosso interesse era ouvir o maior número de profissionais. Por isso,

1 O cálculo teve por base os dados fornecidos pela Escola de Formação e Aperfeiçoamento dos Profissionais da Educação do Estado de São Paulo (Efape), órgão da SEE-SP que, em setembro de 2018, registrava em sua base de dados, aproximadamente, dois mil coordenadores designados na região metropolitana de São Paulo. A Resolução SE n. 18, de 2/5/2019, alterou o siglário da SEE-SP.

optamos por elaborar um *survey* e distribuir o questionário por meio das DRE. A contribuição maior veio por meio do Grupo de Cooperação Técnica e Pesquisa (GCTEC) da Efape, que acolheu a pesquisa e encaminhou um *e-mail* institucional a todos os coordenadores das DRE da capital em atuação no ano letivo de 2018, convidando-os para participar do estudo. Por fim, obtivemos uma participação de 402 sujeitos, dos quais foi necessário excluir 22 respondentes por não atenderem aos critérios deste estudo (ter no mínimo seis meses de atuação na coordenação dos ensinos fundamental e médio). Assim, considerou-se para o levantamento um total de 380 sujeitos, parcela que representa 19% da população de coordenadores da rede pública estadual na capital paulista.

A pergunta principal do questionário foi: “Ao conduzir as reuniões de ATPC, que estratégias de formação docente você costuma utilizar? Fique à vontade para escrever o quanto desejar”. Após extrair do aplicativo *SurveyMonkey* todas as respostas dos coordenadores para essa questão, iniciamos o processo de análise, utilizando como balizador do estudo a técnica de análise de conteúdo (BARDIN, 2016).

Em princípio, nos detivemos nas leituras das cem primeiras respostas na tentativa de criar uma “base” para análise. Nessa amostra foi possível perceber quais assuntos foram mais recorrentes e, assim, selecionamos as respostas mais frequentes em blocos temáticos, o que permitiu a construção de índices e indicadores. Prosseguimos, então, com a classificação das respostas que convergiam para criar categorias preliminares.

O volume de material gerou uma grande quantidade de categorias, o que poderia fragmentar muito os dados. Assim, optamos pelos princípios organizatórios de Franco (2003, p. 54) de “categorias mais amplas ou *molares*, para depois classificar os indicadores (categorias *moleculares*) em módulos interpretativos menos fragmentados”. Mediante tal procedimento, as categorias iniciais, fragmentadas, passaram a ser indicadoras das categorias mais amplas. Feito isso com a amostra, retornamos às demais respostas para, por fim, chegar às categorias definitivas, conforme se observa no Quadro 1.

QUADRO 1 – A condução da ATPC: o que se faz, como e por quê? Um olhar para as respostas por meio da análise de conteúdo

CATEGORIAS	SUBCATEGORIAS
1 Estratégias de formação docente	Prática docente como objeto de estudo; Análise de produções de alunos e propostas didáticas.
2 O que se lê e o que se vê em ATPC?	Fonte, conteúdo e finalidade das leituras; Maneiras de conduzir a leitura de textos e vídeos.
3 Por trás das ações dos coordenadores	Ações definidas com base nos interesses da Escola; Ações definidas com base nas formações da DE; Ações definidas com base nas demandas da SEE-SP.

Fonte: Silva (2019, p. 109), com base na incidência das respostas e no número total de 380 participantes.

As subcategorias (categorias mais amplas) também foram organizadas em quadros para, subsequentemente, desenvolver a análise qualitativa do material. Por último, interpretamos os documentos agrupados com apoio na sistematização dos dados dos quadros, creditando os sentidos que os indicadores inferem ao objeto do estudo.

Na categoria 1 foram agrupadas todas as estratégias apontadas pelos coordenadores, tais como: tematização da prática; estudo de caso; dupla conceitualização; seminários; análise de práticas docentes; dentre outras. Na categoria 2, selecionamos os comentários que apresentavam “o que se lê e o que se vê em ATPC”, isto é, o conteúdo da formação.

Um dado que chamou a atenção ao realizarmos a análise do material foi que muitos coordenadores não se limitaram apenas em informar o que eles consideravam por estratégia formativa, mas relataram também os motivos que os levavam a adotar determinados procedimentos na formação. Assim, foi possível perceber que a definição das estratégias, bem como do que se lê e do que se vê em ATPC, é decorrente de alguns aspectos, tais como plano pedagógico do CP, seu roteiro formativo, projeto pedagógico da escola, currículo, necessidades da escola e demandas da SEE-SP e da Diretoria Regional de Ensino. As declarações que apresentavam os motivos das suas ações formativas em ATPC foram agrupadas na categoria 3 “Por trás das ações dos coordenadores”. Para contemplar os objetivos traçados para este artigo, selecionamos a discussão dos dados referente a essa categoria. A Tabela 1 apresenta mais detalhes.

TABELA 1 – Agrupamento dos eixos temáticos da categoria “Por trás das ações dos coordenadores”

SUBCATEGORIA: AÇÕES DEFINIDAS COM BASE NOS INTERESSES DA ESCOLA		N	%
Eixos temáticos	Definidas pelos coordenadores com base nas necessidades dos professores/escola.	32	9,0
	Definidas pelos professores/elaboradas conjuntamente entre coordenadores e professores.	14	4,0
	Tendo por base o currículo.	28	7,0
Eixo temático	SUBCATEGORIA: AÇÕES DEFINIDAS COM BASE NAS FORMAÇÕES DA DE		
	Adaptação/replicação da pauta formativa elaborada pelos PCNP do núcleo pedagógico da DE.	64	17,0
Eixo temático	SUBCATEGORIA: AÇÕES DEFINIDAS COM BASE NAS DEMANDAS DA SEE-SP		
	Relacionadas às proposições da SEE para análise de índices e melhoria de resultados: Programa MMR; plataforma SED – Foco Aprendizagem; I; Saresp; Idesp.	57	15,0

Fonte: Silva (2019, p. 128).

Nota: As porcentagens foram calculadas com base na incidência de respostas dos 380 coordenadores.

Dos 380 respondentes, 195 indicaram que as ações de formação são definidas com base nos interesses dos professores, bem como do currículo; assim, tais respostas foram agrupadas na subcategoria “Ações definidas com base nos interesses da escola”. As respostas mostraram que 32 coordenadores (9%) consideram os interesses dos professores ao definirem suas estratégias. Como se percebe nestes comentários:

Busco saber dos professores suas maiores dúvidas e necessidades do seu dia a dia. A partir disso, construo propostas de formação utilizando as estratégias mais adequadas, podendo ser a dupla conceitualização, tematização, estudo de casos. Sempre priorizando a troca de experiências. (CP 104)

Procuro identificar, junto ao grupo de professores, quais são as suas necessidades a curto, médio e longo prazo. Dessa forma, consigo priorizar as demandas mais urgentes, mas não deixar de lado as demais. (CP 248)

As formações são organizadas tendo como referência as necessidades formativas dos professores. Identificadas as necessidades, um Plano de Ação é elaborado [...]. (CP 133)

O CP 248 dá um sentido estratégico à formação, pois não se preocupa apenas “com as urgências”, mas encara a formação como um processo permanente, cuja chave é a integração com o grupo.² Esses coordenadores parecem compreender que a efetividade da formação em serviço é muito maior, seja para melhorar as práticas docentes ou para a aprendizagem dos alunos, quando derivadas das necessidades docentes com base em questões concretas da escola (GARCÍA, 1992).

Nos próximos tópicos, nos deteremos na discussão das subcategorias “Ações definidas com base nas demandas da SEE-SP” e “Ações definidas com base nas demandas da Diretoria Regional de Ensino (DRE)”.

POR TRÁS DAS AÇÕES DOS COORDENADORES: COMO AVALIAÇÃO E POLÍTICA DE RESULTADOS CONDICIONAM A ATUAÇÃO DO CP NA ESCOLA?

Dos 380 coordenadores participantes da pesquisa, 17% definem as estratégias e ações da ATPC tendo por base as orientações técnicas realizadas pelos Professores Coordenadores de Núcleo Pedagógico (PCNP) das DRE.

Cada DRE é responsável por elaborar um projeto de formação para ministrar as orientações técnicas aos coordenadores das escolas. Integram-se a esse projeto

2 Mais detalhes sobre o que os coordenadores apontaram por “interesses da escola” podem ser encontrados no estudo de Silva (2019).

as demandas da própria DRE regional, bem como dos órgãos centrais (SÃO PAULO, 2014b). Em casos de implementação de programas de política educacional da SEE, como no caso o MMR, os PCNP recebem formação sobre o conteúdo desses programas para replicá-los na escola. Trata-se de uma lógica de formação em cascata. Gatti e Barreto (2009, p. 202) explicam as características desse modelo:

Não raro o modelo de capacitação segue as características de um modelo “em cascata”, no qual um primeiro grupo de profissionais é capacitado e transforma-se em capacitador de um novo grupo que por sua vez capacita um grupo seguinte. Mediante esse procedimento, que geralmente percorre os diferentes escalões da administração dos extensos sistemas de ensino, corpo técnico-pedagógico, supervisores regionais, professores especialistas, embora permita envolver um contingente profissional bastante expressivo em termos numéricos, tem-se mostrado pouco efetivo quando se trata de difundir os fundamentos de uma reforma em suas nuances, profundidade e implicações.

Numa perspectiva de formação para o desenvolvimento profissional, esse modelo é pouco efetivo, conforme complementam as autoras: “ensinar a ensinar requer estratégias mais complexas e demoradas que as disponibilidades de capacitação oferecem” (GATTI; BARRETO, 2009, p. 202). Para que haja apropriação da aprendizagem, e não apenas um repasse de informações, o corpo de formadores (tanto os de base quanto os intermediários) precisa participar da elaboração das propostas e estar envolvido em todo o processo.

Não foi interesse do presente estudo discutir a pertinência desse modelo, no entanto, ao se propor a estudar a ação do CP da rede pública estadual, não se pode olhar apenas para o micro – escola –, pois ela está inserida em um sistema macro de ensino. De fato, tal aspecto impacta nas condições de trabalho desse profissional. Na lógica da “cascata”, o papel do CP é propagar e implementar na escola conteúdos estabelecidos em programas oficiais que, por vezes, estão desarticulados das necessidades da escola. Para exemplificar essa discussão, retomamos um comunicado publicado nos manuais orientadores de formação dos coordenadores, na ocasião do lançamento do programa São Paulo Faz Escola, o qual instituiu o currículo oficial.

A Secretaria da Educação considera que a coordenação pedagógica constitui-se em um dos pilares estruturais da sua atual política de melhoria da qualidade de ensino e que os Professores Coordenadores devem atuar como gestores implementadores dessa política com os objetivos de: ampliar o domínio dos conhecimentos e saberes dos alunos, elevando o nível de desempenho escolar evidenciado pelos instrumentos

externos e internos de avaliação; intervir na prática de sala de aula, incentivando os docentes a diversificarem as oportunidades de aprendizagem, visando à superação das dificuldades detectadas junto aos alunos; promover o aperfeiçoamento e o desenvolvimento profissional dos professores designados, com vistas à eficácia e à melhoria de seu trabalho. (SÃO PAULO, 2008c, p. 6, grifo nosso)

Os dados revelaram que 7% dos 380 coordenadores parecem assumir esse papel de “gestores implementadores dessa política” em sua atuação na escola, como se observa neste relato:

No ano de 1996, quando foi implantado o professor coordenador, não tinha muita orientação. O trabalho era desenvolvido a partir das necessidades da Escola. A gente ficava muito em cima de discutir os problemas dos alunos. Agora as reuniões, as estratégias, tudo que acontece na escola, vem por conta das orientações da SEE para o Currículo. (CP 222)

Para esse CP, a necessidade da escola representa “discutir problemas dos alunos”. Percebe-se no comentário que as “orientações” se desencadeiam a partir das orientações da SEE-SP para implementação do currículo. Nesse prisma, em que não há clareza do que se entende por formação continuada, as orientações normativas passam a configurar a “cartilha” norteadora das ações.

É certo que não se pode generalizar e afirmar que todos os coordenadores que relataram basear suas ações nas Orientações Técnicas do Núcleo Pedagógico mantêm concepções de formação similares. Muito menos dizer que são reprodutores autômatos dos conteúdos e estratégias advindas dessas orientações. Ao longo de seu estudo, Silva (2019) discutiu que o CP tem capacidade de se apropriar do conhecimento oferecido e atribuir-lhe novos sentidos, a partir da própria história pessoal e particular, adicionando a esse conhecimento saberes experienciais, num processo de ressignificação dialético.

Entretanto, pode-se afirmar que a lógica da formação de “cascata” da qual estão expostos os CP incide fortemente em sua atuação na escola e no seu papel como formador. A lógica da “cascata” é similar à do modelo escolar tradicional de formação docente (DEMAILLY, 1992), pois não há preocupação em abrir espaço para reflexão. A esse respeito, Nóvoa (2014, p. 144) constata que apesar das importantes mudanças pedagógicas ocorridas no último século, “o modelo escolar manteve-se incólume praticamente até os nossos dias, não sendo abusivo afirmar que grande parte dos educadores atuais considera ainda que o seu trabalho consiste em moldar a cera mole”. Apesar dos esforços em romper tal paradigma, ele

se perpetua “segundo uma lógica escolarizante, na qual se trabalha fundamentalmente em torno de uma formação institucionalizada” (NÓVOA, 2014, p. 151).

Outro dado observado é que 15% dos 380 coordenadores afirmam definir suas ações formativas na escola para atender às proposições da SEE no que concerne à análise de índices e melhoria de resultados, tais como o Programa MMR e as avaliações internas como AAP e externas como Saresp.

Algumas declarações dos respondentes enfatizam o valor da avaliação para tomar decisões em relação às estratégias de formação. O desempenho dos estudantes norteia as ações da escola; por isso a recorrência do tema na fala dos coordenadores.

As estratégias partem dos estudos dos resultados. Identificamos as habilidades e competências em defasagem (conforme aparecem na AAP) e estudamos os temas, depois os professores montam atividades coletivamente aplicam em sala de aula e no próximo encontro acontece a socialização com apresentação das evidências, ou seja, do que eles fizeram em sala de aula a partir do que planejamos. Os índices da escola estão melhorando com isso. (CP 226)

De acordo com os resultados obtidos nas avaliações do desenvolvimento das habilidades esperadas para aquele momento (série/ano) em que o aluno se encontra, são definidas as linhas de intervenção como: retomadas de conceitos, intensificar o que não ficou claro ainda e também aplicar o que faltou. É realizado um acompanhamento de acordo com a Matriz Curricular da SEE, Planejamento Anual e Replanejamento, Planos de Ação em Andamento e PPP³ da Unidade Escolar. Planejamento de ações contínuas de formação visando essas necessidades apontadas nos diagnósticos (avaliações) com aprofundamento adequado e inserção na rotina escolar até o assunto estar consolidado dentro do grupo. (CP 98)

Percebe-se nos depoimentos que esses coordenadores parecem fazer uso dos resultados para (re)planejar o que fazer com base neles: no comentário do CP 226 há um movimento circular de avaliar, planejar, aplicar e avaliar; já o CP 98 menciona que os resultados das avaliações servem como indicadores de múltiplos planos de ação, evidenciando como o tema é inserido “na rotina escolar até estar consolidado dentro do grupo”. Em ambos, infere-se que avaliação consome grande parte do tempo do cotidiano dos professores e alunos e mobiliza as ações formativas do CP. Outros depoimentos dos sujeitos também evidenciaram um controle sistemático

dos resultados por meio de planilhas, gráficos e registros, como se observa na declaração do CP 93: “Analiso os dados da AAP, mapa de classe com o gráfico dos resultados, para montar minha pauta formativa”.

O levantamento de habilidades defasadas por série/turma com base nos resultados da AAP é um procedimento que ganha destaque em 35% das respostas de todos os coordenadores. Como faz o CP 40: “Utilizo também para nortear os trabalhos os resultados das avaliações de aprendizagem em processo (AAP). Busco sempre informações que subsidiem o desenvolvimento das competências e habilidades”. Como destacado anteriormente, essa prova avalia apenas as disciplinas de Língua Portuguesa e Matemática.

É preciso estar atento para os efeitos que as políticas de avaliação podem causar na escola. Reduzir toda a ideia de formação em serviço aos resultados de Português e Matemática, sem levar em conta outros componentes curriculares, pode ser prejudicial para aprendizagem dos outros objetos de estudo; além disso, esse estreitamento gera desequilíbrio e hierarquiza as disciplinas. Como os professores dos outros componentes se sentem em relação a isso? Num estudo realizado com 40 professores de três escolas públicas da rede estadual, Silva, Rabello e Almeida (2017, p. 103) constataram que muitos docentes sentem suas disciplinas desvalorizadas pelo sistema, pela escola e pelos próprios colegas docentes, como revelam estes depoimentos:

A começar pelo Estado já percebemos a desvalorização, [...] disciplinas como Português e Matemática são aquelas que ‘valem mais’ não só para professores e coordenadores, como também para os alunos. (Anderson, professor de Sociologia)

Português e Matemática, infelizmente, são o centro da escola. (Lúcia, professora de História)

Conforme observado no estudo, o enfoque nos índices da escola, bem como nos resultados das avaliações, fez com que a SEE-SP implementasse nas escolas o Programa MMR. Para alcançar as metas desse programa, 7% dos 380 coordenadores declararam que promovem as etapas prescritas no cronograma do “planejamento estratégico do MMR” (SÃO PAULO, 2017). É o que se nota, por exemplo, na declaração do CP 134: “As formações são pautadas nas necessidades educacionais que as turmas apresentam na AAP e em consonância com o MMR. Conceitos de avaliação, competências e habilidades e atividades práticas para consolidar as reflexões sobre os temas tratados”.

Portanto, a investigação mais detalhada das respostas agrupadas na categoria 3 permitiu compreender que por trás das ações dos coordenadores, das estratégias utilizadas (ou não) por eles, incidem diversos fatores que devem ser considerados

quando se propõe a discutir a condução da ATPC. Pode-se concluir que as bases que sustentam os processos formativos pelos quais passam os CP nas orientações técnicas promovidas pelos núcleos pedagógicos, bem como as proposições da SEE-SP para análise de índices e melhoria de resultados, afetam decisivamente na concepção de formação desses profissionais e nos modelos de formação continuada que desenvolvem junto aos professores.

Isso corrobora com os achados de Placco, Almeida e Souza (2011, p. 18) em uma pesquisa realizada com coordenadores em âmbito nacional:

Quem planeja e organiza a formação de professores das escolas, em geral, são as Secretarias de Educação do Estado e do Município. A participação dos CPs nesse planejamento e execução da formação, em alguns Estados, se resume ao oferecimento de sugestões e, eventualmente, ao repasse ou proposição de discussões sobre os conteúdos que foram estudados, nas formações das Secretarias.

Nessa maneira de proceder a formação, o coordenador assume o papel de “multiplicador”, propagando conhecimentos que se constituem com base em conteúdos determinados *a priori* pela SEE-SP. A prática vista dessa maneira é entendida apenas como saber fazer, e para a qual os conteúdos teóricos não apresentam nenhuma contribuição.

A esse respeito, Imbernón (2011, p. 64) explica que a formação para assimilar um conhecimento profissional básico atua como uma espécie de “currículo oculto da metodologia”. Ou seja, os modelos com os quais o coordenador aprende assentam a base de suas concepções e estratégias de formação, eles “perpetuam-se com o exercício de sua profissão já que esses modelos se convertem, até de maneira involuntária, em pauta de sua atuação” (IMBERNÓN, 2011, p. 65). O autor considera ainda que, no desenvolvimento do conhecimento profissional, a metodologia deveria fomentar os processos reflexivos sobre a educação e a realidade social por meio de diferentes experiências. Um pensamento parecido encontra-se neste trecho:

Um profissional não pode se contentar com seguir “receitas” ou “aplicar” os conhecimentos teóricos anteriores à ação realizada, pois cada situação profissional que vive é singular e exige de sua parte uma reflexão em e sobre a ação, ação construída em parte pelo profissional que lhe deve dar sentido, precisamente o que Schön denomina *problem setting*. (TARDIF; MOSCOSO, 2018, p. 391)

Por outro lado, constatou-se que embora estejam inseridos num sistema no qual se privilegia a lógica da formação em cascata, para 9% dos 380 coordenadores as ações de formação têm como prioridade os interesses dos professores e da escola.

É certo que o trabalho do coordenador pedagógico não se restringe a atender às demandas da DRE ou dos órgãos centrais no que concerne à melhoria dos resultados das avaliações. Mas a pressão para atingir os índices consome grande parte do tempo desse profissional, tornando-se um dos condicionantes principais de sua atuação na escola. Conforme relatou o CP João em entrevista:

Não é algo que ache ruim, acho que é importante pra escola discutir índices e resultados de avaliações. Mas a gente tem tanta coisa importante pra discutir: os trabalhos interdisciplinares, o grêmio, um projeto envolvendo os pais, as metodologias ativas [...]. Mas tudo fica para segundo plano.

AS PERCEPÇÕES DOS COORDENADORES SOBRE AS DEMANDAS RESULTANTES DA POLÍTICA DE RESULTADOS

Ao se referirem ao Programa MMR, 32% dos 380 coordenadores que participaram do *survey* utilizaram expressões como *é algo imposto; vem de cima; não tem como escapar*. Tais expressões, permeadas de modalizadores discursivos, carregam um valor semântico negativo, criando um efeito de desaprovação e contrariedade.

Buscando entender melhor essas percepções, nos inserimos em quatro escolas e observamos diversas de suas reuniões formativas de ATPC, bem como realizamos entrevistas com os coordenadores dessas escolas. Essa etapa do estudo revelou que esses quatro profissionais não assumem uma posição de subserviência ao que “*vem de cima*”, como revela Eliane em entrevista:

Por conta do MMR a gente teve que fazer essa reflexão para preencher uma planilha com nosso plano de recuperação, então foi algo imposto, mas eu achei bem produtivo, porque a gente teve que parar e pensar: qual habilidade meu aluno está tendo mais dificuldade?

O que tornou a tarefa produtiva para Eliane? Foi o fato de ela conseguir atribuir um sentido pedagógico à demanda. Por isso, ela mobiliza o grupo para “parar e pensar” nas dificuldades de aprendizagem dos alunos.

Na entrevista de João, foi possível perceber que ele também cumpre com as demandas advindas de projetos de política da secretaria, mas acrescenta um olhar crítico à tarefa:

Bem, essa é uma prática que está atrelada ao MMR também, e diante dos números, aí pegando a AAP [Avaliação de Aprendizagem em Processo] a Matemática está com os índices baixos, então o nosso foco para 2018 é a metodologia

diversificada. Mas tudo fica com a nossa cara. [...] não é algo que ache ruim, acho que é importante pra escola discutir índices e resultados de avaliações. Mas a gente tem tanta coisa importante pra discutir.

João parece ter uma visão sistêmica dos índices da escola (provenientes da AAP e MMR) e com base neles propõe um “foco para 2018”. Ele explica que suas ações pedagógicas estão “atreladas ao MMR”, mas que adapta a demanda de um modo que “fique com a cara do grupo”. O CP considera importante discutir os resultados da escola; no entanto, afirma que a discussão não deve apenas girar em torno disso, pois menciona: “Mas a gente tem tanta coisa importante pra discutir”.

Esses e outros recortes suscitaram alguns questionamentos. Acreditamos que a falta de consciência ou a falsa consciência em torno da fragmentação do próprio trabalho pode interferir na ação do coordenador, aprisionando-o em uma alienação em torno de si próprio e do seu fazer (FREIRE, 1983). Mas será que os comentários desses coordenadores revelam uma visão alienada? Que peso têm as relações institucionais (DRE, supervisores, diretores, professores) na aceitação ou na resistência a essas normas de programas da SEE-SP? Que processos são gerados no dia a dia escolar para responder às demandas das políticas educacionais?

É fato que as exigências impostas à escola criam condições para o trabalho do CP e influenciam nas ações de formação. Mas é certo afirmar que o coordenador é apenas um gerencialista a serviço das demandas dos órgãos centrais? É possível encontrar algumas respostas estabelecendo um diálogo com Vaillant e Marcelo (2012, p. 169):

O desenvolvimento profissional docente implica interação com os contextos espacial e temporal. O contexto espacial refere-se ao ambiente social, organizativo e cultural no qual se realiza o trabalho dos docentes. Dentro desses ambientes, são produzidas múltiplas interações sociais, que envolvem companheiros, pais e diretores. Dessa forma, as condições de trabalho influenciam no desenvolvimento profissional docente, promovendo-o ou inibindo-o. Entretanto, essas condições não devem ser entendidas como uma relação de causa e efeito, mas como elementos mediadores.

Não se pode desconsiderar o contexto social e espacial ao investigar a condução da formação do CP na ATPC. Nos achados deste estudo foi possível perceber que as demandas políticas exercem influência nas condições de trabalho desse profissional, porém, como Vaillant e Marcelo (2012) consideram, essas condições não devem ser entendidas como uma relação de causa e efeito, mas sim como

elementos mediadores. A escola é orgânica, as múltiplas interações existentes nesse espaço são atravessadas pela subjetividade de cada pessoa, o que permite aos profissionais que nela atuam realizar adaptações, mediações e transformações nas demandas dos órgãos superiores. O comentário de Adriana na entrevista amplia essa leitura: “Mas é o que eu falo para os professores: nós temos que sim cumprir com toda parte burocrática do MMR, mas, já que nós temos que cumprir isso, vamos utilizar ao nosso favor, entende?”.

A coordenadora tem um posicionamento explícito em relação ao programa MMR, afirmando que é “burocrático”; ao mesmo tempo concorda: “temos que cumprir sim”. No entanto, a expressão “vamos utilizar ao nosso favor” traz subentendida uma apropriação do objeto (MMR), transformando-o para atender aos interesses do grupo que atua como seus interlocutores – os professores.

A esse respeito, uma investigação de seis anos em uma escola pública revela que,

Na mediação, os sujeitos em relação se apropriam dos significados que os outros atribuem aos fatos ou eventos e não dos fatos ou eventos em si, ou seja, configura também o modo como os outros da relação vivenciam a realidade. Ao fazer essa apropriação, no entanto, o fazem de maneira singular, de acordo com sua experiência, com sua história. Então é possível dizer que no processo de significação o sujeito se apropria dos significados que a cultura confere a fatos ou eventos. (SOUZA; PETRONI; DUGNANI, 2015, p. 59)

Os coordenadores compreendem as demandas dos órgãos centrais e entendem seus significados, mas esses profissionais têm a capacidade de fazer a mediação entre tais demandas e o contexto espacial escolar. Os saberes experienciais, a concepção de formação, as teorias implícitas, as vivências, as relações, enfim, diversos são os fatores para a permanente constituição da subjetividade do profissional coordenador, conferindo a ele capacidade de atribuir sentidos a essas demandas e ressignificá-las de maneira singular. Foi isso que fizeram os quatro coordenadores entrevistados para este estudo.

CONSIDERAÇÕES FINAIS

Por meio da análise dos dados deste estudo, constatamos que os modelos de formação aos quais os coordenadores estão mais expostos, como ocorre na lógica da formação em cascata à qual estão inseridos, incidem fortemente na concepção desses profissionais e, subsequentemente, nas propostas de formação continuada que desenvolvem junto aos professores. Por isso é importante que as instâncias da SEE-SP não se preocupem apenas com o repasse de informações de programas

educacionais. É necessária a implantação de políticas públicas relativas a uma formação específica para o coordenador, para que ele possa alicerçar suas concepções educacionais e fundamentar sua prática como formador.

A aceleração da fórmula “avaliação-problema-solução”, tal como proposta no Programa MMR, cria uma circunstância desfavorável que é o fechamento da escuta. Tudo passa a girar em torno de demandas e da solução de demandas; isso pode ser bom para o cumprimento de metas, mas e as pessoas? Nesse cenário, qual espaço para se construir conjuntamente a formação?

Discutir sobre avaliação não obstaculiza a aprendizagem; pelo contrário, pode promover um avanço nos conhecimentos docentes sobre o assunto. Entretanto, é importante que o coordenador planeje situações de formação em que os professores não sejam colocados apenas para ler ou ouvir dados do diagnóstico realizado, pois a simples constatação dos resultados não promove a reflexão. Olhar para os resultados em termos de erros e acertos evidencia uma concepção tradicional de ensino. É necessário debruçar-se sobre os dados e questioná-los: o que indicam sobre o processo de aprendizagem do aluno? Houve retrocesso em comparação à última avaliação? O que isso indica sobre o ensino? Que decisões serão tomadas em conjunto para promover avanços?

A pressão das diferentes instâncias da SEE-SP para melhora dos resultados da escola reverbera na atuação do CP, dado que esses profissionais passaram a ser responsabilizados pela elevação dos níveis de desempenho escolar expressos nos resultados das avaliações externas, como fica evidente no texto normativo das Resoluções SP n. 88/2007 (SÃO PAULO, 2008a) e n. 90/2007 (SÃO PAULO, 2008b). Atribuir valor apenas ao que é mensurável corrobora o discurso de educação utilitarista da racionalidade técnica, para o qual o mais importante é o conhecimento técnico e científico. A avaliação padronizada em larga escala, como no caso do Saesp, tem sua função, mas não se pode considerar que números absolutos são verdades homogeneizantes, pois se deixa de conhecer as singularidades intrínsecas do processo avaliativo. Como afirma Vianna (2005, p. 16),

A avaliação não é um valor em si e não deve ficar restrita a um simples rito da burocracia educacional; necessita integrar-se ao processo de transformação do ensino/aprendizagem e contribuir ativamente para o processo de transformação dos educandos.

Em vista disso, é importante que o coordenador faça sempre um autoquestionamento sobre sua função e sobre o papel que porventura assume nessa questão. A esse respeito, Fernandes (2011) adverte que, a depender de sua atuação, o coordenador pode assumir um papel gerencialista, perpetuando o monitoramento do Estado.

As orientações da SEE-SP sobre os programas de avaliação de resultados, como no caso do MMR, ainda que tenham um caráter vertical descendente, não devem ser simplesmente executadas pela escola. Podem e devem ser discutidas pelos gestores de cada unidade para que haja consenso sobre como tais programas serão apresentados aos docentes e incluídos na pauta de formação. Da mesma forma, as necessidades dos professores que dependam de providências da SEE-SP devem ser levadas pelo coordenador ao diretor para que faça a interlocução com a rede de ensino. Dessa maneira, o CP terá informações mais detalhadas para planejar suas ações.

Por fim, cabe um alerta:

Independentemente das motivações das políticas de avaliação em larga escala, diferentes objetivos e expectativas atreladas às avaliações externas podem se apresentar como dois lados de uma mesma moeda, ou seja, melhorar progressivamente os resultados implica – de alguma forma – tentar modificar e aperfeiçoar práticas de ensino e modos de organização escolares [...]. Embora a efetividade ou os efeitos de uma política educacional não possam ser avaliados apenas por seus objetivos declarados, é importante verificar o tipo de discurso empregado em favor do aperfeiçoamento do trabalho escolar, e que parece ir além do desejável aumento dos resultados. (GIMENES *et al.*, 2013, p.19)

O discurso empregado “em favor do aperfeiçoamento do trabalho escolar” do Programa Gestão em Foco, cujo pilar é o Método de Melhoria de Resultados – MMR, parte da premissa de que o grande problema dos índices e resultados das avaliações escolares é basicamente uma questão de gestão, desconsiderando os grandes problemas estruturais, sociais e econômicos da educação pública.

REFERÊNCIAS

BARDIN, Laurence. *Análise de conteúdo*. São Paulo: Edições 70, 2016.

BAUER, Adriana. *Usos dos resultados do Saresp: o papel da avaliação nas políticas de formação docente*. 2006. Dissertação (Mestrado em Educação) – Universidade de São Paulo, São Paulo, 2006.

BAUER, Adriana; PIMENTA, Claudia Oliveira; NETO, João Luiz H.; SOUSA, Sandra Zákia L. Avaliação em larga escala em municípios brasileiros: o que dizem os números? *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 62, p. 326-352, maio/ago. 2015.

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, p. 373-388, abr./jun. 2012.

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

BRUNO, Eliane B. G.; CHRISTOV, Luiza H. S. Reuniões na escola: oportunidades de comunicação e saber. In: BRUNO, Eliane B. G.; ALMEIDA, Laurinda, R.; CHRISTOV, Luiza H. S. *O coordenador pedagógico e a formação docente*. 4. ed. São Paulo: Loyola, 2003. p. 55-62.

DEMAILLY, Lise C. Modelos de formação contínua e estratégias de mudança. In: NÓVOA, Antônio. *Os professores e sua formação*. Lisboa: Dom Quixote, 1992. p. 139-158.

FERNANDES, Maria José da S. A coordenação pedagógica nas escolas estaduais paulistas: resoluções recentes e atuação cotidiana na gestão e organização escolar. *Revista Brasileira de Política e Administração da Educação*, Goiânia, v. 27, n. 3, p. 361-588, set./dez. 2011.

FERREIRA, Roseli Helena. *O Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo – Saesp: uma análise das provas de leitura e escrita da 4ª. série do ensino fundamental*. 2007. Dissertação (Mestrado em Educação) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2007.

FINI, Maria I. Currículo e avaliação: uma articulação necessária a favor da aprendizagem. In: NEGRI, Barjas; TORRES, Haroldo da G.; CASTRO, Maria Helena G. (org.). *A educação básica no estado de São Paulo: avanços e desafios*. São Paulo: Fundação Seade/FDE, 2014. p. 359-389.

FRANCO, Maria Laura P. B. *Análise de conteúdo*. Brasília: Líber Livro, 2003.

FREIRE, Paulo. *Educação e mudança*. 7. ed. Rio de Janeiro: Paz e Terra, 1983.

GARCÍA, Carlos M. A formação de professores: novas perspectivas baseadas na investigação sobre o pensamento do professor. In: NÓVOA, Antônio. *Os professores e sua formação*. Lisboa: Dom Quixote, 1992. p. 51-76.

GATTI, Bernardete A. Avaliação educacional no Brasil: pontuando uma história de ações. *Eccos Revista Científica*, São Paulo, v. 4, n. 1, p. 17-41, 2002.

GATTI, Bernardete A. Avaliação e qualidade da educação. In: SIMPÓSIO BRASILEIRO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 23., 2007, Porto Alegre. *Anais [...]*. Porto Alegre: Anpae, 2007. Disponível em: http://www.diaadiaeducacao.pr.gov.br/portals/seminariopde/documentos/processo5-para_saber_mais_bernadete_gatti.pdf. Acesso em: 28 set. 2020.

GATTI, Bernardete A.; BARRETO, Elba S. S. *Professores do Brasil: impasses e desafios*. Brasília: Unesco, 2009.

GIMENES, Nelson; SILVA, Vandrê Gomes da; PRÍNCIPE, Lisandra Marisa; LOUZANO, Paula; MORICONI, Gabriela Miranda. Além da prova Brasil: investimento em sistemas próprios de avaliação externa. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 55, p. 12-32, abr./ago. 2013.

IMBERNÓN, Francisco. *Formação docente e profissional: formar-se para a mudança e incerteza*. 9. ed. São Paulo: Cortez, 2011.

MOSCOSO, Javier N. Los métodos mixtos en la investigación en educación: hacia un uso reflexivo. *Cadernos de Pesquisa*, São Paulo, v. 47, n. 164, p. 632-649, abr./jun. 2017.

NÓVOA, Antônio. A formação tem que passar por aqui: as histórias de vida no Projeto Prosalus. In: NÓVOA, Antônio; FINGER, Mathias. *O método (auto)biográfico e a formação*. 2. ed. Natal: EDUFRRN, 2014. p. 143-175.

PESTANA, Maria Inês Gomes S. A experiência em avaliação de sistemas educacionais. Em que avançamos? In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origem e pressupostos*. Florianópolis: Insular Livros; Fundação Carlos Chagas, 2013. p. 117-133. (Ciclo de Debates v. 1).

PLACCO, Vera M. N. S.; ALMEIDA, Laurinda R. S.; SOUZA, Vera L. T. *O coordenador pedagógico e a formação de professores: intenções, tensões e contradições*. Relatório. São Paulo: Fundação Victor Civita, 2011. Disponível em: <http://www.fvc.org.br/pdf/apresentacao-coordenadores-qualitativo.pdf>. Acesso em: 18 ago. 2017.

SÃO PAULO (Estado). Secretaria da Educação. *Resolução SE - 88, de 19 de dezembro de 2007*. Dispõe sobre a função gratificada de Professor Coordenador. 2008a. Disponível em: http://siau.edunet.sp.gov.br/ItemLise/arquivos/88_07.HTM. Acesso em: 10 fev. 2020.

SÃO PAULO (Estado). Secretaria da Educação. *Resolução SE - 90, de 19 de dezembro de 2007*. Dispõe sobre função gratificada de Professor Coordenador nas quatro séries finais do ensino fundamental e no ensino médio, em escolas da rede estadual de ensino. 2008b. Disponível em: http://siau.edunet.sp.gov.br/ItemLise/arquivos/90_07.HTM. Acesso em: 10 jan. 2020.

SÃO PAULO (Estado). Secretaria da Educação. *Currículo do Estado de São Paulo: Caderno do Gestor. Gestão do Currículo na Escola*. São Paulo: SEE, 2008c. v. 3.

SÃO PAULO (Estado). Secretaria da Educação. *Avaliação da aprendizagem em processo: comentários e recomendações pedagógicas. Subsídios para o professor – Matemática. 1ª série do ensino médio – prova 2*. São Paulo: SEE, 2012. Disponível em: <https://midiasstoragesec.blob.core.windows.net/001/2019/05/mat-prova2-1-em-comentrios-e-recomendaes.pdf>. Acesso em: 11 fev. 2020.

SÃO PAULO (Estado). Secretaria da Educação. *Avaliação da aprendizagem em processo: comentários e recomendações pedagógicas. Subsídios para o professor de Língua Portuguesa. 7º ano / 2º semestre*. São Paulo: SEE, 2014a.

SÃO PAULO (Estado). Secretaria da Educação. *Aula de trabalho pedagógico coletivo (ATPC) em destaque*. São Paulo: Coordenadoria de Gestão da Educação Básica/SEE, 2014b. (Documento Orientador CGEB, n. 20).

SÃO PAULO (Estado) Secretaria da Educação. *Gestão em Foco – MMR*. 2017. Disponível em: <https://www.educacao.sp.gov.br/gestaoemfoco>. Acesso em: 10 abr. 2020.

SÃO PAULO (Estado) Secretaria da Educação. *Resolução SE n. 18, de 02 de maio de 2019*. Dispõe sobre o siglário a ser utilizado, no âmbito da Secretaria de Estado da Educação, pelas unidades que compõem a sua nova estrutura organizacional, e dá providências correlatas. 2019. Disponível em: <http://www.educacao.sp.gov.br/lise/sislegis/detresol.asp?strAto=201905020018>. Acesso em: 10 abr. 2020.

SILVA, Jeanny M. S. *Diferentes caminhos para formação docente: estratégias empregadas por coordenadores pedagógicos*. 2019. 279f. Tese (Doutorado em Educação: Psicologia da Educação) – Pontifícia Universidade Católica de São Paulo, São Paulo, 2019.

SILVA, Jeanny M. S.; RABELLO, Kátia M.; ALMEIDA, Laurinda R. As relações interpessoais nos contextos escolares: as várias faces do jogo coletivo. In: PLACCO, Vera M. N. S.; ALMEIDA, Laurinda R. *O coordenador pedagógico e a legitimidade de sua atuação*. São Paulo: Edições Loyola, 2017. p. 95-116.

SOUSA, Sandra Z.; OLIVEIRA, Romualdo P. Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. *Cadernos de Pesquisa*, São Paulo, v. 40, n. 141, p. 793-822, set./dez. 2010.

SOUZA, Maria Alba. O uso dos resultados da avaliação externa da escola: relação entre os resultados da avaliação externa e a avaliação interna dos alunos. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origem e pressupostos*. Florianópolis: Insular Livros; Fundação Carlos Chagas, 2013. p. 163-176. (Ciclo de Debates, v. 1).

SOUZA, Vera L. T.; PETRONI, Ana P.; DUGNANI, Lilian A. C. A dimensão do trabalho coletivo na escola: intervenções com a equipe gestora. In: PLACCO, Vera Maria N. S.; ALMEIDA, Laurinda Ramalho de. *O coordenador pedagógico no espaço escolar: articulador, formador e transformador*. São Paulo: Loyola, 2015. p. 53-72.

SZEWCZUK, Valéria Nunes F. *Formação de professores em ATPC: uma leitura a partir dos dados do Saesp*. 2013. Dissertação (Mestrado em Educação) – Universidade Metodista de São Paulo, São Bernardo do Campo, 2013.

TARDIF, Mauricio; MOSCOSO, Javier N. A noção de “profissional reflexivo” na educação: atualidade, usos e limites. *Cadernos de Pesquisa*, São Paulo, v. 48, n. 168, p. 388-411, abr./jun. 2018.

VAILLANT, Denise; MARCELO, Carlos. *Ensinando a ensinar*. As quatro etapas de uma aprendizagem. Curitiba: Editora UTFPR, 2012.

VIANNA, Heraldo M. Avaliações nacionais em larga escala: análises e propostas. *Estudos em Avaliação Educacional*, São Paulo, v. 54, n. 27, p. 41-76, jan./jun. 2003.

VIANNA, Heraldo M. *Fundamentos de um programa de avaliação educacional*. Brasília: Liber Livro, 2005.

Recebido em: 27 JUNHO 2020

Aprovado para publicação em: 23 OUTUBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7332>

INDICADORES DE MATRÍCULAS DO PÚBLICO-ALVO DA EDUCAÇÃO ESPECIAL NO ESTADO DE SÃO PAULO

-  RÉGIS HENRIQUE DOS REIS SILVA^I
 JOICE MELO VIEIRA^{II}
 JULIANA SALLES MADEIRA^{III}
 LUCIANA CRISTINA SALVATTI COUTINHO^{IV}

^I Universidade Estadual de Campinas (Unicamp), Campinas-SP, Brasil; regishsilva@gmail.com

^{II} Universidade Estadual de Campinas (Unicamp), Campinas-SP, Brasil; joicemv@unicamp.br

^{III} Universidade Estadual de Campinas (Unicamp), Campinas-SP, Brasil; juliana_madeira@yahoo.com.br

^{IV} Universidade Federal de São Carlos (UFSCar), campus Sorocaba, Sorocaba-SP, Brasil; lucscoutinho@gmail.com

RESUMO

Este artigo tem como objetivo analisar os indicadores de matrículas do Público-Alvo da Educação Especial no estado de São Paulo no período de 2007 a 2015. Para tanto, utilizamos como fonte de informações os microdados dos censos escolares do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira dos anos de 2007, 2011 e 2015, bem como as principais legislações e normas orientadoras de políticas nacionais e do estado de São Paulo no período analisado. Os dados demonstram que, se por um lado tem crescido o acesso do Público-Alvo da Educação Especial na rede estadual de ensino de São Paulo, por outro lado há, internamente, processos marcadamente excludentes, segundo os dados de fluxo escolar. Com efeito, desenvolvemos algumas hipóteses interpretativas para essa realidade educacional e concluímos que são muitos os desafios e barreiras que inviabilizam a efetivação da educação como direito social ao Público-Alvo da Educação Especial na rede estadual de ensino de São Paulo.

PALAVRAS-CHAVE EDUCAÇÃO ESPECIAL • INDICADORES EDUCACIONAIS • POLÍTICAS EDUCACIONAIS • SÃO PAULO.

INDICADORES DE MATRÍCULAS DEL PÚBLICO-META DE LA EDUCACIÓN ESPECIAL EN EL ESTADO DE SAN PABLO

RESUMEN

Este artículo analiza los indicadores de matrículas del Público-Meta de la Educación Especial en el Estado de San Pablo en el período de 2007 a 2015. Para tanto, utilizamos los microdatos de los censos escolares del Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [Instituto Nacional de Estudios e Investigaciones Educativas Anísio Teixeira] de los años de 2007, 2011 y 2015, así como legislaciones y normas orientadoras de Políticas Nacionales y del Estado de San Pablo. Los datos demuestran que creció el acceso del Público-Meta de la Educación Especial en la red estadual de enseñanza de San Pablo, sin embargo, hay, internamente, procesos notoriamente excluyentes según los datos de flujo escolar. En efecto, concluimos que son muchos los desafíos y las barreras que no hacen viable la efectivización de la educación como derecho social para ese Público-Meta en la red de enseñanza analizada.

PALABRAS CLAVES EDUCACIÓN ESPECIAL • INDICADORES EDUCACIONALES • POLÍTICAS EDUCACIONALES • SAN PABLO.

INDICATORS OF ENROLLMENT OF SPECIAL EDUCATION STUDENTS IN SAO PAULO STATE (BRAZIL)

ABSTRACT

This article analyzes indicators of enrollment of Special Education students in Sao Paulo state from 2007 to 2015. To this end, we used the microdata from 2007, 2011 and 2015 school censuses of Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [Anísio Teixeira National Institute of Educational Studies and Research], as well as laws and guidelines of national and Sao Paulo state policies. The data shows that the access of Special Education students (in Brazil called “Público-Alvo da Educação Especial”, which loosely translates as Target Public of Special Education) has grown in Sao Paulo state school system. However, according to student flow data, internally there are marked exclusion processes. Indeed, many challenges and barriers prevent the realization of the right to education of the target public of Sao Paulo state school system.

KEYWORDS SPECIAL EDUCATION • EDUCATION INDICATORS • EDUCATION POLICIES • SAO PAULO.

INTRODUÇÃO

Este artigo tem como objetivo analisar os indicadores de matrículas do Público-Alvo da Educação Especial (PAEE)¹ no estado de São Paulo no período de 2007 a 2015. Para tanto, utilizamos como fonte de informações os microdados dos censos escolares do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) dos anos de 2007, 2011 e 2015, bem como as principais legislações e normas que orientaram as políticas nacionais e do estado de São Paulo no período analisado.

O critério de escolha do marco temporal do artigo considerou a disponibilidade de dados individualizados. Assim, trabalhamos com as informações dos anos-calendário correspondentes ao primeiro ano de mandato das últimas três gestões encerradas do governo estadual paulista, a fim de analisar os indicadores de matrículas no contexto das políticas educacionais direcionadas ao PAEE ao longo dos sucessivos governos.

Embora as políticas direcionadas para a escolarização dos estudantes PAEE sejam de longa data e, especialmente nos últimos 30 anos, tenham ganhado maior projeção no contexto social, político, econômico e cultural do Brasil, avanços mais significativos ocorreram a partir dos anos 2000, quando políticas de Educação Especial (EE) foram de fato implantadas sob princípios inclusivistas,² ou, como enunciado no texto da política nacional de 2008, na perspectiva da educação inclusiva (BRASIL, 2008a).

Nos dois mandatos presidenciais de Luiz Inácio Lula da Silva (2003-2006 e 2007-2010), deu-se continuidade à difusão dos princípios de inclusão educacional presentes nos mandatos de Fernando Henrique Cardoso (1995-1998 e 1999-2002); mas, pelas características da política de conciliação de classes via “coesão social” (SILVA JÚNIOR, 2005) que caracterizam as gestões do governo Lula, os próprios nomes dos planos plurianuais (PPA) incorporaram os princípios inclusivistas, a saber: “Plano Brasil de Todos: participação e inclusão” (2004-2007) e “Plano Desenvolvimento com inclusão social e educação de qualidade” (2008-2011). É nesse contexto que a educação foi apresentada como prioridade nacional e como instância que objetivava

1 Na área educacional, utilizam-se diferentes denominações, como “pessoas em situação de deficiência”, “pessoas com deficiência”, “deficientes”, “pessoas portadoras de deficiências”, entre outros. Assim, procuraremos manter, sempre que possível, os termos originais utilizados pelos autores e leis/normas citados neste texto. No entanto, optamos pelo termo “Público-Alvo da Educação Especial”, conforme preconizado na política nacional de 2008, que compreende estudantes com deficiência, transtornos globais do desenvolvimento (TGD) e altas habilidades/superdotação (AH/SD).

2 Segundo Stainback e Stainback (1992 *apud* DORÉ; WAGNER; BRUNET, 1997, p. 176), “trata-se de um novo paradigma. [...] a noção de *full inclusion* prescreve a educação de todos os alunos nas classes e escolas de bairro. [...] reflete mais clara e precisamente o que é adequado: todas as crianças devem ser incluídas na vida social e educacional da escola e classe de seu bairro, e não somente colocadas no curso geral (*mainstream*) da escola e da vida comunitária, depois de ele já ter sido excluído”.

a promoção da equidade através da valorização da diversidade e da viabilização da inclusão social. Como observa Kassar (2011), tais planos fazem parte de um discurso coerente, difundido pela Organização das Nações Unidas (ONU). No caso da Educação Especial, as Convenções da Guatemala (2001) e de Nova Iorque (2006) intensificaram a divulgação de uma perspectiva inclusiva para a educação.³ Assim, ao longo do governo Lula, a Política Nacional de Educação Especial (PNEE) recebeu novas ações com o desenvolvimento de três programas implantados nacionalmente, dois voltados para a educação básica e um para a educação superior: 1) Programa de Implantação de Salas de Recursos Multifuncionais;⁴ 2) Programa Educação Inclusiva: direito à diversidade; 3) Programa Incluir.⁵ Esses programas colaboraram significativamente para a expansão dos princípios inclusivistas na política de Educação Especial no Brasil na última década (GARCIA; MICHELS, 2011).

Garcia e Michels (2011), ao analisarem os três programas supracitados, observam que o Programa de Implantação de Salas de Recursos Multifuncionais foi uma ação governamental de criação de equipamentos públicos de Educação Especial no âmbito da escola/universidade concretizada através da sistemática de editais, que também foi utilizada para os outros dois programas.

Os anos que se seguiram a 2008 têm sido decisivos para a definição de uma mudança de curso nas diretrizes da política nacional de Educação Especial (GARCIA; MICHELS, 2011), pois, desde a publicação do documento orientador da Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva (PNEE-EI), em 2008, a área vem passando por redefinições. Por exemplo, existe uma mudança na compreensão do PAEE, não obstante a manutenção do termo “alunos com necessidades especiais”. Tais estudantes são compreendidos, pela nova definição, como

- 3 Convenção de Guatemala – Convenção Interamericana para a Eliminação de Todas as Formas de Discriminação contra as Pessoas Portadoras de Deficiência, realizada na Guatemala em 2001 e aprovada no Brasil por meio do Decreto Legislativo n. 198, de 13 de junho de 2001; Convenção de Nova Iorque – Convenção sobre os Direitos das Pessoas com Deficiência e de seu Protocolo Facultativo, assinada em Nova Iorque, em 30 de março de 2007, aprovada pelo Senado Federal pelo Decreto Legislativo n. 186/2008 e pelo Decreto Executivo n. 6.949/2009 (SILVA, 2016).
- 4 A Educação Especial foi tratada como um serviço educacional especializado complementar, suplementar ou substitutivo à educação comum a partir da Resolução CNE/CEB n. 2/2001. O Programa de Implementação de Salas de Recursos Multifuncionais (2007) e o Decreto n. 6.571/2008 definem o *locus* privilegiado do Atendimento Educacional Especializado e reiteram a compreensão de Educação Especial como serviço complementar e suplementar à educação comum (GARCIA; MICHELS, 2011, p. 111).
- 5 Segue exemplos de outros programas que constituíram a política na época, mas que aqui podem ser considerados de menor escopo em termos da gestão da educação especial: Programa Escola Acessível, que consiste na liberação de recursos financeiros mediante o Programa Dinheiro Direto na Escola (PDDE) para a adequação de espaço físico em escolas estaduais e municipais visando à acessibilidade física; Programa BPC na Escola, que realiza o acompanhamento do acesso e da permanência na escola das pessoas com deficiência, favorecidas pelo Benefício da Prestação Continuada da Assistência Social (BPC), na faixa etária de 0 a 18 anos, por meio da articulação das políticas de educação, saúde, assistência social e direitos humanos (GARCIA; MICHELS, 2011, p. 111).

aqueles que apresentam deficiências, altas habilidades/superdotação e transtornos globais do desenvolvimento (BRASIL, 2008a), o que, segundo Garcia e Michels (2011), retoma os termos presentes no documento PNEE, de 1994, quando a política se apoiava no princípio da integração (BRASIL, 1994). O conceito de Educação Especial também foi redefinido, pois, na nova diretriz, abandonou-se a ideia de Educação Especial como uma proposta pedagógica para se centrar na disponibilização de recursos e serviços (GARCIA; MICHELS, 2011).

No referido documento, o Atendimento Educacional Especializado (AEE) é definido como um “conjunto de atividades, recursos de acessibilidade e pedagógicos organizados institucionalmente, prestado de forma complementar ou suplementar à formação dos alunos no ensino regular” (BRASIL, 2008b).

A Resolução n. 4/2009 do Conselho Nacional de Educação/Câmara de Educação Básica (CNE/CEB), que institui as diretrizes operacionais para o AEE na educação básica, prioritariamente na Sala de Recursos Multifuncionais, reafirmou a Educação Especial como modalidade educacional e enfatizou o AEE (BRASIL, 2009). Na mesma perspectiva, o Decreto n. 7.611 “dispõe sobre a Educação Especial, o Atendimento Educacional Especializado e dá outras providências” (BRASIL, 2011, p. 1).

Em conformidade com os dois planos plurianuais anteriores, o PPA para o período de 2012-2015, do governo da presidenta Dilma Rousseff, também estabeleceu ações prioritárias sob a perspectiva inclusiva (BRASIL, 2011).

Apesar de o contexto político ser outro, no primeiro mandato de Dilma (2011-2014) foi aprovado o Plano Nacional de Educação (PNE) para o período de 2014-2024, e sua meta 4 prevê a universalização:

[...] para a população de 4 (quatro) a 17 (dezessete) anos com deficiência, transtornos globais do desenvolvimento e altas habilidades ou superdotação, o acesso à educação básica e ao atendimento educacional especializado, *preferencialmente* na rede regular de ensino, com a garantia de sistema educacional inclusivo, de salas de recursos multifuncionais, classes, escolas ou serviços especializados, públicos ou conveniados. (BRASIL, 2014, p. 24, grifo nosso)

Como se pode observar, o PNE (2014-2024) retoma o termo “preferencialmente” e, com isso, toda uma discussão em torno das classes e escolas especiais, apresentadas pelo Ministério da Educação (MEC) em proposta recente de revisão da PNEE-EI (KASSAR; REBELO; OLIVEIRA, 2019; REBELO; KASSAR, 2018; SILVA; MACHADO; SILVA, 2019).

Em síntese, o ideário inclusivo e, particularmente, o discurso da educação inclusiva passaram a ser de interesse dos governantes – entre 1990 e 2003 – não só pelas vantagens educacionais e formativas que apresentavam aos estudantes

PAEE, mas também pelo fato de que representavam a possibilidade real de diminuição do uso de recursos públicos na área. O ideário inclusivo começou a fazer parte da agenda política de gestores nas diferentes esferas do Executivo – municipal, estadual e federal –, de modo que, a partir dos anos 1990, muitos municípios e estados revisaram suas políticas e iniciaram um processo de transferência dos estudantes PAEE para as escolas comuns.

A partir de 2003, a educação inclusiva passou a ser assumida como política do governo pela esfera federal e, por meio da estrutura de repasse de recursos para os governos municipais e estaduais, iniciou-se um processo de indução da política de “inclusão escolar” no país (GARCIA, 2013).

Nesse contexto, é oportuno observar que o estado de São Paulo, como parte integrante do sistema federativo do país, também passou a regulamentar a Educação Especial dentro de sua esfera de atuação e a prover as condições para o atendimento prescrito no âmbito das leis e diretrizes federais. A Secretaria de Educação do Estado de São Paulo (SEE-SP), atenta às políticas nacionais e, principalmente, à obtenção e administração dos recursos financeiros repassados pelo governo federal, elaborou várias resoluções nos últimos trinta anos, dentre as quais merecem destaque aquelas formuladas após 2003. O Quadro 1, a seguir, elenca essas resoluções.

QUADRO 1 – Resoluções da SEE-SP alinhadas com o PNEE e o PNEE-EI

RESOLUÇÃO	DATA	COMENTÁRIO
Resol. SE n. 135	18/07/1994	Cria o Centro de Apoio Pedagógico para Atendimento ao Deficiente Visual (CAP-DV).
Resol. SE n. 61	05/04/2002	Programa de Inclusão Escolar – Transforma o CAP-DV em Centro de Apoio Pedagógico Especializado (Cape).
Resol. SE n. 32	23/05/2007	Ações do Programa de AEE.
Resol. SE n. 72	09/10/2009	Parcerias da Secretaria com instituições sem fins lucrativos. Serviu de base para as resoluções seguintes.
Resol. SE n. 14	07/02/2012	Celebra convênio com entidades de fins não econômicos, para proporcionar atendimento e apoio a alunos com deficiência, matriculados em escolas da rede estadual de ensino, e dá providências correlatas.
Resol. SE n. 70	29/06/2012	Altera dispositivo da Resolução SE n. 54, de 12/08/2011, que dispõe sobre a celebração de convênios com instituições, sem fins lucrativos, atuantes em educação especial.
Resol. SE n. 59	22/11/2016	Dispõe sobre o processo de credenciamento de organizações da sociedade civil sem fins lucrativos, a que alude o art. 30, inciso VI, da Lei Federal n. 13.019, de 31/07/2014.
Resol. SE n. 68	12/12/2017	Dispõe sobre o atendimento educacional aos alunos PAEE na rede estadual de ensino.
Resol. SE n. 15	16/02/2018	Dispõe sobre o processo de credenciamento de organizações da sociedade civil sem fins lucrativos, a que alude o art. 30, inciso VI, da Lei Federal n. 13.019, de 31/07/2014.
Decreto n. 57.141	18/07/2011	Reestrutura a SEE/SP – Centro de Atendimento Especializado (Caesp), que passa a aglutinar o Cape e o Núcleo de Inclusão (Ninc).
Resol. SE n. 81	07/08/2012	Aceleração de estudantes com altas habilidades e superdotação.
Resol. SE n. 32	17/05/2013	Define as atribuições do Núcleo de Apoio Pedagógico Especializado – Cape.

(Continua)

RESOLUÇÃO	DATA	COMENTÁRIO
Resol. SE n. 18	03/04/2014	Trata do Termo de Ajuste de Conduta com o Ministério Público do estado de São Paulo sobre acessibilidade nos prédios escolares.
Resol. SE s/n	08/12/2016	Trata da Educação Especial no sistema estadual de ensino - Deliberação do Conselho Estadual de Educação n. 149/2016.
Resol. SE n. 68	12/12/2017	Dispõe sobre o AEE.

Fonte: Elaboração própria. Base de legislações da Secretaria de Educação do estado de São Paulo.

Além dessas resoluções, a SEE-SP desenvolveu programas que incluíam ações de escolarização dos estudantes PAEE na rede estadual de ensino do estado de São Paulo (REE/SP); são eles: o Programa Melhoria da Qualidade do Ensino Fundamental (2000-2011), o Programa Ler e Escrever e o Programa Gestão Pedagógica da Educação Básica (2012-2016), ambos do Departamento de Desenvolvimento Curricular e de Gestão da Educação Básica.

Assim, passados pouco mais de trinta anos entre políticas integracionistas e, principalmente, inclusivistas, vários pesquisadores da área, professores e profissionais que trabalham com esse público vêm chamando a atenção para os indicadores educacionais censitários como estratégia de avaliação das políticas de Educação Especial. Isso porque os dados obtidos pelo censo escolar do Inep, por meio de sinopses ou microdados estatístico-educacionais, são empregados para fins de formulação de políticas públicas e de execução de programas na área educacional, inclusive para definição dos valores dos recursos públicos federais transferidos para os estados e municípios (BRASIL, 2018). Destacam-se, a esse respeito, os trabalhos de Meletti e Bueno (2010), Mendes (2010), Meletti (2014), Meletti e Ribeiro (2014), Rebelo e Kassar (2018) e Santos e Mendes (2019).

Enfim, no tocante à política de Educação Especial no estado de São Paulo, faz-se necessário analisar os números fornecidos pelos censos escolares do Inep questionando se esses representam aumento do acesso à escola e garantia de direito à escolarização dos estudantes PAEE.

METODOLOGIA

Para este estudo foram utilizados como fonte os microdados dos censos escolares referentes aos anos de 2007, 2011 e 2015. Embora seja chamado de censo escolar pelo Inep – posto que ambiciona contabilizar anualmente, desde 1995, todas as matrículas realizadas no sistema educacional brasileiro –, este é um registro administrativo, uma vez que o formulário do censo é inteiramente respondido através de plataforma digital por funcionários das escolas, sem que exista qualquer tipo de entrevista direta com os estudantes.

Vale destacar que, entre 1995 e 2006, os dados eram coletados de forma agregada, ou seja, cada escola informava quantos estudantes apresentavam determinada

característica e, por conseguinte, não é possível realizar o cruzamento de duas ou mais características dos estudantes. Em 2007, o levantamento sofreu significativa alteração de formato e passou-se a coletar dados individualizados por estudante, prática que se mantém até o momento presente. O censo conta atualmente com quatro formulários, que apresentam distintas unidades básicas de análise: matrículas, turmas, escolas e docentes.

No período analisado, é possível verificar mudanças nos quesitos que dizem respeito ao Público-Alvo da Educação Especial. Em 2007, por exemplo, a síndrome de Down constava entre as categorias de necessidade educacional especial; porém, a partir de 2009, ela passou a ser contabilizada na categoria de deficiência mental,⁶ ocasionando um aumento considerável de tais casos.

Além dos dados sobre matrículas na educação básica do PAEE, também são de interesse, para fins comparativos, as informações referentes aos estudantes que não foram incluídos como PAEE nos censos (deficiência sensorial, física, mental/intelectual, transtorno ou altas habilidades/superdotação), por não fazerem parte das categorias abarcadas por essas políticas públicas, denominados neste trabalho Não Público-Alvo da Educação Especial (NPAEE).

As variáveis selecionadas para esta pesquisa foram ano de nascimento do estudante, ano do censo, série/ano da etapa de ensino que frequenta, se apresenta alguma necessidade especial e dependência administrativa da escola, isto é, se pertence à rede federal, estadual, municipal ou privada de ensino. Quanto à comparação entre categorias de dependência administrativa, optou-se por contrastar, sobretudo, as redes estadual e privada.

CONSTRUÇÃO DO BANCO DE DADOS E TRATAMENTO DAS VARIÁVEIS

Uma vez selecionados os dados para o estado de São Paulo, foram realizadas análises descritivas para dimensionar o PAEE e o NPAEE.

A partir da aplicação de critérios utilizados pelo Inep e pelo Instituto Brasileiro de Geografia e Estatística (IBGE), foi realizada a construção da taxa de distorção idade-série, que expressa o percentual de estudantes com idade superior à idade recomendada em cada série/ano dos ensinos fundamental e médio, cuja fórmula é dada por:

$$TDIS_{ks} = \frac{M_{ksa_sup}}{M_{ks}} * 100 \quad (1)$$

Nessa fórmula, $TDIS_{ks}$ é a taxa de distorção idade-série no nível de ensino k

6 Termo utilizado pelo Inep até 2013.

em sua série/ano; M_{ksa_sup} é o número de matrículas de estudantes que nasceram em anos anteriores ao ano de nascimento dos estudantes que cumprem a idade recomendada para o nível de ensino k e na série/ano no ano de realização do censo escolar; M_{ks} é o número total de matrículas no nível de ensino k na série/ano. Na prática,

[...] considerando o Censo Escolar do ano t e a série k do ensino fundamental, cuja a idade adequada é de i anos, então o indicador será expresso pelo quociente entre o número de alunos que, no ano t , completam $i + 2$ anos ou mais (nascimento antes de $t - [i + 1]$), e a matrícula total na série k . A justificativa deste critério é que os alunos que nasceram em $t - [i + 1]$, completam $i + 1$ anos no ano t e, portanto, em algum momento deste ano (de 1º de janeiro a 31 de dezembro) ainda permanecem com i anos e, por isso, o critério aqui adotado, considera estes alunos como tendo idade adequada para esta série. Os que nasceram depois de $t - [i + 1]$ completam, no ano t , i anos ou menos. (BRASIL, 2012)

Deve-se destacar que foram consideradas para a análise as matrículas nas turmas seriadas dos ensinos fundamental e médio, sendo desconsideradas as matrículas realizadas no ensino médio nas modalidades magistério, profissional, técnico e Educação de Jovens e Adultos (EJA). Atenta-se para o fato de que, ao longo do período em análise, a legislação foi alterada, e o ensino fundamental passou a conter nove anos de duração, em vez de oito, conforme a Lei n. 11.274/2006 (BRASIL, 2006), de modo que, para apresentar os resultados, foi necessário fazer a compatibilização entre essas duas lógicas, que ainda coexistem, através da soma das matrículas realizadas na primeira série do ensino de oito anos com a soma das matrículas no segundo ano do ensino de nove anos, e assim por diante, tendo o ensino fundamental de nove anos como referência.

Após a recodificação das variáveis necessárias para a avaliação da distorção idade/série entre os estudantes da educação básica, as análises descritivas apresentadas contrastam a situação de distorção idade/série dos matriculados na educação básica no estado de São Paulo, PAEE e NPAEE, segundo dependência administrativa das escolas, enfatizando as diferenças entre as redes estadual e privada.

Para além da distorção idade/série, foram avaliados aspectos referentes à política de educação inclusiva, tais como: se esses estudantes estão matriculados em classes exclusivas (modalidade substitutiva) ou classes comuns, tal como defendido pela perspectiva da educação inclusiva; se frequentam o AEE; e se as escolas que ofertam o AEE possuem salas de recursos multifuncionais. Em todas as etapas do trabalho foi utilizado o *software* estatístico SPSS.

RESULTADOS

A Tabela 1 traz dados sobre o total de matrículas de alunos PAEE e sua proporção em relação ao total de matrículas realizadas no estado de São Paulo nos anos de 2007, 2011 e 2015.

TABELA 1 – São Paulo, 2007, 2011 e 2015: Prevalência de matrículas de estudantes PAEE na educação básica

ANO	MATRÍCULAS DE ESTUDANTES PAEE	TOTAL DE MATRÍCULAS	PROPORÇÃO DE MATRÍCULAS DE ESTUDANTES PAEE NO TOTAL DE MATRÍCULAS
2007	156.443	10.651.119	1,5
2011	210.263	10.797.800	1,9
2015	219.925	11.162.580	2,0

Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 – Inep.

Segundo os dados apresentados, houve aumento no número de estudantes PAEE ao longo do período analisado: em 2007, os estudantes PAEE representavam 1,5% do total de matrículas no estado de São Paulo; em 2011, essa proporção aumentou para 1,9%; e, em 2015, o índice subiu para 2,0%. Entre 2017 e 2015, a representatividade de estudantes PAEE aumentou 0,5% entre os alunos matriculados; em números absolutos, houve um aumento de 63.482 matrículas quando se comparam os anos de 2007 e 2015. Portanto, há um aumento do acesso das pessoas que são PAEE à rede de ensino no estado de São Paulo.

Com o objetivo de adensar as análises, faz-se necessário identificar o modelo de escolarização utilizado pela rede de ensino paulista – classes comuns com apoio do AEE ou classes exclusivas – para o atendimento dos estudantes PAEE, tendo em vista que a legislação brasileira, especialmente o PNE e a Lei n. 13.146 (BRASIL, 2015), conhecida como Lei Brasileira de Inclusão (LBI), retoma a ambiguidade histórica da área, possibilitando que as matrículas dos estudantes PAEE sejam identificadas em classes comuns ou exclusivas – apesar de o PNEE-EI definir a matrícula no AEE ou Centro de Atendimento Educacional Especializado (CAEE) como não substitutivo às classes comuns.

A Tabela 2 contém dados relativos aos estudantes PAEE que frequentam classe exclusiva no estado de São Paulo. Ainda que se tenha verificado *missing* de 20,5% em 2011 e de 24% em 2015, em relação aos dados de 2007, considera-se pertinente apresentá-los, pois permitem refletir, por aproximação, sobre as condições de inclusão (modelo de escolarização) das pessoas PAEE.

TABELA 2 – São Paulo, 2007, 2011 e 2015: percentual de estudantes PAEE que frequentam classe exclusiva

ANO	NÚMERO DE MATRÍCULAS	PERCENTUAL
2007	64.749	41,4%
2011	49.998	29,9%*
2015	40.771	24,4%*

Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 – Inep.

*Considerando-se apenas os casos válidos, pois há 20,5% de *missing* em 2011 e 24% de *missing* em 2015 na variável que identifica classe exclusiva.

Verifica-se, na tabela acima, uma tendência de diminuição do atendimento educacional em classes exclusivas: em 2007, dos 156.443 estudantes PAEE matriculados, 41,4% estavam em classes exclusivas; em 2011, esse percentual diminuiu para 29,9%; e, em 2015, 24,4% dos estudantes PAEE matriculados na rede de ensino paulista estavam em classes exclusivas.

Infere-se que esse movimento de diminuição do atendimento do PAEE na rede de ensino paulista em classes exclusivas pode ter sido impulsionado, a partir de 2008, pela PNEE-EI, que induziu, por meio de financiamentos obtidos em editais do governo federal, as redes de ensino brasileiras a aderirem à política inclusiva.

Considerando ainda que, a partir de 2008, define-se o AEE como forma de se materializar uma educação inclusiva do PAEE no ensino comum, contando com serviços complementares ou suplementares, é pertinente trazer à tona, a seguir, os dados relativos ao AEE.

TABELA 3 – São Paulo, 2007, 2011 e 2015: percentual de estudantes PAEE que recebem atendimento educacional especializado

ANO	NÚMERO DE MATRÍCULAS	PERCENTUAL
2007	Informação não coletada	-
2011	38.123	18,1%
2015	39.241	17,8%

Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 – Inep.

Definido por lei a partir de 2008, não há dados relativos ao AEE no ano de 2007. Em 2011, dos 210.263 estudantes PAEE matriculados, 38.123 tiveram AEE, o que representa somente 18,1% do total de matrículas nessa modalidade de ensino. Em 2015, 17,8% dos 219.925 estudantes matriculados receberam AEE.

Esse decréscimo do número de estudantes PAEE matriculados nas classes comuns e que recebem apoio do AEE nos dados de 2015, conforme Silva, Machado e Silva (2019), pode estar relacionado com as mudanças ocorridas no contexto sociopolítico e econômico do Brasil a partir de 2014. Tais mudanças geraram consequências, inclusive, na definição da meta 4 do PNE (2014-2024), que retomou o

termo “preferencialmente” e toda sua ambiguidade histórica para a definição dos modelos de escolarização dos estudantes PAEE.

Na Tabela 4, a seguir, buscamos mapear, a partir dos dados do censo escolar, a distribuição dos estudantes PAEE e NPAEE nos anos/séries do ensino fundamental (1º ao 9º ano) e ensino médio (1º ao 3º ano), o que nos permite, a partir desses dados, realizar uma comparação do fluxo de estudantes matriculados, tanto em relação ao ritmo de escolarização entre aqueles estudantes Público-Alvo da Educação Especial quanto destes em relação aos demais estudantes.

Cabe lembrar que, com a Lei n. 11.274/2006 (BRASIL, 2006), foi instituído o ensino fundamental de 9 anos, o que exigiu a compatibilização dos dados. Além disso, em 2008, foi criada a PNEE-EI que, dentre as mudanças propostas, insere a redefinição da classificação do PAEE. Essas mudanças legais foram consideradas para a análise do movimento das matrículas, sobretudo dos dados de 2007 e 2011.

TABELA 4 - São Paulo, 2007, 2011 e 2015: número de matrículas do PAEE e NPAEE conforme ano e etapa do ensino

ANO E NÍVEL	PAEE			NPAEE		
	2007	2011	2015	2007	2011	2015
1º ano EF	14.870	23.147	14.360	203.624	572.968	599.865
2º ano EF	32.864	16.940	14.649	690.495	564.730	602.838
3º ano EF	33.081	12.096	16.762	763.492	493.934	627.618
4º ano EF	15.192	15.484	14.833	706.299	619.055	580.704
5º ano EF	17.252	17.988	16.602	748.231	664.535	594.089
6º ano EF	9.544	14.308	13.357	733.277	703.927	586.727
7º ano EF	6.861	12.363	10.372	722.281	748.338	499.259
8º ano EF	4.661	9.319	12.075	686.673	708.836	625.853
9º ano EF	4.141	7.568	11.240	700.795	721.387	648.995
1º ano EM	2.243	4.785	8.486	685.229	727.749	687.265
2º ano EM	1.502	3.041	5.949	551.727	608.363	605.865
3º ano EM	1.067	1.958	3.831	467.254	504.457	498.783
Total	143.278	138.997	142.516	7.659.377	7.638.279	7.157.861

Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 - Inep.

Quanto ao fluxo de escolarização dos estudantes PAEE, destaca-se que, enquanto no ciclo do ensino fundamental I (1º ao 5º ano) há, mesmo com alguma diferenciação, relativa estabilidade no número de matrículas, a partir do ensino fundamental II (6º ao 9º ano) há uma tendência de decréscimo nas matrículas.

Comparando-se os dados das matrículas dos estudantes PAEE ao final da educação básica com o total de matriculados nessa modalidade de ensino, os dados demonstram que, em 2007, os que chegam ao 3º ano do ensino médio representam 0,74% do total; em 2011, esse percentual cresce um pouco, chegando a 1,41%; e, em 2015, mais que triplica em relação a 2007, alcançando o índice de 2,69%

dos estudantes matriculados como PAEE. De forma menos acentuada, percebe-se essa tendência de decréscimo nas matrículas ao longo dos níveis de escolarização, destacando-se o ensino médio e, sobretudo, o 3º ano do ensino médio.

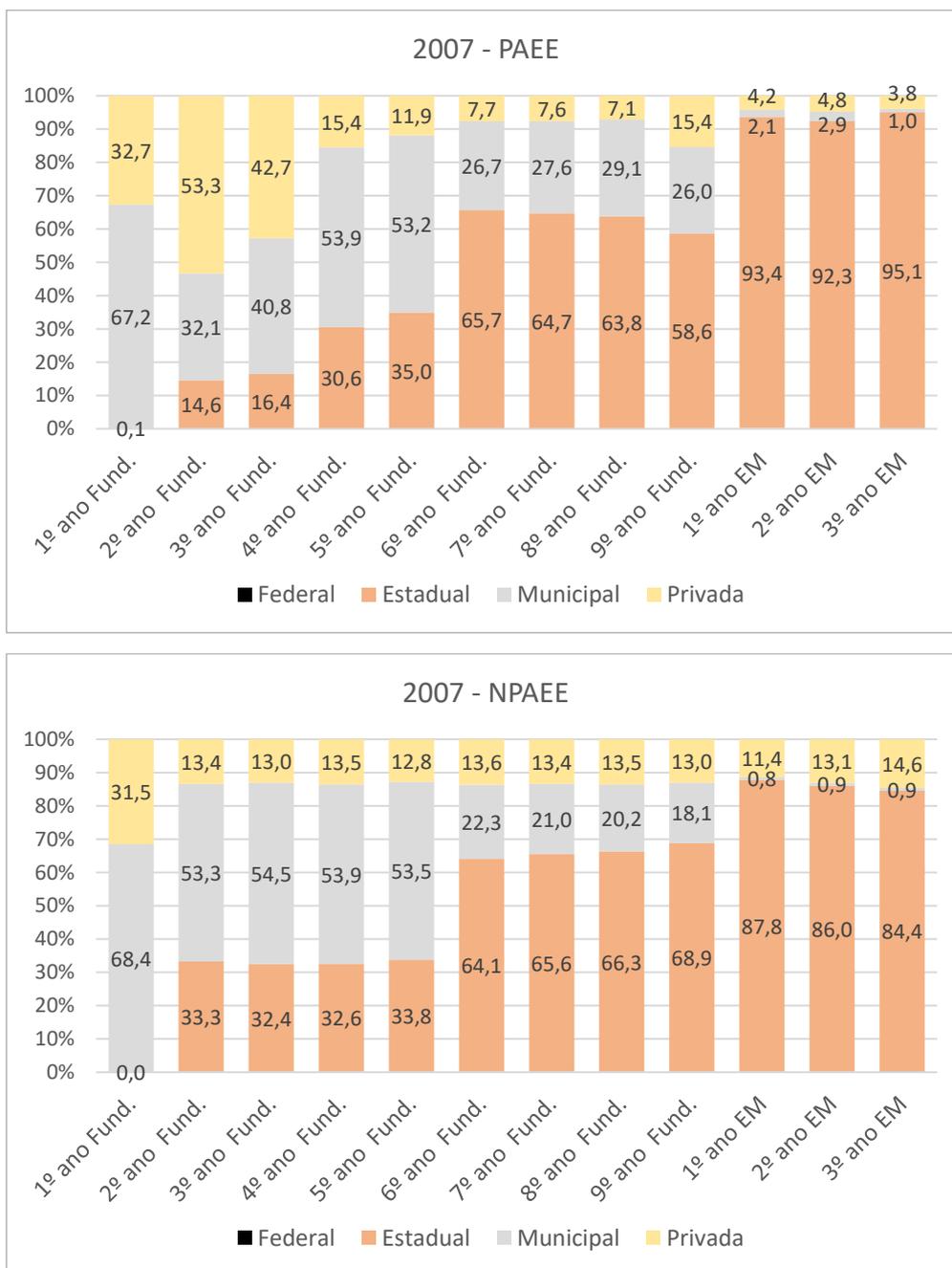
A comparação entre os dados dos estudantes matriculados no 3º ano do ensino médio com o total de matrículas nos anos de 2007, 2011 e 2015 demonstra que, em geral, estes representam pouco mais de 6% do total de matrículas: em 2007, 6,1% dos estudantes NPAEE matriculados encontravam-se no 3º ano do ensino médio; em 2011, esse percentual subiu para 6,6% e, em 2015, chegou a 6,97%.

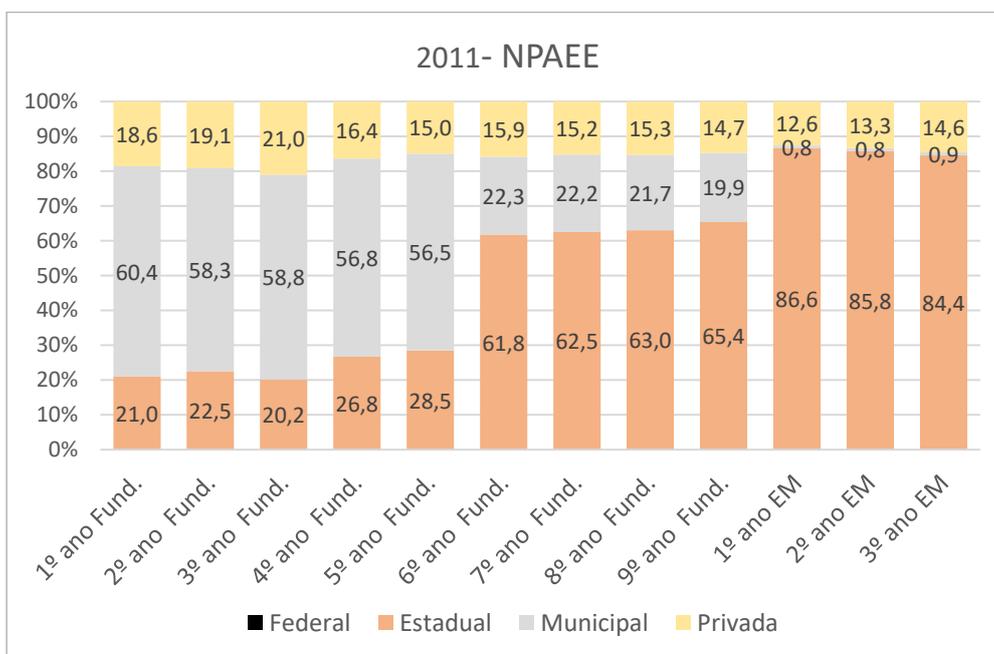
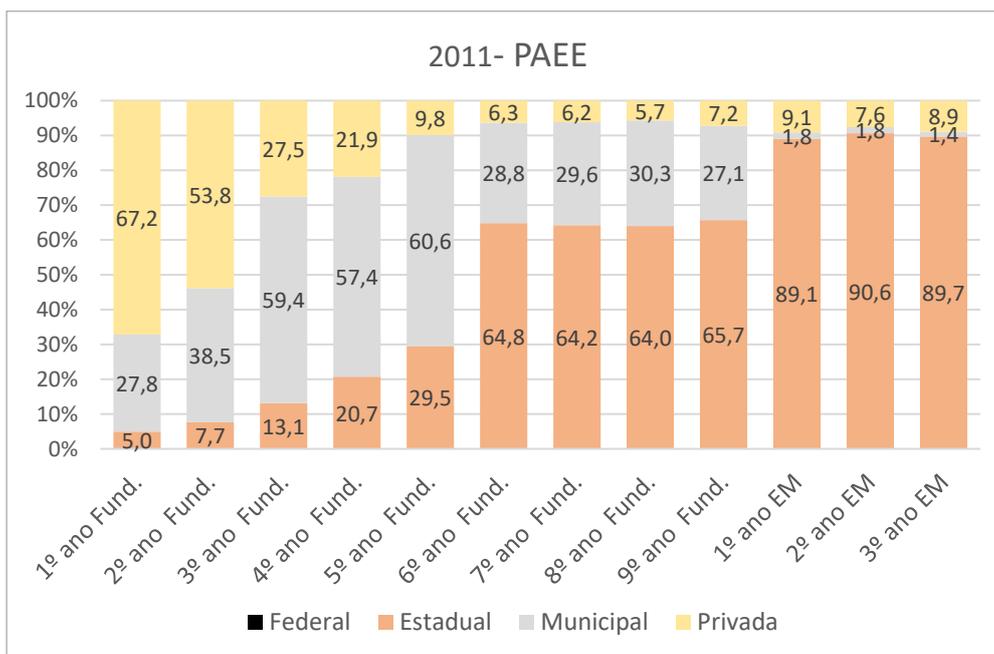
Os dados revelam que, em relação ao NPAEE, o fluxo de escolarização dos estudantes PAEE é mais afunilado, o que permite inferir que os mecanismos que provocam evasão e/ou repetência impactam mais acentuadamente os estudantes em situação de deficiência. Pergunta-se: o modelo vigente de organização do trabalho educativo da escola contemporânea poderia ser um dos mecanismos internos que explicariam esses dados?

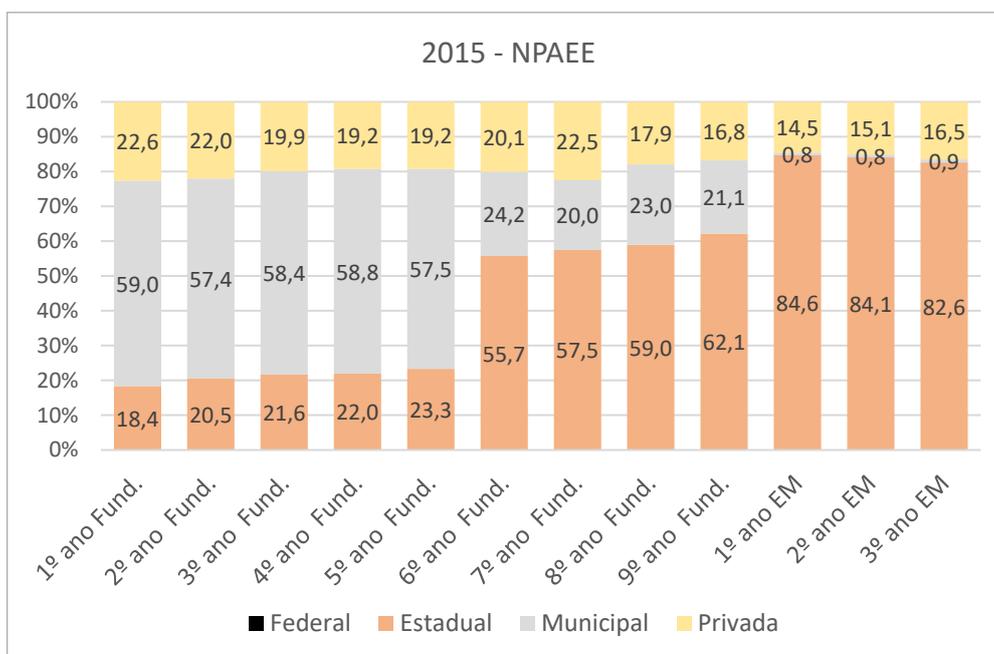
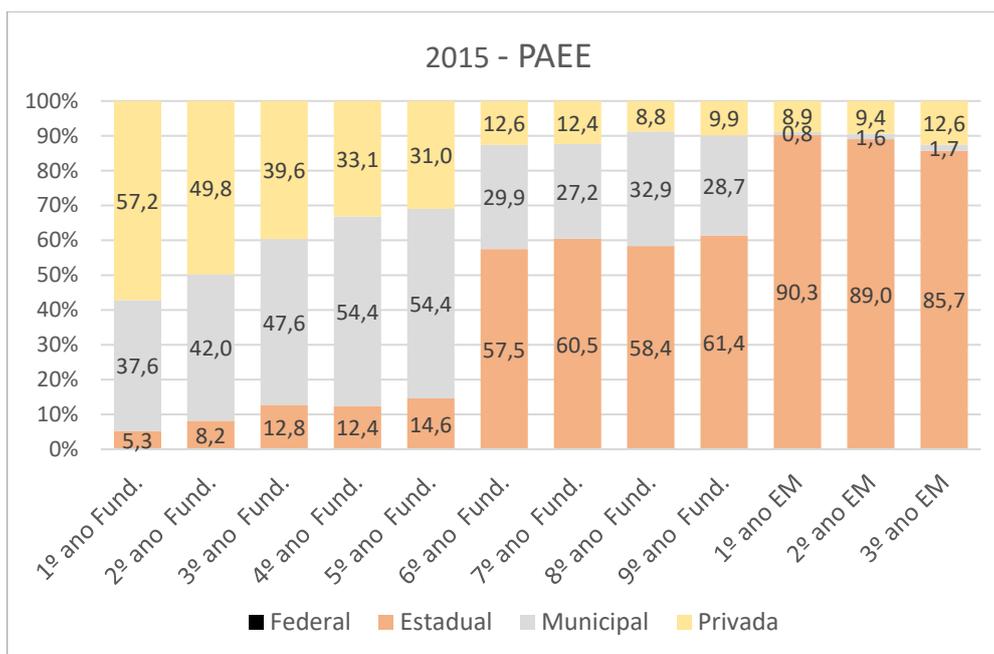
Abaixo, selecionamos os dados relativos aos estudantes PAEE e NPAEE distribuídos por dependência administrativa (federal, estadual, municipal e privada), buscando, com isso, identificar onde, no ensino comum, prioritariamente, os estudantes têm sido atendidos.

Classificam-se os “sistemas” de ensino como públicos quando ofertados pelo Estado, e privados quando mantidos por instituições educacionais constituídas juridicamente na esfera da sociedade civil. Além disso, tendo em vista que o Brasil é um país organizado na forma federativa, o atendimento estatal se dá de modo colaborativo entre os entes federados (federação, estados, distrito e municípios). Cabe aos municípios a oferta da educação infantil e, prioritariamente, ensino fundamental. Aos estados, cabe a oferta do ensino médio e, de forma colaborativa com os municípios, atuar, também, na garantia do ensino fundamental. Ao Distrito Federal aplicam-se as competências relativas tanto aos estados quanto aos municípios. O governo federal, em termos de atendimento, é responsável pela rede federal de ensino (BRASIL, 1996).

FIGURA 1 - São Paulo, 2007, 2011 e 2015: Distribuição percentual das matrículas de estudantes PAEE e NPAEE segundo dependência administrativa por etapa de ensino







Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 - Inep.

As figuras acima, ao trazerem dados relativos às matrículas de estudantes PAEE e NPAEE nos anos de 2007, 2011 e 2015, distribuídas por dependência administrativa, revelam algumas tendências gerais do atendimento escolar.

Em relação à rede federal de ensino, não há dados que permitam a análise, visto que a educação básica não é de sua competência.

Quanto à rede estadual, percebe-se um aumento no percentual de matrículas de estudantes PAEE a partir do 4º ano do ensino fundamental em todos os anos analisados, mantendo-se oferta majoritária no ensino fundamental II (6º ao 9º

ano), com pouco mais de 50% das matrículas em todos os anos e, sobretudo, no ensino médio, com a quase totalidade dos estudantes.

No entanto, cabe destacar algumas variações nos índices no que se refere à diminuição gradativa de oferta no ensino fundamental I (1º ao 5º ano). Assim, os dados obtidos do 2º ao 5º ano do ensino fundamental permitem verificar os seguintes resultados: em 2007, 14,6% das matrículas no 2º ano encontravam-se na rede estadual; em 2011, esse percentual diminuiu para 7,7%; e, em 2015, houve uma pequena elevação, chegando a 8,2%. O mesmo movimento se dá em relação ao 3º ano do ensino fundamental: em 2007, verificou-se um atendimento de 16,4%, que caiu para 13,1% em 2011 e chegou a 12,8% em 2015; no 4º ano do ensino fundamental, de um percentual de 30,6% de matrículas em 2007, houve uma diminuição de 20,7% em 2011 para 12,4% em 2015; no 5º do ensino fundamental, essa tendência se repetiu: em 2007 havia 35,0% de matrículas na rede estadual, que passaram a 29,5% em 2011 e 14,6% em 2015. Essa tendência de diminuição na oferta na primeira etapa do ensino fundamental I se manteve, também, nas matrículas dos estudantes NPAEE, com pouco mais de 30% de matrículas nos anos correspondentes ao ensino fundamental I em 2007 e chegou, em 2015, a cerca de 20% das matrículas.

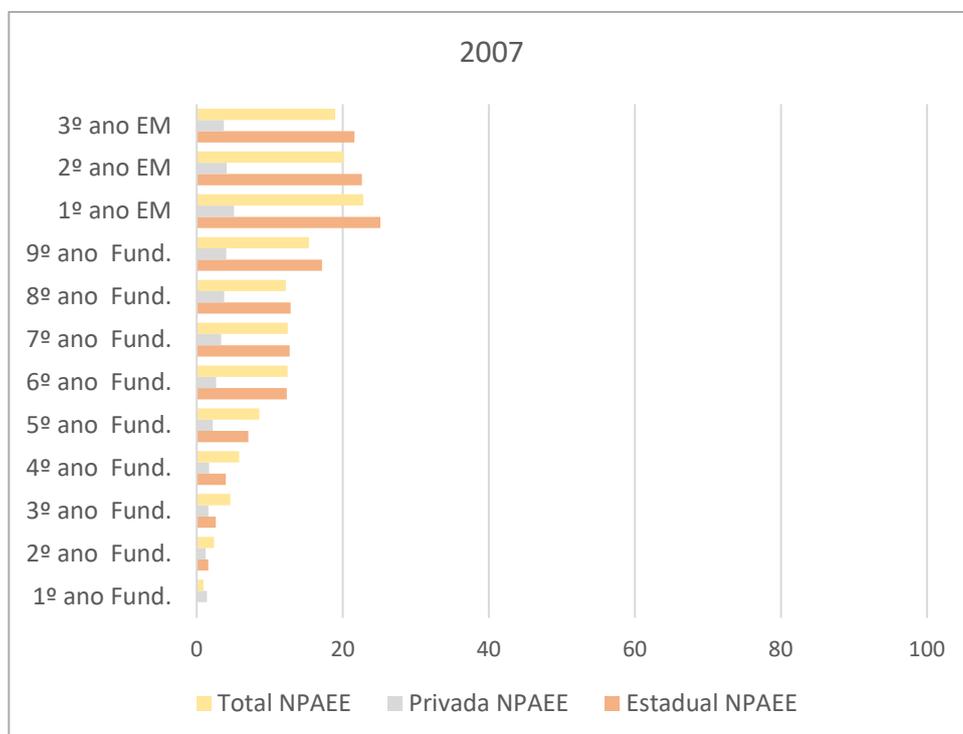
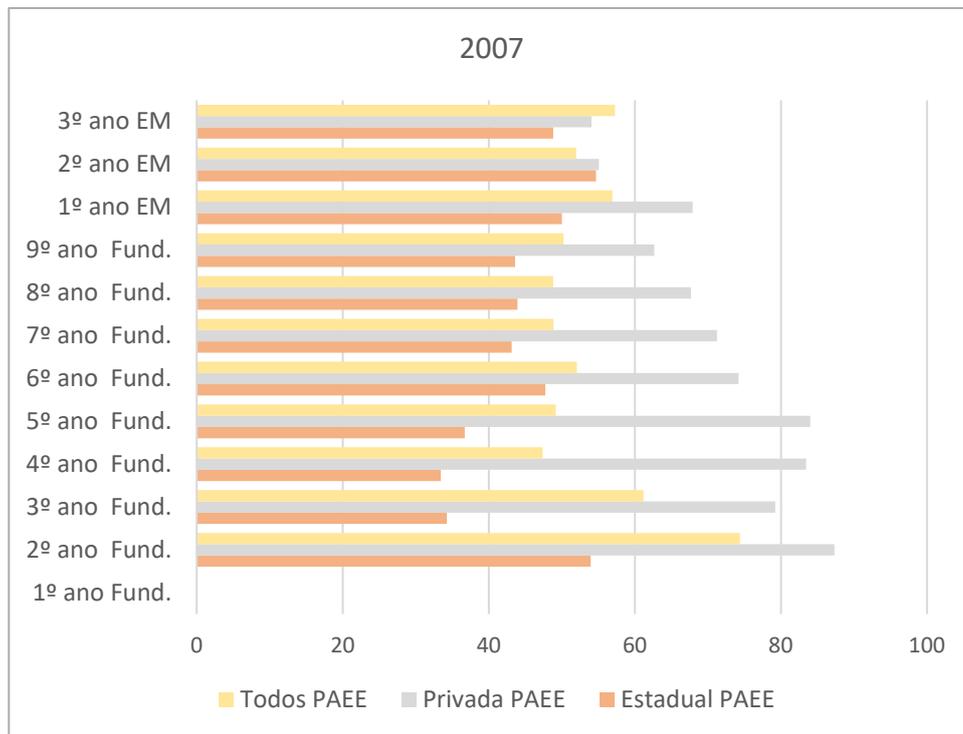
Percebe-se, pela Figura 1, que a diminuição de matrículas na rede estadual relativas ao ensino fundamental I (1º ao 5º ano) foi acompanhada por um aumento do número de matrículas nas redes municipal e privada de ensino, tanto no que se refere ao PAEE quanto ao NPAEE.

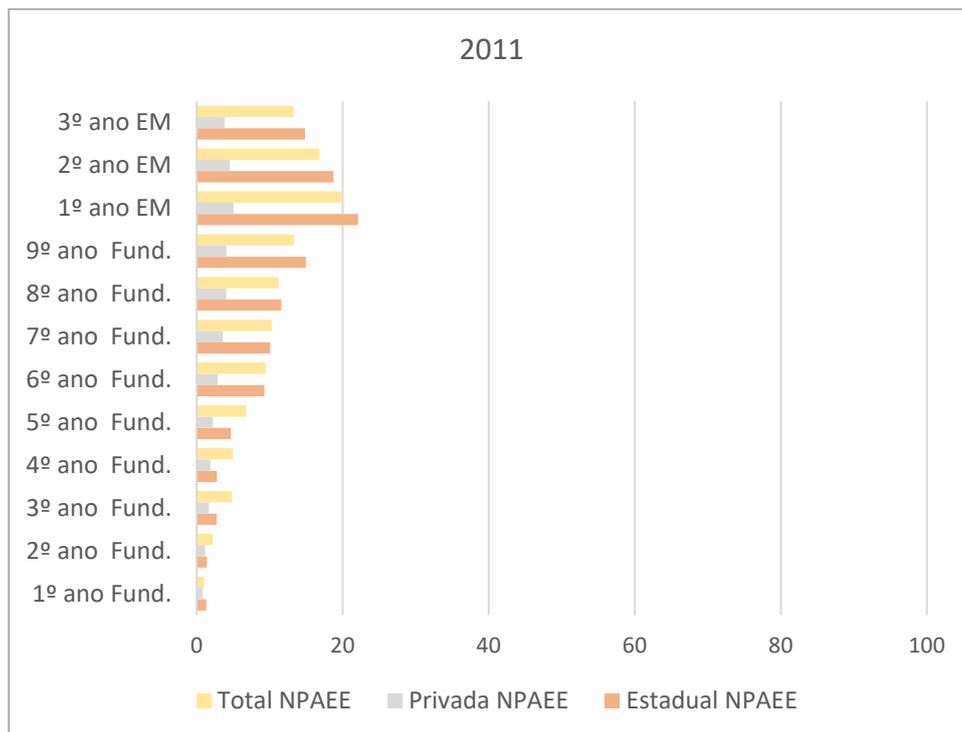
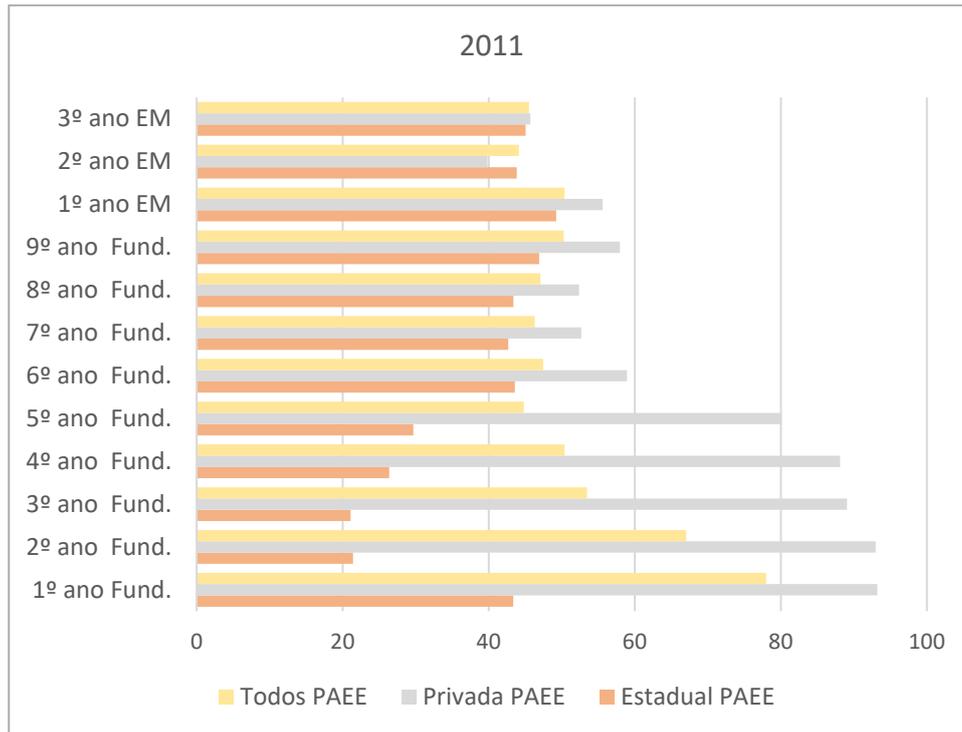
Em relação às redes municipal e privada, ainda, evidenciou-se, tanto no PAEE quanto no NPAEE, uma queda no percentual de matrículas no ensino fundamental II (6º ao 9º ano), tendência acentuada no ensino médio.

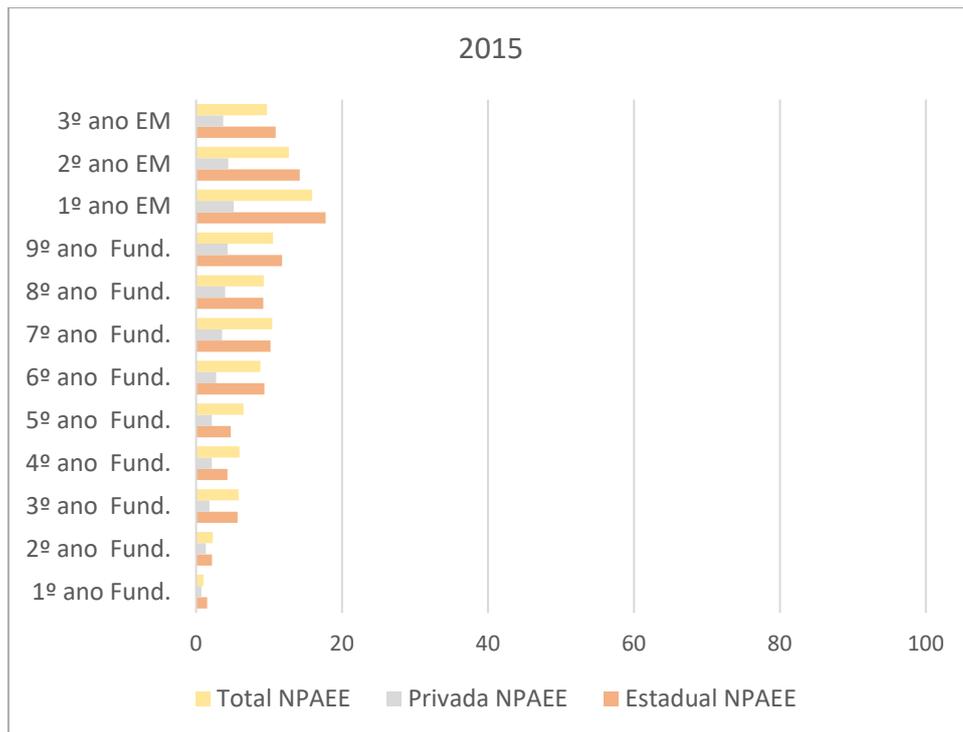
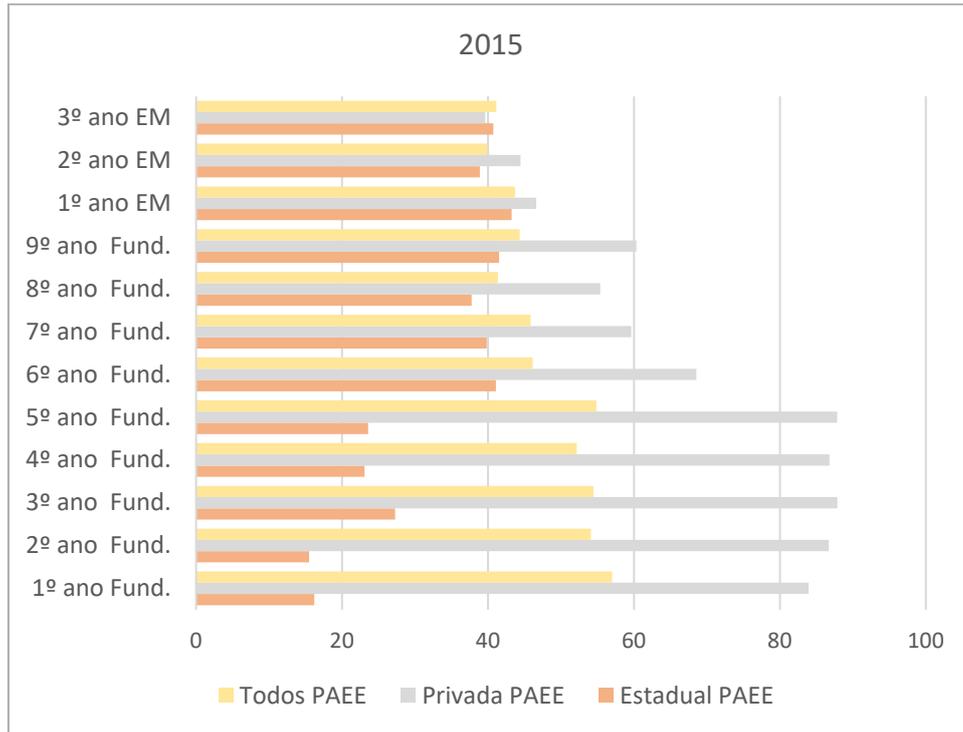
Os dados da Figura 1 e as análises deles decorrentes permitem inferir, em relação à educação básica, que as redes municipal e privada de ensino detêm a maioria das matrículas no ensino fundamental I, enquanto a rede estadual passa, a partir do ensino fundamental II e, sobretudo, no ensino médio, a atender a maioria dos estudantes matriculados.

Desse modo, será apresentado a seguir o percentual de matrículas de estudantes em situação de distorção idade/série distribuído por dependência administrativa tanto para o PAEE quanto para o NPAEE, segundo a classificação de ensino público (estatal) e privado. Essa escolha se justifica pelo fato de que os estudos e pesquisas na área de Educação Especial têm privilegiado dados relativos às redes municipais (CAIADO; JESUS; BAPTISTA, 2018), além de a REE/SP hegemonizar o atendimento educacional nos níveis fundamental II e médio.

FIGURA 2 – São Paulo, 2007, 2011 e 2015: distribuição percentual das matrículas de estudantes em situação de distorção idade/série segundo PAEE e NPAEE e dependência administrativa por etapa de ensino







Fonte: Elaboração própria. Base de matrículas dos censos escolares de 2007, 2011 e 2015 – Inep.

Destacou-se, preliminarmente, na Figura 2, o contraste na distorção idade/série dos estudantes PAEE em relação aos NPAEE. O percentual referente ao PAEE é significativamente maior. Pode-se levantar, ao menos, três hipóteses, não excludentes, mas complementares, para explicar esse fenômeno. Uma diz respeito

à possibilidade de as crianças/jovens em idade escolar em situação de deficiência iniciarem mais tarde a sua escolarização. A segunda hipótese se refere aos efeitos que os processos de avaliação realizados ao longo do processo de escolarização podem causar sobre a trajetória escolar dos estudantes PAEE. Finalmente, atrelado à avaliação, encontra-se o modelo homogeneizador de organização do trabalho educativo na escola contemporânea, que toma como pressuposto um padrão de desenvolvimento e aprendizagem em que todos os seres humanos considerados “normais” estariam aptos, evidenciando as possíveis “deficiências” daqueles que não alcançarem as metas previamente definidas.

Quaisquer dessas hipóteses, contudo, não são passíveis de verificação com base em dados do censo escolar. Fazem-se necessárias, para tanto, pesquisas empíricas sobre a trajetória escolar desses estudantes e a verificação dos modos de organização do trabalho educativo das instituições escolares. Ressalta-se, todavia, a importância de se realizar uma análise como aqui proposta, tendo em vista que ainda está em causa o direito ao acesso à escola por pessoas em situação de deficiência.

Destaca-se, ainda, o fato de que a distorção idade/série dos estudantes Público-Alvo da Educação Especial é maior na rede privada de ensino se comparada com a rede estadual paulista. Essa situação se inverte ao se analisar os percentuais relativos aos estudantes NPAEE: a distorção é mais acentuada na rede estadual do que na rede privada de ensino no estado de São Paulo.

Os dados aqui analisados, em confronto com as diretrizes emanadas da Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva, permitem refletir sobre os desafios históricos enfrentados no processo de integração/inclusão das pessoas Público-Alvo da Educação Especial.

CONSIDERAÇÕES FINAIS

Considerando a histórica luta das pessoas em situação de deficiência por direitos sociais, incluindo o direito à educação, que desembocou, no Brasil, no PNEE-EI, em 2008, a escolarização dos estudantes Público-Alvo da Educação Especial nas escolas comuns passou a ser inserida na pauta política da REE/SP.

Assim, julgamos pertinente verificar as condições de acesso e permanência do PAEE na REE/SP por meio das informações disponibilizadas pelos censos escolares do Inep, principalmente aqueles relativos às matrículas estudantis.

Dessa maneira, os dados de matrícula demonstram um aumento gradativo de estudantes PAEE nas escolas comuns da REE/SP e revelam, pelos dados do fluxo escolar, que há, internamente, processos marcadamente excludentes.

Levantamos a hipótese de que esses processos internos de exclusão, caracterizados pela evasão/retenção, podem estar relacionados ao modelo homogeneizador

da organização do trabalho educativo das escolas e são aprofundados pelas políticas de avaliação e currículo da REE/SP.

Outra hipótese que pode ser considerada na interpretação dos resultados obtidos é o fato de que, em função do histórico da área – comum *versus* especial, preceitos médico-clínicos *versus* preceitos sociopedagógicos – e das legislações – Constituição Federal de 1988 (BRASIL, 1988); Lei de Diretrizes e Bases da Educação Nacional –, Lei n. 9.394 (BRASIL, 1996); PNEE-EI de 2008; PNE 2014-2024; LBI; dentre outras –, mantém-se a ambiguidade na definição do modelo e a instituição responsável pela escolarização do PAEE, bem como a forma de relacionamento entre a educação comum e a especial.

Em síntese, não obstante um processo histórico de luta pela escolarização por parte do PAEE, ainda são muitos os desafios e as barreiras (as dicotomias e as estruturas sociais excludentes) a serem superadas para que todas as pessoas, inclusive aquelas em situação de deficiência e consideradas PAEE, consigam garantir o direito social à educação na REE/SP.

REFERÊNCIAS

BRASIL. Ministério da Educação. Secretaria de Educação Especial. *Política Nacional de Educação Especial*. Brasília: MEC/SEESP, 1994.

BRASIL. Presidência da República. *Lei n. 9.394, de 20 de dezembro de 1996*. Estabelece as diretrizes e bases da educação nacional. Brasília, 1996. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L9394.htm. Acesso em: dez. 2020.

BRASIL. Presidência da República. Secretaria Geral. Subchefia para Assuntos Jurídicos. *Lei n. 11.274, de 06 de fevereiro de 2006*. Altera a redação dos arts. 29, 30, 32 e 87 da Lei n. 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, dispondo sobre a duração de 9 (nove) anos para o ensino fundamental, com matrícula obrigatória a partir dos 6 (seis) anos de idade. Brasília, 2006.

BRASIL. Ministério da Educação. Secretaria de Educação Especial. *Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva*. Brasília: MEC/SEESP, 2008a.

BRASIL. Presidência da República. *Decreto n. 6.571, de 17 de setembro de 2008*. Dispõe sobre o atendimento educacional especializado, regulamenta o parágrafo único do art. 60 da Lei n. 9.394, de 20 de dezembro de 1996, e acrescenta dispositivo ao Decreto n. 6.253, de 13 de novembro de 2007. Brasília, 2008b.

BRASIL. Conselho Nacional de Educação. Câmara de Educação Básica. *Resolução n. 4, de 2 de outubro de 2009*, que institui diretrizes operacionais para o atendimento educacional especializado na educação básica, modalidade educação especial. Brasília, 2009.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Planejamento e Investimentos Estratégicos. *Plano Plurianual 2012-2015*. Mensagem presidencial. Brasília, 2011.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep. *Taxas de distorção idade-série Escolar na Educação Básica*. Brasília, 2012. Disponível em: <https://dados.gov.br/dataset/taxas-de-distorcao-idade-serie-escolar-na-educacao-basica>. Acesso em: jun. 2018.

BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. *Lei n. 13.005, de 25 de junho de 2014*. Aprova o Plano Nacional de Educação – PNE e dá outras providências. Brasília, 2014.

BRASIL. Presidência da República. Secretaria Geral. Subchefia para Assuntos Jurídicos. *Lei n. 13.146, de 06 de julho de 2015*. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, 2015.

BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. *Constituição da República Federativa do Brasil de 1988*. Brasília: Presidência da República, 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 14 fev. 2020.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep. *Censo Escolar*. Brasília, 2018. Disponível em: <http://inep.gov.br/censo-escolar>. Acesso em: jun. 2018.

CAIADO, Katia Regina Moreno; JESUS, Denise Meyrelles; BAPTISTA, Claudio Roberto. Educação especial na perspectiva da educação inclusiva em diferentes municípios. *Cadernos Cedes*, Campinas, SP, v. 38, n. 106, p. 261-265, set./dez. 2018.

DORÉ, Robert; WAGNER, Serge; BRUNET, Jean-Pierre. A integração escolar: os principais conceitos, os desafios e os fatores de sucesso no secundário. In: MANTOAN, Maria Teresa Egler (org.). *A integração de pessoas com deficiência: contribuições para uma reflexão sobre o tema*. São Paulo: Menom/SENAC, 1997. p. 174-183.

GARCIA, Rosalba Maria Cardoso. Política de educação especial na perspectiva inclusiva e a formação docente no Brasil. *Revista Brasileira de Educação*, Rio de Janeiro, v. 18, n. 52, p. 101-119, jan./mar. 2013.

GARCIA, Rosalba Maria Cardoso; MICHELS, Maria Helena. A política de educação especial no Brasil (1991-2011): uma análise da produção do GT 15 – educação especial da ANPED. *Revista Brasileira de Educação Especial*, Marília, SP, v. 17, n. especial 1, p. 105-123, maio/ago. 2011.

KASSAR, Mônica de Carvalho Magalhães. Percursos da constituição de uma política brasileira de educação especial inclusiva. *Revista Brasileira de Educação Especial*, Marília, SP, v. 17, n. especial 1, p. 41-58, maio/ago. 2011.

KASSAR, Mônica de Carvalho Magalhães; REBELO, Andressa Santos; OLIVEIRA, Regina Tereza Cestari de. Embates e disputas na política nacional de Educação Especial brasileira. *Educação e Pesquisa*, São Paulo, v. 45, p. 1-19, 2019.

MELETTI, Silvia Márcia Ferreira. Indicadores educacionais sobre a educação especial no Brasil e no Paraná. *Educação e Realidade*, Porto Alegre, v. 39, n. 3, p. 789-809, jul./set. 2014.

MELETTI, Silvia Márcia Ferreira; BUENO, José Geraldo Silveira. Escolarização de alunos com deficiência: uma análise dos indicadores sociais no Brasil (1997-2006). In: REUNIÃO ANUAL DA ANPED, 33., 2010, Caxambu, MG. *Anais [...]*. Caxambu, MG, 2010, p. 1-17.

MELETTI, Silvia Márcia Ferreira; RIBEIRO, Karen. Indicadores educacionais sobre a educação especial no Brasil. *Cadernos Cedes*, Campinas, SP, v. 34, n. 93, p. 175-189, maio/ago. 2014.

MENDES, Enicéia Gonçalves. Breve histórico da educação especial no Brasil. *Revista Educación y Pedagogía*, Medellín, v. 22, n. 57, p. 93-109, mayo/ago., 2010.

REBELO, Andressa Santos; KASSAR, Mônica de Carvalho Magalhães. Indicadores educacionais de matrículas de alunos com deficiência no Brasil (1974-2014). *Estudos em Avaliação Educacional*, São Paulo, v. 29, n. 70, p. 276-307, jan./abr. 2018.

SANTOS, Vivian; MENDES, Enicéia Gonçalves. Distorção idade-série de estudantes paulistas com e sem necessidades educacionais especiais. *Estudos em Avaliação Educacional*, São Paulo, v. 30, n. 74, p. 486-507, maio/ago. 2019.

SÃO PAULO. *Resolução SE n. 135, de 18 de julho de 1994*. Cria, no âmbito da Secretaria da Educação, o Centro de Apoio Pedagógico para Atendimento ao Deficiente Visual – CAP. São Paulo, 1994.

SÃO PAULO. *Resolução SE n. 61, de 05 de abril de 2002*. Dispõe sobre ações referentes ao Programa de Inclusão Escolar. São Paulo, 2002.

SÃO PAULO. *Resolução SE n. 32, de 23 de maio de 2007*. Dispõe sobre o desenvolvimento das ações do programa de atendimento aos alunos da rede pública com necessidades educacionais especiais. São Paulo, 2007.

SÃO PAULO. *Resolução SE n. 72, de 09 de outubro de 2009*. Estabelece orientações e procedimentos para a celebração de convênios com instituições, sem fins lucrativos, atuantes em educação especial, e dá providências correlatas. São Paulo, 2009.

SÃO PAULO. *Decreto n. 57.142, de 18 de julho de 2011*. Reestrutura a SEE/SP – Centro de Atendimento Especializado (Caesp), que passa a aglutinar o Cape e o Núcleo de Inclusão (Ninc). São Paulo, 2011.

SÃO PAULO. *Resolução SE n. 14, de 07 de fevereiro de 2012*. Dispõe sobre a celebração de convênio com entidades de fins não econômicos, para proporcionar atendimento e apoio a alunos com deficiência, matriculados em escolas da rede estadual de ensino, e dá providências correlatas. São Paulo, 2012.

SÃO PAULO. *Resolução SE n. 70, de 29 de junho de 2012*. Altera dispositivo da Resolução SE 54, de 12-08-2011, que dispõe sobre a celebração de convênios com instituições, sem fins lucrativos, atuantes em educação especial. São Paulo, 2012.

SÃO PAULO. *Resolução SE n. 81, de 07 de agosto de 2012*. Dispõe sobre o processo de aceleração de estudos para alunos com altas habilidades/superdotação na rede estadual de ensino e dá providências correlatas. São Paulo, 2012.

SÃO PAULO. *Resolução SE n. 32, de 17 de maio de 2013*. Dispõe sobre as atribuições do Núcleo de Apoio Pedagógico Especializado – CAPE, em diretorias de ensino, e dá providências correlatas. São Paulo, 2013.

SÃO PAULO. *Resolução SE n. 18, de 03 de abril de 2014*. Institui Comissão de Gestão do Termo de Ajustamento de Conduta celebrado entre a Secretaria da Educação e o Ministério Público do Estado de São Paulo, visando à acessibilidade nos prédios escolares da rede estadual de ensino. São Paulo, 2014.

SÃO PAULO. Resolução SE n. 59, de 22 de novembro de 2016. Dispõe sobre o processo de credenciamento de organizações da sociedade civil sem fins lucrativos, a que alude o art. 30, inciso VI, da Lei Federal n. 13.019, de 31/07/2014. São Paulo, 2016.

SÃO PAULO. Resolução SE s/n, de 08 de dezembro de 2016. Homologando, com fundamento no artigo 9º da Lei 10.403, de 6-7-1971, a Deliberação CEE 149/2016, que “Estabelece normas para a educação especial no sistema estadual de ensino”. São Paulo, 2016.

SÃO PAULO. Resolução SE n. 68, de 12 de dezembro de 2017. Dispõe sobre o atendimento educacional aos alunos, público-alvo da Educação Especial, na rede estadual de ensino. São Paulo, 2017.

SÃO PAULO. Resolução SE n. 15, de 16 de fevereiro de 2018. Dispõe sobre o processo de credenciamento de organizações da sociedade civil sem fins lucrativos, a que alude o art. 30, inciso VI, da Lei Federal n. 13.019, de 31/07/2014. São Paulo, 2018.

SILVA JÚNIOR, João dos Reis. A racionalidade mercantil da pós-graduação: a produção da ciência pragmática e do individualismo profissional. In: QUARTIERO, Elisa Maria; BIANCHETTI, Lucídio (org.). *Educação corporativa: mundo do trabalho e do conhecimento: aproximações*. Santa Cruz do Sul, SC: EDUNISC; São Paulo: Cortez, 2005. p. 288-312.

SILVA, Régis Henrique dos Reis. Apontamentos histórico-críticos sobre a educação especial brasileira: vicissitudes, problemas e perspectivas. In: ZANOLLA, Sílvia Rosa da Silva; BAPTISTA, Tadeu João Ribeiro (org.). *Educação, cultura, diversidade e formação*. Campinas, SP: Mercado de Letras, 2016, p. 159-190.

SILVA, Régis Henrique dos Reis; MACHADO, Robson; SILVA, Ribamar Nogueira da. Golpe de 2016 e a educação no Brasil: implicações nas políticas de educação especial na perspectiva da educação inclusiva. *Revista HISTEDBR On-line*, Campinas, SP, v. 19, p. 1-23, 2019. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/histedbr/article/view/8655755>. Acesso em: maio 2020.

NOTA: Os autores contribuíram para a elaboração do artigo da seguinte maneira: Régis Henrique dos Reis Silva preparou a atualização da literatura e, juntamente com Luciana Salvatti Coutinho, colaborou no estudo, análise e sistematização escrita das legislações nacional e estadual relacionadas ao tema do artigo. Joice Melo Vieira e Juliana Salles Madeira fizeram a preparação do banco de dados e produção de gráficos e tabelas. Todos os autores colaboraram na conclusão das análises (complementação de informações) e na elaboração e revisão final do artigo.

Recebido em: 29 MAIO 2020

Aprovado para publicação em: 8 DEZEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7381>

UMA MATRIZ DE (AUTO)AVALIAÇÃO DA INTERNACIONALIZAÇÃO DO ENSINO SUPERIOR NO BRASIL¹

 GABRIEL BRITO AMORIM^I

 KYRIA REBECA FINARDI^{II}

^I Universidade Federal do Espírito Santo (Ufes), Vitória-ES, Brasil; gabrielbamorim@icloud.com

^{II} Universidade Federal do Espírito Santo (Ufes), Vitória-ES, Brasil; kyria.finardi@gmail.com

RESUMO

O estudo apresenta uma proposta de matriz de (auto)avaliação do processo de internacionalização de instituições de ensino superior, pilotada em uma universidade federal brasileira. A matriz engloba 86 indicadores, distribuídos nos três pilares da universidade, a saber, nas dimensões Ensino, Pesquisa e Extensão, analisadas em relação às categorias Política Linguística, Mobilidade Acadêmica e Internacionalização em Casa, sugerindo três possíveis classificações: Internacionalizada, Engajada e Emergente. Os resultados da análise da matriz sugerem que ela representa uma contribuição relevante para a (auto)avaliação das instituições de ensino superior brasileiras como um instrumento unificado pertinente de (auto)avaliação da internacionalização.

PALAVRAS-CHAVE INTERNACIONALIZAÇÃO DA EDUCAÇÃO • AVALIAÇÃO DA EDUCAÇÃO • MÉTODOS DE AUTOAVALIAÇÃO.

1 Este artigo representa um recorte da pesquisa de doutorado do primeiro autor (AMORIM, 2020) sob a orientação do segundo autor, sendo que ambos contribuíram para a escrita do artigo. O segundo autor agradece o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) na forma de Bolsa de Produtividade em Pesquisa (PQ).

UNA MATRIZ DE (AUTO)EVALUACIÓN DE LA INTERNACIONALIZACIÓN DE LA EDUCACIÓN SUPERIOR EN BRASIL

RESUMEN

El estudio presenta una propuesta de matriz de (auto)evaluación del proceso de internacionalización de instituciones de educación superior, aplicada en una universidad federal brasileña. La matriz engloba 86 indicadores, distribuidos en los tres pilares de la universidad, a saber, en las dimensiones de Enseñanza, Investigación y Extensión, analizadas en relación a las categorías de Política Lingüística, Movilidad Académica e Internacionalización en Casa, sugiriendo tres posibles clasificaciones: Internacionalizada, Comprometida y Emergente. Los resultados del análisis de la matriz sugieren que ella representa una contribución relevante para la (auto)evaluación de las instituciones de educación superior brasileñas como un instrumento unificado pertinente de (auto)evaluación de la internacionalización.

PALABRAS CLAVE INTERNACIONALIZACIÓN DE LA EDUCACIÓN • EVALUACIÓN DE LA EDUCACIÓN • MÉTODOS DE AUTOEVALUACIÓN.

A (SELF)ASSESSMENT MATRIX OF THE INTERNATIONALIZATION OF HIGHER EDUCATION IN BRAZIL

ABSTRACT

The study proposes a matrix of (self)evaluation of the process of internationalization of higher education institutions piloted in a Brazilian federal university. The matrix comprises 86 indicators, distributed in the three pillars of the university, namely, Teaching, Research and Outreach dimensions analyzed in relation to Language Policy, Academic Mobility and Internationalization at Home, suggesting three possible classifications: Internationalized, Engaged and Emerging. The analysis of the matrix suggests that it represents a relevant contribution to the (self)evaluation of Brazilian higher education institutions as a unified instrument for (self)evaluation of internationalization.

KEYWORDS INTERNATIONALIZATION OF EDUCATION • EVALUATION OF EDUCATION • SELF-ASSESSMENT METHODS.

INTRODUÇÃO

A internacionalização do ensino superior, definida como o processo de integração intencional de uma dimensão internacional, intercultural ou global no Ensino, Pesquisa e Extensão do ensino superior (DE WIT, 2019; KNIGHT, 2003), pode ser entendida também como um processo “inteligente” que exige a articulação e integração de pesquisas, práticas e políticas no âmbito do ensino superior. Para que a internacionalização seja compreendida como um processo inteligente (RUMBLEY, 2015), os participantes, agentes e tomadores de decisão desse processo devem ter acesso a dados e informações relevantes, a fim de avaliá-lo, planejá-lo, transformá-lo e guiá-lo.

Na última década, no Brasil, a internacionalização tem ocupado lugar de destaque nas agendas de discussão das instituições de ensino superior (IES) brasileiras, principalmente desde a criação do programa Ciência sem Fronteiras (CsF). Outro efeito causado pelo CsF foi o impulso para a criação de outros programas, políticas e incentivos governamentais para a internacionalização, tais como o programa Inglês sem Fronteiras, em 2012, repaginado em 2014 como Idiomas sem Fronteiras (IsF), para incluir outros idiomas, e mais recentemente o Programa Institucional de Internacionalização da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (Capes PrInt) (FINARDI; ARCHANJO, 2018).

No Brasil, teóricos da área, como Lima e Maranhão (2009), por exemplo, descrevem o processo de internacionalização das IES brasileiras como um processo passivo na medida em que se concentra no envio de acadêmicos para o exterior, como ocorreu no CsF, ou seja, mais alunos brasileiros vão para o exterior, enquanto um processo de mobilidade acadêmica ativo tende a receber mais acadêmicos do exterior. O processo de internacionalização das IES brasileiras também é descrito como sendo fortemente induzido por programas e políticas nacionais, como o CsF, o IsF e o Capes PrInt, além de modelos importados do norte global, constituindo-se, por isso mesmo, um processo reativo² (FINARDI; GUIMARÃES, 2017).

Em relação à análise desse processo, teóricos da área de avaliação (RAUPP; REICHLÉ, 2003; STUFFLEBEAM, 1968, 1971) entendem que o desenvolvimento de estratégias, políticas e programas deve ser retroalimentado por processos avaliativos (tanto informativos quanto formativos), para dar conta de um planejamento eficiente, bem como da análise dos efeitos e impactos dessas políticas, estratégias e programas. Para esses teóricos, somente a avaliação pode fornecer informações relevantes para o planejamento e administração (inteligente) de recursos e ações.

2 Os conceitos de sul global e norte global são epistemológicos e se referem à situação geopolítica e não necessariamente geográfica de alguns países. Esses termos serão utilizados neste trabalho segundo a perspectiva de Boaventura de Sousa Santos (2011).

Considerando a carência de instrumentos de avaliação do processo de internacionalização das IES brasileiras, identificada em Amorim (2020), o autor propôs a criação de um instrumento de (auto)avaliação desse processo, contribuindo, assim, para a avaliação da internacionalização das IES brasileiras. O arcabouço teórico que informou a construção da matriz proposta por Amorim (2020) incluiu as ideias de Selasi (2014), em relação à importância da multidimensionalidade de um contexto; de Blommaert (2010, 2013), em relação aos conceitos de globalização/internacionalização; e de Stufflebeam (1968, 1971) e Raupp e Reichle (2003), em relação à necessidade de avaliação contextualizada para fins de diagnóstico/planejamento estratégico.

No que se refere ao uso de matrizes para a avaliação do processo de internacionalização do ensino superior no mundo, Knight (2004) propôs diferentes abordagens para avaliar a internacionalização, dentre as quais podemos destacar uma abordagem nacional que a pesquisadora divide em cinco níveis: 1) programas; 2) análise de princípios e motivações; 3) análise *ad hoc*; 4) políticas; e 5) estratégias. Um outro tipo de abordagem proposto por Knight (2004) é a institucional, que engloba seis pilares de avaliação, quais sejam: 1) atividade; 2) resultados; 3) princípios e motivações; 4) processo; 5) internacionalização em casa/local; e 6) exterior. A área afirma que as diretrizes sugeridas são apenas referências ou pontos de partida (ou chegada) para aqueles que conduzem as avaliações de seus processos de internacionalização, sejam eles governos, nações ou instituições, e também dependem das razões e motivações de quem as conduz.

Em se tratando do uso de matrizes de avaliação do processo de internacionalização no Brasil, Amorim e Finardi (2017), com base na matriz proposta por Knight (2004), propuseram uma matriz tridimensional para análise do processo de internacionalização de uma universidade federal do Sudeste, a mesma universidade onde a matriz de Amorim (2020) foi pilotada, a Universidade Federal do Espírito Santo (Ufes). A matriz de Amorim e Finardi (2017) tinha como objetivo analisar o processo de internacionalização em três níveis, a saber: no nível macro – analisando políticas e programas nacionais como o CsF e o IsF; no nível meso – analisando programas e ações institucionais; e no nível micro – analisando o engajamento da comunidade acadêmica como um todo. A fim de coletar dados para análise do nível micro (comunidade acadêmica), Amorim e Finardi (2017) enviaram um questionário para toda a comunidade acadêmica da Ufes, sendo que um total de 1.649 participantes responderam e oito participantes foram também entrevistados para triangulação dos dados. A análise levou em conta dados quantitativos e qualitativos, triangulando as percepções de participantes da comunidade acadêmica sobre as ações de internacionalização da universidade investigada, no nível micro; com as ações de internacionalização implementadas pela IES, no nível meso; e

com as políticas linguísticas e de internacionalização concretizadas em programas nacionais como o CsF e o IsF, no nível macro. De maneira geral, os resultados de Amorim e Finardi (2017) sugerem que, no nível macro, a internacionalização afeta e é afetada pela globalização, sendo que o inglês tem um papel diferenciado dos outros idiomas nesse processo. No nível meso, os resultados sugeriram que a IES analisada, apesar de almejar a internacionalização, ainda tinha um processo incipiente e reativo, sendo induzido principalmente por programas e políticas nacionais, tal como relatado anteriormente por Finardi, Santos e Guimarães (2016), que investigaram o papel dos idiomas e do programa IsF na criação da Divisão de Línguas da Secretaria de Relações Internacionais da Ufes, mesma IES investigada em Amorim e Finardi (2017).

Outro estudo que desenvolveu uma matriz de avaliação para o contexto brasileiro foi o de Finardi *et al.* (2020). Após uma metanálise de 21 estudos da área da internacionalização, os pesquisadores desenvolveram uma matriz tridimensional subdividida em três camadas: 1) internacionalização em casa; 2) princípios; e 3) políticas linguísticas, mobilidade e educação. A primeira camada, internacionalização em casa, refere-se a políticas de internacionalização da instituição que contemplem os três níveis pilares da universidade – Ensino, Pesquisa e Extensão. Já a segunda camada, princípios, está relacionada a valores globais e locais que devem perpassar o desenho das ações e políticas implantadas na instituição. Por último, a terceira camada envolve três aspectos importantes do processo de internacionalização: 1) políticas linguísticas, que leva em conta o papel das línguas estrangeiras; 2) mobilidade, que perpassa ações do tipo envio de estudantes para o exterior (mobilidade *out*) e recebimento de estudantes estrangeiros (mobilidade *in*) e também virtual; e, por fim, 3) o ensino englobando aspectos como internacionalização do currículo e formação de professores. A matriz tridimensional de Finardi *et al.* (2020) forneceu relevantes contribuições para o campo da avaliação da internacionalização, ainda que a matriz não detalhe as camadas nem elabore indicadores para uma avaliação mais pormenorizada e aprofundada desse processo tão complexo.

A fim de suprir parcialmente essa lacuna, a matriz multidimensional de Amorim (2020) se inspirou na matriz de Knight (2004), e nos modelos de Amorim e Finardi (2017) e de Finardi *et al.* (2020), indo além dessas matrizes ao propor três dimensões, três categorias e três classificações de internacionalização institucional, que serão descritas mais adiante. Além desses estudos, Amorim (2020) levou em consideração documentos norteadores da internacionalização do ensino superior no Brasil, como as Orientações para a Elaboração de Políticas Linguísticas para a Internacionalização do Ensino Superior, da Associação Brasileira de Educação

Internacional (Faubai) (2017);³ o levantamento, feito pela Faubai/Conselho Britânico, de cursos ofertados em inglês nas IES brasileiras, publicado no Guia EMI⁴ 2018-2019⁵ (FAUBAI, 2018); e os critérios de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (Capes) para programas de pós-graduação.⁶ O Quadro 1 resume as contribuições desses documentos para a avaliação da internacionalização no Brasil:

QUADRO 1 – Resumo das matrizes e documentos de avaliação da internacionalização no Brasil

MATRIZES DE AVALIAÇÃO DA INTERNACIONALIZAÇÃO NO BRASIL	CONTRIBUIÇÕES
Amorim e Finardi (2017)	Avaliou a internacionalização nos níveis macro (nacional), meso (institucional) e micro (comunidade); de modo geral a internacionalização no Brasil é incipiente.
Finardi <i>et al.</i> (2020)	Conduziu metanálise com 21 estudos da área e propôs matriz tridimensional para internacionalização com três camadas: 1) internacionalização em casa; 2) princípios; e 3) políticas linguísticas, mobilidade e educação.
Orientações para a Elaboração de Políticas Linguísticas para a Internacionalização do Ensino Superior da Faubai (FAUBAI, 2017)	Perguntas norteadoras para a elaboração de políticas linguísticas para a internacionalização das IES brasileiras.
Guia EMI 2018-2019 (FAUBAI, 2018)	Mapeamento de programas de graduação e pós-graduação de IES públicas e particulares no Brasil que oferecem cursos em língua inglesa.
Crítérios de avaliação da Capes para programas de pós-graduação	Estabelece critérios para avaliação de programas de pós-graduação no Brasil que variam de 1 a 7, sendo 6 e 7 as classificações mais altas e os ditos programas de "excelência", e cujos critérios obrigatoriamente incluem internacionalização.

Fonte: Elaboração dos autores.

Além desses estudos e documentos, Amorim (2020) se guiou pelo modelo de classificação de Foskett (2010) ao propor as três classificações do processo de internacionalização de uma IES. Foskett (2010) conduziu estudo com sete universidades no Reino Unido e 16 na Ásia, com o intuito de traçar o perfil dessas instituições, chegando a cinco classificações: 1) Universidades Domésticas, instituições que estão focadas no seu contexto nacional e regional, podendo haver recrutamento de alunos internacionais, mas investem pouco em recrutamento ativo; 2) Universidades Imperialistas, instituições que possuem estratégias de recrutamento internacional bem fortes, mas não se preocupam em fazer nenhum tipo de adaptação em

³ Disponível em: <http://faubai.org.br/pt-br/wp-content/uploads/2019/01/Documento-do-GT-de-Pol%C3%ADticas-Lingu%C3%ADsticas-da-FAUBAI.pdf>. Acesso em: 12 maio 2020.

⁴ *English Medium Instruction* (EMI).

⁵ Disponível em: <http://faubai.org.br/britishcouncilfaubaiguide2018.pdf>. Acesso em: 12 maio 2020.

⁶ Disponível em: <https://capes.gov.br/images/stories/download/avaliacao/27032017-Portaria-59-21-03-2017-Regulamento-da-Avaliacao-Quadrienal.pdf>. Acesso em: 12 maio 2020.

sua organização, espaço físico ou serviços para atender esse público internacional; 3) Universidades Internacionalmente-Conscientes, instituições que estão mudando sua organização e cultura para ter um perfil internacional, mas ainda não estão completamente engajadas com recrutamento e parcerias internacionais; 4) Universidades Internacionalmente-Engajadas, aquelas que já estão operando na arena internacional por meio de recrutamento e parcerias internacionais; e, finalmente, 5) Universidades Internacionalmente-Focadas, cujos planos de internacionalização são fortes e desenvolvidos tanto nacionalmente (internacionalização em casa) quanto no exterior.

O estudo de Foskett (2010), apesar de abranger um número significativo de instituições, foi conduzido em países que apresentam uma realidade diferente da que encontramos no Brasil – a saber, Reino Unido/norte global e Ásia –, razão pela qual o estudo serviu de base para a proposta de critérios de classificação da internacionalização institucional, tendo sido, no entanto, adaptado por Amorim (2020) para o contexto brasileiro. No que segue, o presente artigo apresenta a construção e análise da matriz de Amorim (2020) concluindo com uma discussão de implicações para o processo de avaliação e internacionalização no Brasil.

CONSTRUÇÃO DA MATRIZ DE (AUTO)AVALIAÇÃO

A metodologia do desenvolvimento (COUTINHO; CHAVES, 2001; RICHEY; KLEIN; NELSON, 2001; VAN DEN AKKER, 1999; VAN DEN AKKER *et al.*, 2006; VAN DER MAREN, 1996) tem como foco principal propor uma solução ou modelo protótipo para um problema ou uma situação real e foi usada para a construção da matriz de Amorim (2020). Van den Akker (1999) aponta que a metodologia do desenvolvimento difere das demais abordagens de investigação por causa da finalidade a que se propõe. Van den Akker (1999, p. 8-9) ainda esclarece que “a investigação com fins de desenvolvimento visa dar ao mesmo tempo contributos práticos e científicos”. Perguntas como “é possível conceber uma intervenção prática para um problema real?” orientam um processo de “aproximações sucessivas” ou de “evolução de um protótipo” rumo a uma “intervenção ideal” para um problema real (VAN DEN AKKER, 1999, p. 8-9). Segundo Bottentuit Junior (2011), as características e *design* dessa metodologia são:

- 1) Investigação preliminar: no caso da matriz de Amorim (2020), foi feito um estudo intenso da literatura sobre a internacionalização no mundo e no Brasil para identificação de (possíveis) similaridades e discrepâncias entre os diferentes contextos. Nessa fase, objetivou-se também identificar pontos nevrálgicos no contexto brasileiro que servissem de ímpeto para o processo de composição da matriz por meio de consulta à literatura da área de internacionalização e de avaliação desse processo.

- 2) Embebimento teórico: com base na revisão de literatura sobre internacionalização e avaliação realizada em 1) e, sobretudo, nas propostas de classificação de Foskett (2010) e da matriz de Knight (2004), Amorim e Finardi (2017) e Finardi *et al.* (2020), o design da matriz multidimensional de (auto)avaliação (solução/modelo protótipo) de Amorim (2020) foi proposto.
- 3) Testagem empírica: aplicação da solução/modelo protótipo ao grupo-alvo em contextos reais, ou seja, pilotagem da matriz de (auto)avaliação em uma instituição de ensino superior. A matriz de Amorim (2020) foi pilotada na Ufes em 2019, em virtude de essa IES já ter sido objeto de estudos e análises anteriores (AMORIM; FINARDI, 2017; FINARDI; ORTIZ, 2015), permitindo assim uma comparação entre estudos realizados em dois momentos distintos nessa IES.
- 4) Análise e reflexão: após a testagem empírica e focal da matriz, procedeu-se à análise do instrumento discutindo seu potencial de generalização para outros contextos além do da pilotagem.

A MATRIZ MULTIDIMENSIONAL DE (AUTO)AVALIAÇÃO

A matriz proposta por Amorim (2020) é composta por três dimensões, três categorias e três classificações, conforme se pode ver no Quadro 2:

QUADRO 2 - Dimensões, categorias e classificações da matriz de Amorim (2020)

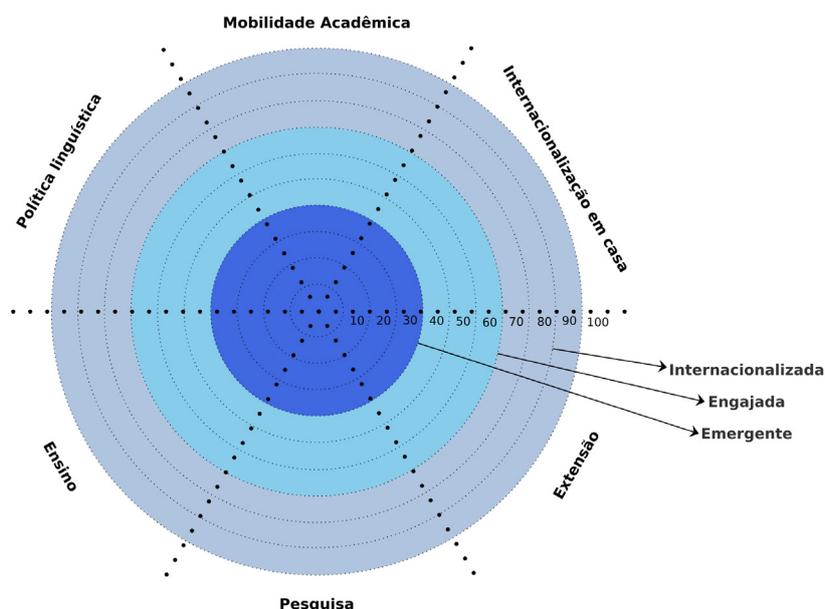
DIMENSÃO	CATEGORIA	CLASSIFICAÇÃO
Ensino	Política Linguística	Internacionalizada
Pesquisa	Mobilidade Acadêmica	Engajada
Extensão	Internacionalização em Casa	Emergente

Fonte: Adaptado de Amorim (2020).

A Figura 1 ilustra a matriz de (auto)avaliação de Amorim (2020). O círculo, similar à roda da vida, representa a ideia de processo, de constante avaliação para mudança e planejamento estratégico. As linhas pontilhadas indicam a relação de comunicação entre as dimensões e as categorias, uma vez que entendemos que dentro de uma IES essas atividades estão todas interligadas e são dificilmente dissociadas. As dimensões (Ensino, Pesquisa e Extensão) foram dispostas na parte inferior (*covert*), pois, na prática, elas são aspectos implícitos do processo e não visíveis, por isso consideramos que estão inseridas nas categorias. Já as categorias (Política Linguística, Mobilidade Acadêmica e Internacionalização em Casa) foram dispostas na parte superior (*overt*), pois representam, na prática, os aspectos explícitos do processo, ou que vemos de fato, e serão discutidas nesta seção. Da margem para o centro, temos as três classificações representadas por diferentes cores, sendo a mais exterior a classificação *Internacionalizada*; a subsequente, a *Engajada*;

e a mais interior, a classificação *Emergente*. De acordo com a proposta de (auto) avaliação, quanto mais respostas “sim” aos indicadores (mais sobre isso adiante), maior a classificação.

FIGURA 1 - A matriz multidimensional de (auto)avaliação da internacionalização do ensino superior



Fonte: Amorim (2020).

Na categoria de Política Linguística, pensou-se no papel das línguas dentro da/ para a comunidade acadêmica. Pensou-se no uso das línguas nas atividades que são desenvolvidas dentro de uma IES, a saber: nas línguas de admissão, nas línguas de instrução, nas línguas de pesquisa, nas línguas de comunicação externa e nas línguas de administração (FINARDI; CSILLAGH, 2016; FINARDI, 2017), conforme discutido em tese sobre políticas linguísticas para internacionalização (GUIMARÃES, 2020). Foram também avaliadas ações/atividades curriculares e extracurriculares tanto no âmbito da graduação quanto da pós-graduação, no que concerne ao ensino de línguas estrangeiras para a população acadêmica, incluindo Português como Língua Estrangeira (PLE). Um dos pontos levantados por Amorim e Finardi (2017) foi a questão do acesso às línguas estrangeiras (gratuito ou não) e da forma (curricular ou extracurricular), também considerado nessa categoria da matriz de Amorim (2020).

Algumas das perguntas/questões de documentos norteadores, tais como as Orientações para Elaboração de Políticas Linguísticas para Internacionalização da Faubai, o Guia EMI da Faubai (FAUBAI, 2016, 2018, respectivamente), os quesitos de avaliação da Capes para programas de pós-graduação e o Edital Capes PrInt (CAPES, 2017a) também foram usados na elaboração dos indicadores da matriz.

Na categoria Mobilidade Acadêmica, Amorim (2020) analisou os programas de mobilidade em termos de população atendida (discentes, docentes, técnicos-administrativos), áreas de conhecimento contempladas e local de origem ou de destino, bem como o resultado dessas mobilidades e o impacto na IES. Além disso, o equilíbrio entre os dois tipos de mobilidade (para o país e para o exterior) – e se há países/universidades prioritários para esses programas de mobilidade (vide Edital Capes PrInt) – também foi contemplado na construção da matriz. A categoria relativa à mobilidade propicia um levantamento do tipo de internacionalização que uma IES pretende implementar ou que já possui e se ela é mais ativa ou passiva, segundo a classificação de Lima e Maranhão (2009), ou periférica, segundo Morosini (2011), e mais focada nos programas de excelência da pós-graduação, conforme Ramos (2018).

A construção da matriz teve, ainda, a preocupação de olhar para as possibilidades de diferentes modelos de mobilidade tais como a virtual, também conhecida como COIL (do inglês, Collaborative Online International Learning) (HILDEBLANDO JÚNIOR; FINARDI, 2018). A matriz prevê interfaces na análise de categorias como, por exemplo, a análise de programas de mobilidade levando em conta questões relacionadas a políticas linguísticas.

Na categoria Internacionalização em Casa foram consideradas atividades de engajamento intercultural, assim como questões de acesso à informação no *campus* em termos de seus espaços físicos, programas, corpo docente, entre outros. Novamente podemos ver aqui uma interface entre essa categoria e a categoria Política Linguística, uma vez que os idiomas usados (ou não) para essas trocas também falam do potencial de internacionalização/interculturalidade de uma IES, conforme sugerem Guimarães e Finardi (2018). Num modelo de internacionalização ativa (LIMA; MARANHÃO, 2009), em casa (BAUMVOL, 2019; BAUMVOL; SARMENTO, 2016), mais abrangente e “local” (ABBA; STRECK, 2019; RAMOS, 2018; STRECK; ABBA, 2018), deve-se atentar para questões interculturais que permeiam as trocas entre a comunidade acadêmica e os acadêmicos estrangeiros.

INDICADORES DE AVALIAÇÃO DA INTERNACIONALIZAÇÃO

Na busca por critérios/indicadores coerentes com o contexto do ensino superior brasileiro para a construção da matriz de Amorim (2020), a forma como as IES brasileiras são avaliadas tanto na graduação quanto na pós-graduação foi considerada. Em relação à graduação, os Indicadores de Qualidade da Educação Superior e o Guia EMI da Faubai foram usados e, em relação à pós-graduação, foram usados os critérios de avaliação e editais da Capes.

Os Indicadores de Qualidade da Educação Superior, por exemplo, são compostos por quatro itens, todos administrados pelo Instituto Nacional de Estudos

e Pesquisas Educacionais “Anísio Teixeira” (Inep) e enquadrados no Sistema Nacional de Avaliação da Educação Superior (Sinaes); são eles: 1) o Índice Geral de Cursos (IGC), que é um indicador que avalia a qualidade das IES; 2) o Conceito Preliminar de Curso (CPC), que é um indicador que avalia a qualidade dos cursos superiores; 3) o Indicador de Desempenhos Observado e Esperado (IDD), que busca mensurar o valor agregado pelo curso ao desenvolvimento dos estudantes concluintes, considerando seus desempenhos no Exame Nacional de Desempenho dos Estudantes (Enade) e no Exame Nacional do Ensino Médio (Enem), como medida *proxy* (aproximação) das suas características de desenvolvimento ao ingressar no curso de graduação avaliado; e, por fim, 4) o Conceito Enade, que é um indicador que busca avaliar o desempenho dos alunos com base no resultado obtido na aplicação do exame.⁷ Novas IES ou IES que desejam abrir novos cursos devem solicitar ao Inep uma avaliação presencial feita por uma comissão designada pelo órgão, que segue um manual disponibilizado no sítio eletrônico do Inep. As IES também devem, no decorrer dos cursos, solicitar reavaliação e reconhecimento de seus cursos.

Já em relação à pós-graduação, o Regulamento para Avaliação Quadrienal de Programas de Pós-Graduação da Capes (CAPES, 2017b) estabelece princípios e diretrizes para a avaliação dos programas de mestrado e doutorado de IES públicas e privadas brasileiras que são classificatórios e comparativos. A avaliação é classificatória no sentido de que estabelece diferentes níveis de qualidade de desempenho dos programas em cada área da avaliação, e é comparativa entre as diferentes áreas. Embora preserve a especificidade de cada área, a avaliação pauta-se pelo mesmo conjunto de quesitos e itens avaliativos e deve promover a equivalência de qualidade de desempenho entre programas com as mesmas notas nas distintas áreas de avaliação. Esse mesmo regulamento também estabelece parâmetros, que são: 1) a adoção de padrões internacionais de desenvolvimento da área como referência; 2) a adequação dos referenciais de avaliação adotados, ajustando-os ao desenvolvimento do conhecimento da área e do programa; e 3) garantia da qualidade dos relatórios sobre o desempenho de cada programa, obedecendo aos requisitos de fundamentação técnica, clareza, estruturação, coerência e precisão. De acordo com esse regulamento, os programas são avaliados com notas que vão de 1 a 7, baseadas em conceitos “muito bom”, “bom”, “regular”, “fraco” e “insuficiente” atribuídos a cada item de uma ficha de avaliação; os resultados ficam disponíveis para consulta no sítio eletrônico da Capes.⁸

⁷ Disponível em: <http://portal.inep.gov.br/web/guest/indicadores-de-qualidade>. Acesso em: 17 out. 2018.

⁸ Disponível em: <http://cad.capes.gov.br/ato-administrativo-detallar?idAtoAdmElastic=240#anchor>. Acesso em: 5 maio 2020.

No levantamento de indicadores feito por Amorim (2020), a avaliação, com exceção do Edital Capes PrInt,⁹ lançado em dezembro de 2017, que se refere especificamente ao Projeto Institucional de Internacionalização das IES brasileiras selecionadas para participar e receber o recurso disponível, observou que, grosso modo, há rastros de indicadores aqui e ali tornando o trabalho de *stakeholders* mais difícil no que tange ao estabelecimento de uma política ou plano de internacionalização institucional alinhado a políticas/planos nacionais. O referido edital é destinado a IES com programas de pós-graduação que obtiveram notas superiores a 5 na última avaliação da Capes e que submeteram proposta e foram escolhidos para receber o fomento destinado à internacionalização de seus programas. Ou seja, não se trata de um programa inclusivo ou abrangente, já que não contempla, dentro de uma mesma IES, todos os seus programas de pós-graduação (e linhas de pesquisa), dimensões (graduação e extensão ficam de fora) e, claro, toda a comunidade acadêmica. Tal realidade colabora com as políticas de internacionalização fragmentadas (RAMOS, 2018), passivas (LIMA; MARANHÃO, 2009), reativas (FINARDI; ARCHANJO, 2018), periféricas (MOROSINI, 2011), engessadas (STRECK; ABBA, 2018) e colonizadas (ABBA; STRECK, 2019) que vemos nas IES brasileiras.

Tomando por base Coombe, Folse e Hubley (2007), a matriz de Amorim (2020) reúne, em um instrumento único, indicadores baseados nos documentos/instrumentos supracitados, com o intuito de aumentar/fomentar: 1) a utilidade, pois, uma vez compilados os indicadores em um só instrumento, a matriz será útil para os *stakeholders* em seus processos de mapeamento de potencialidades e fraquezas e, porventura, planejamento estratégico, já que a eles é destinada; 2) a validade do instrumento avaliativo, pois os indicadores aqui serão específicos para a internacionalização das IES brasileiras nas suas três dimensões, três categorias e três classificações; 3) a confiabilidade da matriz, pois será possível usar o instrumento para verificar a consistência de resultados (se houve mudança, por exemplo); 4) a praticidade na avaliação da internacionalização, pois a matriz pode ser respondida em forma de questionário *on-line* (Google Forms), que por sua vez compila os dados para a IES avaliada; 5) o retorno, pois, uma vez coletados os dados, é possível se iniciar um processo de reflexão e planejamento para melhora das atividades e melhor aproveitamento das potencialidades e recursos; 6) a autenticidade, pois o instrumento lida com uma situação real e gerará dados que, caso a IES deseje, poderão impactar um cenário real; 7) a transparência, pois as informações a serem coletadas – e de quem serão coletadas – ficarão claras desde o início do processo; e,

⁹ Disponível em: <http://www.capes.gov.br/cooperacao-internacional/multinacional/programa-institucional-de-internacionalizacao-capes-print>. Acesso em: 19 out. 2018.

por último, 8) a segurança, uma vez que leva em consideração os administradores do instrumento de avaliação.

As categorias e seus indicadores foram organizados de forma didática e visualmente acessível para a compilação das informações em uma determinada IES. A matriz de avaliação para internacionalização pode ser visualizada no Apêndice A.

CLASSIFICAÇÕES DA MATRIZ

Amorim (2020) adaptou a proposta de classificação de Foskett (2010) para a realidade brasileira, propondo uma classificação em três níveis. Para chegar a esses três níveis é importante lembrar que a matriz compreende um total de 86 indicadores para as três categorias, sendo que a categoria Política Linguística possui 26 ($\cong 30,2\%$), a categoria Mobilidade Acadêmica possui 20 ($\cong 23,7\%$) e a categoria Internacionalização em Casa possui 40 ($\cong 46\%$) indicadores. Todos os indicadores possuem o mesmo peso com três opções de respostas: “sim”, “não” e “não sei”.

Com base nos indicadores da matriz, Amorim (2020) propôs a seguinte classificação do processo de internacionalização institucional: 1) *Internacionalizada*, para IES que marcarem entre 70 e 86 respostas “sim”; 2) *Engajada*, para IES que marcarem entre 50 e 69 respostas “sim”; e 3) *Emergente*, para IES que marcarem menos de 50 respostas “sim”. Somente as respostas “sim” marcam 1 ponto. As demais respostas não pontuam, apenas mostram para as IES que ou elas têm uma área que precisa de atenção (em caso de uma resposta “não”) ou têm uma área que precisam conhecer mais (em caso de uma resposta “não sei”). De todas as formas, as respostas “negativas” servem de reflexão formativa para a IES avaliada.

Para definir os critérios para “pontos de corte”, Amorim (2020) tomou como base os documentos revisados na seção anterior, em especial os critérios da Capes (já mencionados e disponíveis em sítio eletrônico) para avaliação dos programas de pós-graduação. A Capes faz um apanhado para avaliar os programas em cinco quesitos, de 1 a 7, mas os programas só existem se obtiverem uma nota a partir de 3, e os programas 6 e 7 são minoria no Brasil (ditos de excelência, internacionalizados, como descrito em Ramos (2018)); por isso, Amorim (2020) descartou as classificações da Capes 1, 2, 6 e 7 e se ateve apenas às faixas 3, 4 e 5, que correspondem à maioria dos programas de pós-graduação no Brasil.

Assim, o Programa Nota 3 da Capes (menor que 60%, menos de 50 pontos) corresponderia à classificação Emergente, o Programa Nota 4 ($\cong 60\%$ a 84%, entre 50 e 69 pontos) corresponderia à classificação Engajada e, por fim, o Programa Nota 5 ($\cong 85\%$ a 100%, entre 70 e 86 pontos) corresponderia à classificação Internacionalizada. Para chegar a esses valores e porcentagens, Amorim (2020) usou uma regra de três arredondando para mais. Ainda que uma IES alcance uma

pontuação/porcentagem de 85% de respostas “sim” para os indicadores (ou seja, não atingindo 100%), é possível ser considerada internacionalizada com possibilidades de aperfeiçoamento em alguma das dimensões e/ou categorias conforme resumo na Tabela 1.

TABELA 1 – Resumo das sugestões de classificação da matriz

CLASSIFICAÇÕES	NOTA CAPES	PORCENTAGEM (≅)	PONTOS (“SIM”)
Internacionalizada	Nota 5	85% - 100%	70 - 86
Engajada	Nota 4	60% - 84%	50 - 69
Emergente	Nota 3	Menor que 60%	Menos de 50

Fonte: Amorim (2020).

Cumprе salientar que esses critérios podem não corresponder às expectativas em relação a uma avaliação formativa (objetivo principal do instrumento proposto). Por se tratar de uma matriz de caráter formativo (COOMBE; FOLSE; HUBLEY, 2007) e cujo objetivo é propiciar às IES uma reflexão sobre seus processos de internacionalização, a classificação ou a pontuação obtida tem como finalidade fornecer informações orientadas para as IES, muito mais que julgamentos ou ranqueamentos.

ANÁLISE DA MATRIZ MULTIDIMENSIONAL

A matriz proposta por Amorim (2020) foi pilotada na Ufes em 2019 com três respondentes, sendo dois servidores da Secretaria de Relações Internacionais (SRI) da Ufes, ambos também doutorandos na área de internacionalização, e a terceira respondente, uma pesquisadora em estágio pós-doutoral na mesma universidade, cujo projeto de pesquisa se relacionava com a avaliação do processo de internacionalização dessa IES. A pilotagem da matriz de Amorim (2020) ocorreu em duas etapas, sendo a primeira feita por meio da administração de um questionário enviado via Google Forms para validar os indicadores em relação às dimensões, categorias e classificações propostas na matriz, e a segunda composta pelo questionário da matriz com os indicadores de dimensões, categorias e classificações para ser respondido.

A análise empírica da matriz de Amorim (2020) realizada na Ufes sugeriu desmembrar alguns dos indicadores como, por exemplo, os indicadores 1, 3 e 5 da categoria de Política Linguística. Outro resultado importante da pilotagem da matriz foi a sugestão de evitar o uso de termos específicos da área, uma vez que eles podem não ser de conhecimento do avaliador. Esse foi o caso, por exemplo, dos termos sul/norte global nas categorias Mobilidade Acadêmica e Internacionalização em Casa. Essa mesma sugestão também foi dada em relação à definição de alguns termos como “visitante”, no indicador 2.6 da categoria Internacionalização em Casa.

Outro ponto importante mostrado pela pilotagem da matriz foi a necessidade de se pensar nos respondentes da matriz, ou seja, quem vai conduzir a coleta de dados da avaliação, uma vez que as respostas “não sei” são mais “negativas” do que as respostas “não”. Tal assertiva parte do pressuposto de que a matriz se destina a gestores responsáveis pelo planejamento e avaliação envolvidos com a internacionalização da instituição e que têm poder decisivo no planejamento estratégico das ações nos três pilares da universidade, quais sejam, Ensino, Pesquisa e Extensão.

Em relação a outros estudos sobre o processo de internacionalização na Ufes no período entre 2015 e 2020, ou seja, em um período de cinco anos, podemos ver um padrão que se confirmou em Amorim (2020). Iniciando com Finardi e Ortiz (2015), que analisaram as motivações para internacionalização, partindo para os resultados de Finardi, Santos e Guimarães (2016), em relação ao papel das línguas em geral e do inglês em particular no processo de internacionalização e na criação da SRI da Ufes, seguindo para os resultados de Amorim e Finardi (2017) em relação à análise do processo de internacionalização da Ufes em seus três níveis (macro – nacional, meso – institucional e micro – comunidade acadêmica), passando ainda pela meta-análise realizada por Finardi *et al.* (2020) e, finalmente, culminando nos resultados da pilotagem da matriz de Amorim (2020) na Ufes, vemos que há indícios de um engajamento dessa IES, que demonstra forte intenção de se internacionalizar, sendo que a Internacionalização em Casa parece ser o maior desafio nesse sentido.

De maneira geral, a análise da matriz de Amorim (2020) sugere que os três pilares principais do ensino superior – Ensino, Pesquisa e Extensão – estão contemplados nas categorias representativas Política Linguística, Mobilidade Acadêmica e Internacionalização em Casa da matriz. Ademais, a análise da matriz sugere que uma IES de posse de uma (auto)avaliação bem informada e orientada tem condições de (re)pensar as motivações, estratégias e objetivos que tem para seu processo de internacionalização.

De posse de um instrumento de avaliação do processo de internacionalização que permita fazer uma (auto)análise, talvez possamos seguir o proposto por Brandenburg e De Wit (2011), preocupando-nos mais com questões de fundo e de motivações do que com instrumentos de avaliação e estratégias de internacionalização. Dito de outra forma, uma vez que tenhamos condições de avaliar o que estamos fazendo, talvez possamos partir para questões mais complexas, a exemplo de por que estamos fazendo o que fazemos. Conforme colocado por De Wit (2019), podemos partir para perguntas como: podemos pensar em uma mudança de paradigma da internacionalização mais focada na cooperação do que na concorrência? O foco na internacionalização para o exterior predomina ainda na IES às custas do foco na internacionalização em casa? A internacionalização é um agente de mudança e responsabilidade social global e local do ensino superior?

CONSIDERAÇÕES FINAIS

Este artigo teve como objetivo principal apresentar a matriz proposta por Amorim (2020) para a (auto)avaliação do processo de internacionalização do ensino superior. Essa matriz engloba 86 indicadores distribuídos nos três pilares da universidade – Ensino, Pesquisa e Extensão –, que por sua vez se relacionam com três categorias de análise, quais sejam: Política Linguística, Mobilidade Acadêmica e Internacionalização em Casa. Com base na análise e triangulação das dimensões e categorias da matriz, três classificações são sugeridas para a (auto)avaliação das IES: Internacionalizada, Engajada e Emergente.

Espera-se que o uso da matriz de (auto)avaliação de Amorim (2020) propicie espaços de discussão e ressignificação do processo de internacionalização, a fim de que ele seja mais localmente informado/planejado para ser, em dimensão regional, mais relevante, inclusivo e abrangente. Na pilotagem feita por Amorim (2020), percebeu-se que o conceito/a categoria Internacionalização em Casa é um gargalo que precisa ser olhado com mais atenção (essa categoria recebe o maior número de indicadores da matriz e recebeu as menores pontuações na pilotagem), por possuir itens complexos como “internacionalização do currículo” e “desenvolvimento de competência intercultural” e também por ter recebido o maior número de respostas “não sei” dos respondentes.

Tendo em vista os resultados da análise de Amorim (2020), propomos que o processo de internacionalização institucional seja informado por pesquisas e avaliações locais que levem em consideração, pelo menos, as três dimensões e categorias propostas por Amorim (2020), podendo compreender muitas outras. Em última análise, essa proposta tem como fim (res)significar o processo de internacionalização institucional, promovendo uma internacionalização ativa e informada, sensível – tanto global quanto localmente – a contextos e realidades multidimensionais.

REFERÊNCIAS

ABBA, Maria Julieta; STRECK, Danilo R. Interculturality and internationalization. *SFU Educational Review*, v. 12, n. 3, p. 110-126, 2019.

AMORIM, Gabriel Brito. *A internacionalização do ensino superior no Brasil: uma proposta de matriz multidimensional de (auto)avaliação*. 2020. 145 f. Tese (Doutorado em Linguística) – Universidade Federal do Espírito Santo, Vitória, 2020.

AMORIM, Gabriel Brito; FINARDI, Kyria Rebeca. Globalização e internacionalização do ensino superior: evidências de um estudo de caso nos níveis micro, meso e macro. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas; Sorocaba, v. 3, n. 2, p. 614-632, set./dez. 2017.

BAUMVOL, Laura Knijnik. Advancing internationalization at home from different roles. *SFU Educational Review*, v. 12, n. 3, p. 48-53, 2019.

BAUMVOL, Laura Knijnik; SARMENTO, Simone. A internacionalização em casa e o uso de inglês como meio de instrução. In: BECK, Magali Sperling; MARTINS, Maria Lúcia Milléo; HEBERLE, Viviane (org.). *Echoes: further reflections on language and literature*. Florianópolis: EdUFSC, 2016. p. 65-82.

BLOMMAERT, Jan. *The sociolinguistics of globalization: Cambridge approaches to language contact*. Nova York: Cambridge University Press, 2010.

BLOMMAERT, Jan. Citizenship, language, and superdiversity: towards complexity. *Journal of Language, Identity & Education*, v. 12, n. 3, p. 193-196, 2013.

BOTTENTUIT JUNIOR, João Batista. *Concepção, avaliação e dinamização de um portal educacional de WebQuests em língua portuguesa*. 2011. Tese (Doutorado em Ciências da Educação) – Universidade do Minho, Braga, Portugal, 2011.

BRANDENBURG, Uwe; DE WIT, Hans. The end of internationalisation. *International Higher Education*, n. 62, p. 15-17, 2011.

CAPES. *Programa Institucional de Internacionalização – Capes-PrInt Edital 41/2017*. Brasília: Ministério da Educação, 2017a. Disponível em: <https://www1.capes.gov.br/images/stories/download/editais/10112017-Edital-41-2017-Internacionalizacao-PrInt-2.pdf>. Acesso em: dez. 2020.

CAPES. *Regulamento Quadrienal de Programas de Pós-Graduação da Capes, Portaria n. 59 de 21 de Março de 2017*. Brasília: Ministério da Educação, 2017b. Disponível em: <https://uab.capes.gov.br/images/stories/download/avaliacao/27032017-Portaria-59-21-03-2017-Regulamento-da-Avaliacao-Quadrienal.pdf>. Acesso em: dez. 2020.

COOMBE, Christine A.; FOLSE, Keith S.; HUBLEY, Nancy J. *A practical guide to assessing English language learners*. Ann Arbor, MI: University of Michigan Press, 2007.

COUTINHO, Clara P.; CHAVES, José Henrique. Investigação em tecnologia educativa na Universidade do Minho: uma abordagem temática e metodológica às dissertações de mestrado concluídas nos cursos de mestrado em educação. In: ESTRELA, Albano; FERREIRA, Júlia (org.). *Tecnologias em educação: estudos e investigações: actas do Colóquio AFIRSE/AIPELF*, 10., Lisboa, Portugal, 2001. p. 289-302.

DE WIT, Hans. Internationalization in higher education, a critical review. *SFU Educational Review*, v. 12, n. 3, p. 9-17, 2019.

FAUBAI. *Guide to English as medium of instruction in Brazilian higher education institutions*, 2016. Disponível em: https://www.britishcouncil.org.br/sites/default/files/guide_brazilian_highered_courses_inenglish.pdf. Acesso em: dez. 2020.

FAUBAI. *Orientações para a elaboração de políticas linguísticas para a internacionalização do ensino superior*, 2017. Disponível em: <http://faubai.org.br/pt-br/wp-content/uploads/2019/01/Documento-do-GT-de-Pol%C3%ADticas-Lingu%C3%ADsticas-da-FAUBAI.pdf>. Acesso em: 12 maio 2020.

FAUBAI. *Guide to English as medium of instruction in brazilian higher education institutions 2018-2019*, 2018. Disponível em: <http://faubai.org.br/britishcouncilfaubaiguide2018.pdf>. Acesso em: 12 maio 2020.

FINARDI, Kyria Rebeca. What Brazil can learn from Multilingual Switzerland and its use of English as a multilingua franca. *Acta Scientiarum. Language and Culture*, Maringá, PR, v. 39, n. 2, p. 219-228, abr./jun. 2017.

FINARDI, Kyria Rebeca; ARCHANJO, Renata. Washback effects of the Science without Borders, English without Borders and Language without Borders programs in Brazilian language policies and rights. In: SIINER, Maarja; HULT, Francis M.; KUPISCH, Tanja (ed.). *Language policy and language acquisition planning*. Cham: Springer, 2018. p. 173-185.

FINARDI, Kyria Rebeca; CSILLAGH, Virag. Globalization and linguistic diversity in Switzerland: insights from the roles of national languages and English as a foreign language. In: GRUCZA, Sambor; OLPISKA, Magdalena; ROMANOWSKI, Piotr (org.). *Advances in understanding multilingualism*. Frankfurt am Main; Warsaw: Peter Lang GmbH; International Academic Publishers, 2016, v. 24. p. 41-56.

FINARDI, Kyria Rebeca; GUIMARÃES, Felipe. Internacionalização, rankings e publicações em inglês: a situação do Brasil na atualidade. *Estudos em Avaliação Educacional*, São Paulo, v. 28, n. 68, p. 600-626, maio/ago. 2017.

FINARDI, Kyria Rebeca; GUIMARÃES, Felipe; AMORIM, Gabriel; PICCIN, Gabriela. Internationalization and language policies in Brazil: quality and assessment evidence from a glocal proposal. In: CALVO, Luciana C. S.; EL KADRI, Michele S.; PASSONI, Taisa (org.). *Language policies and internationalization of higher education institutions in Brazil: Contributions from applied Linguists*. 1. ed. Campinas, SP: Pontes, 2020.

FINARDI, Kyria Rebeca; ORTIZ, Ramon A. Globalization, internationalization and education: What is the connection? *IJAEDU: International e-Journal of Advances in Education*, v. 1, n. 1, p. 18-25, 2015.

FINARDI, Kyria Rebeca; SANTOS, Jane M.; GUIMARÃES, Felipe. A relação entre línguas estrangeiras e o processo de internacionalização: evidências da coordenação de letramento internacional de uma universidade federal. *Interfaces Brasil/Canadá: Revista Brasileira de Estudos Canadenses*, São Paulo, v. 16, n. 1, p. 233-255, 2016.

FOSKETT, Nick. Global markets, national challenges, local strategies: the strategic challenge of internationalization. In: MARINGE, Felix; FOSKETT, Nick. *Globalization and internationalization in higher education: Theoretical, strategic and management perspectives*. Londres: Continuum International Publishing Group, 2010. p. 36-50.

GUIMARÃES, Felipe Furtado. *Internacionalização e multilinguismo: uma proposta de política linguística para universidades federais*. 2020. 256 f. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal do Espírito Santo, Vitória, 2020.

GUIMARÃES, Felipe F.; FINARDI, Kyria R. Interculturalidade, internacionalização e intercompreensão: qual a relação? *Ilha do Desterro*, Florianópolis, v. 71, n. 3, p. 15-37, 2018.

HILDEBLANDO JÚNIOR, Carlos A.; FINARDI, Kyria Rebeca. Internationalization and virtual collaboration: Insights from COIL experiences. *Ensino em Foco*, Salvador, v. 1, n. 2, p. 19-33, 2018.

KNIGHT, Jane. Updating the definition of internationalization. *International Higher Education: The Boston College Center for International Higher Education*, Boston, n. 33, p. 2-3, 2003.

KNIGHT, Jane. Internationalization remodeled: Definition, approaches, and rationales. *Journal of Studies in International Education*, v. 8, n. 1, p. 5-31, 2004.

LIMA, Manolita C.; MARANHÃO, Carolina M. S. de A. O sistema de educação superior mundial: entre a internacionalização ativa e passiva. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas; Sorocaba, v. 14, n. 3, p. 583-610, 2009.

MOROSINI, Marília Costa. Internacionalização na produção de conhecimento em IES brasileiras: cooperação internacional tradicional e cooperação internacional horizontal. *Educação em Revista*, Belo Horizonte, v. 27, n. 1, p. 93-112, 2011.

RAMOS, Milena Yumi. Internacionalização da pós-graduação no Brasil: lógica e mecanismos. *Educação e Pesquisa*, São Paulo, v. 44, p. 1-22, 2018.

RAUPP, Magdala; REICHLE, Adriana. *Avaliação: ferramenta para melhores projetos*. Santa Cruz do Sul, RS: Edunisc, 2003.

RICHEY, Rita T.; KLEIN, James D.; NELSON, Wayne A. developmental research studies of instructional design and development. In: JONASSEN, David H. (ed.). *Handbook of research for educational communications and technology*. New Jersey: Taylor & Francis, 2001. p. 1099-1130. Disponível em: <http://www.aect.org/edtech/ed1/41.pdf>. Acesso em: jul. 2016.

RUMBLEY, Laura E. "Intelligent internationalization": A 21st century imperative. *International Higher Education*, Chestnut Hill, MA, n. 80, p. 16-17, 2015.

SANTOS, Boaventura S. *A universidade do século XXI: para uma reforma democrática e emancipatória*. 3. ed. São Paulo: Cortez, 2011.

SELASI, Taiye. Don't ask me where I'm from, ask me where I'm a local. *TED Global*, Oct. 2014. Disponível em: https://www.ted.com/talks/taiye_selasi_don_t_ask_where_i_m_from_ask_where_i_m_a_local. Acesso em: dez. 2020.

STRECK, Danilo R.; ABBA, Maria Julieta. Internacionalização da educação superior e herança colonial na América Latina. In: KORSUNSKY, Lionel; DEL VALLE, Damián; MIRANDA, Estela; SUASNÁBAR, Claudio (comp.). *Internacionalización y producción de conocimiento: el aporte de las redes académicas*. Ciudad Autónoma de Buenos Aires: IEC – CONADU; CLACSO, 2018. p. 113-131.

STUFFLEBEAM, Daniel L. *Evaluation as enlightenment for decision-making*. Columbus, OH: Ohio State University, 1968.

STUFFLEBEAM, Daniel L. *The relevance of the CIPP evaluation model for educational accountability*. Columbus, OH: Ohio State University, 1971.

VAN DEN AKKER, Jan. Principles and methods of development research. In: VAN DEN AKKER, Jan; BRANCH, Robert; GUSTAFSON, Kent; NIEVEEN, Nienke; PLOMP, Tjeerd (ed.). *Design approaches and tools in education and training*. Dordrecht, Netherlands: Kluwer Academic Publisher, 1999. p. 1-14.

VAN DEN AKKER, Jan; GRAVEMEIJER, Koeno; MCKENNY, Susan; NIEVEEN, Nienke. Introducing educational design research. In: VAN DEN AKKER, Jan; GRAVERMEIJER, Koeno; MCKENNY, Susan; NIEVEEN, Nienke. *Educational design research*. New York: Routledge, 2006. p. 1-8.

VAN DER MAREN, Jean-Marie. *Méthodes de recherche pour l'éducation*. 2. ed. Bruxelles: DeBoeck Université, 1996. (Collection Méthodes en Sciences Humaines).

APÊNDICE

APÊNDICE A - Indicadores da Matriz por Categorias

CATEGORIA: POLÍTICA LINGUÍSTICA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
1	A IES possui uma política linguística institucional aprovada que contempla os três pilares da universidade (Ensino, Pesquisa e Extensão).			
2	O papel das línguas estrangeiras nos três pilares da universidade (ensino, pesquisa e extensão) é bem definido no plano de internacionalização da IES.			
3	A IES adota diferentes línguas para diferentes finalidades (admissão, ensino, instrução, pesquisa e comunicação) e possui diretrizes específicas para cada dimensão de uso das línguas (admissão, ensino, instrução, pesquisa e comunicação) em seu documento de política linguística institucional.			
4	A IES, por meio de sua política linguística institucional, promove a pluralidade das línguas nos campi da IES.			
5	O(s) teste(s) de proficiência, bem como os scores exigidos para admissão tanto na graduação quanto na pós-graduação da IES, está/ estão descrito(s) na política linguística da IES e é/são chancelado(s) pela comunidade acadêmica internacional, dialoga(m) com outros exames internacionalmente conhecidos e tem/têm pontuação padronizada em toda a IES.			
6	A IES possui curso(s) de Letras.			
6.1	A IES possui curso(s) de Letras com licenciatura dupla.			
6.2	A IES possui um curso de Letras com licenciatura única.			
6.3	O departamento de Letras da IES oferta disciplinas eletivas e optativas para outros cursos em línguas estrangeiras.			
7	A IES oferta cursos de línguas estrangeiras a seu quadro de funcionários via Departamento de Desenvolvimento Pessoal (DDP).			
7.1	A IES incentiva seus docentes e técnicos-administrativos (por meio de cursos de capacitação e/ou contagem de pontos para progressão, etc.) a participarem de cursos de línguas estrangeiras no Núcleo de Línguas - Idiomas sem Fronteiras (NuLi-IsF).			
7.2	A IES incentiva seus docentes e técnicos-administrativos (por meio de cursos de capacitação e/ou contagem de pontos para progressão, etc.) a participarem de cursos de línguas estrangeiras no Núcleo/Centro/Curso de Línguas.			
7.3	A IES oferta ações de capacitação de seus gestores e servidores para lidar com documentos estrangeiros em diferentes línguas estrangeiras.			
8	A IES possui um Núcleo/Centro/Curso de Línguas Extracurricular para a comunidade acadêmica.			
8.1	Há oferta de diferentes cursos de línguas estrangeiras no Núcleo de Línguas da IES para a comunidade acadêmica e externa.			
8.2	Há distribuição de bolsas tanto para a comunidade acadêmica quanto para a comunidade externa para estudar línguas estrangeiras no Núcleo/Centro/Curso de Línguas da IES.			
9	A IES é/já foi credenciada ao Programa Idiomas sem Fronteiras (IsF).			
9.1	A IES possui um NuLi-IsF.			
9.2	A IES oferta cursos presenciais e <i>on-line</i> via IsF em línguas estrangeiras para sua comunidade acadêmica.			
9.3	O NuLi-IsF e o Núcleo de Línguas oferecem cursos de línguas de forma integrada.			

(Continua)

(Continuação)

CATEGORIA: POLÍTICA LINGUÍSTICA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
10	A IES incentiva e prepara seus docentes para ofertar disciplinas específicas em nível de graduação e pós-graduação de variadas áreas em línguas estrangeiras periodicamente por meio de cursos de capacitação, progressão funcional, etc.			
11	Há oferta de ações de extensão e pesquisa em línguas estrangeiras periodicamente para toda a comunidade acadêmica.			
12	A IES incentiva seus discentes a participarem de disciplinas específicas em línguas estrangeiras na graduação por meio de horas complementares, disciplinas eletivas/optativas, etc.			
13	A IES incentiva seus discentes a participarem de disciplinas específicas em línguas estrangeiras na pós-graduação por meio de horas complementares, disciplinas eletivas/optativas, etc.			
14	A IES possui um setor responsável por receber e encaminhar documentos em língua estrangeira.			
15	A IES possui servidores capacitados ou serviços para traduzir documentos oficiais em línguas estrangeiras.			
CATEGORIA: MOBILIDADE ACADÊMICA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
1	A IES possui uma coordenação de mobilidade <i>in</i> e <i>out</i> centralizada e capacitada para orientar o fluxo de entrada, permanência e saída de acadêmicos na IES.			
2	A IES possui programa(s) de mobilidade com e sem bolsas para discentes da graduação.			
3	A IES possui programa(s) de mobilidade com e sem bolsas para discentes da pós-graduação.			
4	A IES possui programa(s) de mobilidade com e sem bolsas ou outras formas de auxílio para docentes.			
5	A IES possui programa(s) de mobilidade com e sem bolsas ou outras formas de auxílio para servidores técnico-administrativos.			
6	Há equilíbrio entre programas e número de acadêmicos em mobilidade do tipo <i>in</i> e mobilidade do tipo <i>out</i> na IES.			
7	A IES tem parcerias com IES do “sul global” em seus programas de mobilidade.			
8	A IES tem parcerias com IES do “norte global” em seus programas de mobilidade.			
9	A IES oferta disciplina de Português como Língua Estrangeira (PLE) para seus estrangeiros na matriz curricular e como ações de extensão e pesquisa.			
10	A IES incentiva colaboração acadêmica (virtual ou não) para docentes, discentes e servidores com outras universidades do mundo.			
11	A IES possui um plano de atração de discentes, docentes, pesquisadores e servidores estrangeiros em várias áreas.			
11.1	A IES possui docentes estrangeiros atuando na graduação.			
11.2	A IES possui docentes estrangeiros atuando na pós-graduação.			
11.3	A IES possui discentes estrangeiros do “sul global” na graduação.			
11.4	A IES possui discentes estrangeiros do “sul global” na pós-graduação.			
11.5	A IES possui discentes estrangeiros do “norte global” na graduação.			
11.6	A IES possui discentes estrangeiros do “norte global” na pós-graduação.			
11.7	A IES possui servidores estrangeiros do quadro técnico-administrativo advindos do “sul global” nos seus vários setores periodicamente.			

(Continua)

CATEGORIA: MOBILIDADE ACADÊMICA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
11.8	A IES possui servidores estrangeiros do quadro técnico-administrativo advindos do “norte global” nos seus vários setores periodicamente.			
12	A IES capacita sua comunidade e gestores para a mobilidade (tipos <i>in</i> e <i>out</i>) por meio de cursos de capacitação, oficinas de aprendizado, etc.			
CATEGORIA: INTERNACIONALIZAÇÃO EM CASA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
1	A IES possui um plano de internacionalização documentado/aprovado por conselho superior.			
1.1	O plano de internacionalização da IES faz referência aos três pilares da universidade: ensino, pesquisa e extensão.			
1.2	O plano de internacionalização da IES faz referência a ações com IES do “sul global”.			
1.3	O plano de internacionalização da IES faz referência a ações com IES do “norte global”.			
1.4	O plano de internacionalização da IES faz referência a ações que englobem a realidade local da IES.			
1.5	A IES possui diferentes motivações e essas estão bem delineadas em seu plano de internacionalização, para se internacionalizar.			
1.6	As ações de internacionalização da IES são direcionadas a docentes, discentes (graduação e pós-graduação) e servidores de diferentes níveis.			
1.7	O plano de internacionalização da IES faz referência a gestores de todos os níveis e pilares da universidade (ensino, pesquisa e extensão).			
1.8	A IES possui um setor de relações internacionais.			
1.9	O plano de internacionalização da IES faz referência a ações de mobilidade contemplando os três pilares da universidade (ensino, pesquisa e extensão) e os dois níveis de ensino (graduação e pós-graduação).			
1.10	O plano de internacionalização da IES faz referência a ações de internacionalização em casa contemplando os três pilares da universidade (ensino, pesquisa e extensão).			
2	A IES possui um programa de acolhimento abrangente para seus visitantes estrangeiros.			
2.1	A IES tem sinalização e avisos importantes em línguas estrangeiras para acolher/informar visitantes estrangeiros.			
2.2	A IES desenvolve atividades para integrar estrangeiros no <i>campus/</i> nos <i>campi</i> .			
2.3	A IES disponibiliza laboratórios, bibliotecas (físicas e digitais) e outros recursos a acadêmicos estrangeiros.			
2.4	A IES disponibiliza atendimento psicológico gratuito para seus visitantes estrangeiros.			
2.5	A IES disponibiliza atendimento pedagógico gratuito para seus visitantes estrangeiros.			
2.6	A IES disponibiliza auxílio-saúde gratuito para seus visitantes estrangeiros.			
2.7	A IES disponibiliza um programa de alimentação gratuito ou restaurante universitário ou de baixo custo para seus visitantes estrangeiros.			
2.8	A IES possui estrutura ou programa de auxílio de hospedagem para seus visitantes estrangeiros.			

(Continua)

(Continuação)

CATEGORIA: INTERNACIONALIZAÇÃO EM CASA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
2.9	A IES possui estrutura ou programa de auxílio de transporte de seus visitantes estrangeiros para o(s)/do(s) campus/campi.			
2.10	A IES possui um manual/guia de informações para o visitante estrangeiro.			
3	A relação global-local é levada em consideração na formulação dos currículos da graduação da IES.			
3.1	Os currículos da graduação são pensados/formulados de maneira a suscitar a colaboração (virtual ou não) intercultural entre universidades do “sul global”.			
3.2	Os currículos da graduação são pensados/formulados de maneira a suscitar a colaboração (virtual ou não) intercultural entre universidades do “norte global”.			
3.3	A IES incentiva seus discentes de graduação a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “sul global” por meio de auxílio para tradução, fomento, etc.			
3.4	A IES incentiva seus discentes de graduação a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “norte global” por meio de auxílio para tradução, fomento, etc.			
4	A relação global-local é levada em consideração na formulação dos currículos da pós-graduação da IES.			
4.1	Os currículos da pós-graduação são pensados/formulados de maneira a suscitar a colaboração (virtual ou não) intercultural entre universidades do “sul global”.			
4.2	Os currículos da pós-graduação são pensados/formulados de maneira a suscitar a colaboração (virtual ou não) intercultural entre universidades do “norte global”.			
4.3	A IES incentiva seus discentes de pós-graduação a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “sul global” por meio de auxílio para tradução, fomento, etc.			
4.4	A IES incentiva seus discentes de pós-graduação a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “norte global” por meio de auxílio para tradução, fomento, etc.			
5	A IES oferece atividades de capacitação a seu quadro de docentes e servidores para interações com acadêmicos estrangeiros.			
5.1	A IES incentiva seus docentes a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “sul global”.			
5.2	A IES incentiva seus docentes a formarem parcerias para publicação internacional em línguas estrangeiras com acadêmicos de universidades do “norte global”.			
5.3	A IES incentiva seu quadro funcional técnico-administrativo a formarem parcerias para publicação internacional em línguas estrangeiras com profissionais de universidades do “sul global”.			
5.4	A IES incentiva seu <i>staff</i> técnico-administrativo a formar parcerias para publicação internacional em línguas estrangeiras com profissionais de universidades do “norte global”.			
6	A IES possui uma política de incentivo a revalidação/reconhecimento de créditos acadêmicos obtidos no exterior para incentivar seus docentes, discentes e quadro de funcionários técnico-administrativos a buscarem especialização acadêmica ou outras ações de pesquisa, ensino e extensão no exterior.			

(Continua)

(Continuação)

CATEGORIA: INTERNACIONALIZAÇÃO EM CASA				
Nº	INDICADORES	SIM	NÃO	NÃO SEI
7	A cultura local é pensada e integrada às atividades de imersão cultural estrangeira no campus/nos campi.			
8	A competência intercultural é levada em consideração nas interações com acadêmicos estrangeiros.			

Fonte: Amorim (2020).

Recebido em: 14 MAIO 2020

Aprovado para publicação em: 6 OUTUBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7096>

CONTORNOS DO ESTADO AVALIADOR NO BRASIL

 MARILDA PASQUAL SCHNEIDER^I

 ELINA RENILDE DE OLIVEIRA RIBEIRO^{II}

^I Universidade do Oeste de Santa Catarina (Unoesc), Joaçaba-SC, Brasil; marilda.schneider@unoesc.edu.br

^{II} Universidade do Vale do Rio dos Sinos (Unisinos), São Leopoldo-RS, Brasil; elinalameira@hotmail.com

RESUMO

Este texto tem por objetivo discutir alguns aspectos relacionados às conformações do Estado Avaliador no Brasil pela análise da evolução histórica do Sistema de Avaliação da Educação Básica (Saeb). No plano teórico-conceitual, evidencia relações entre as reconfigurações globais do papel do Estado, os modos de regulação da administração pública que emergem em consonância a essas reconfigurações, o conceito de governança e a constituição de um determinado modelo de Estado Avaliador. O estudo assume uma perspectiva explanatória fundamentada na análise da política educacional e de seus mecanismos fundamentais, para o que se vale de levantamento e análise de portarias e de um decreto promulgados pelos órgãos de regulação do Saeb, no período compreendido entre 1994 e 2020.

PALAVRAS-CHAVE ESTADO AVALIADOR • SISTEMA DE AVALIAÇÃO • EDUCAÇÃO BÁSICA • ACCOUNTABILITY.

CONTORNOS DEL ESTADO EVALUADOR EN BRASIL

RESUMEN

El estudio tiene como objetivo discutir algunos aspectos relacionados con las configuraciones del Estado Evaluador en Brasil, mediante el análisis de la evolución histórica del Sistema de Evaluación de la Educación Básica (Saeb). A nivel teórico-conceptual, destaca la correlación entre la remodelación global del rol del Estado, los medios de regulación de la gestión pública que surgen en consecuencia de esa remodelación, el concepto de gobernanza y la constitución de un modelo particular de Estado Evaluador. El estudio toma una perspectiva explicativa basada en el análisis de la política educativa y de sus mecanismos fundamentales, por medio de encuesta y análisis de ordenanzas y de un decreto promulgado por los órganos reguladores del Saeb entre 1994 y 2020.

PALABRAS CLAVE ESTADO EVALUADOR • SISTEMA DE EVALUACIÓN • EDUCACIÓN BÁSICA • ACCOUNTABILITY.

OUTLINE OF THE EVALUATIVE STATE IN BRAZIL

ABSTRACT

The present study aims to discuss some aspects of the Evaluative State structure in Brazil, by analyzing the historical evolution of the Sistema de Avaliação da Educação Básica [Basic Education Assessment System] (Saeb). At the theoretical-conceptual level, it shows the correlation between the global reshaping of the role of the state, the means of regulation of public administration that emerge accordingly, the concept of governance and the establishment of a particular model of the Evaluative State. The study takes on an explanatory perspective based on the analysis of educational policy and its leading mechanisms, through the survey and analysis of ordinances and a decree promulgated by the regulatory bodies of the Saeb between 1994 and 2020.

KEYWORDS EVALUATIVE STATE • ASSESSMENT SYSTEM • BASIC EDUCATION • ACCOUNTABILITY.

INTRODUÇÃO

O debate sobre o Estado Avaliador, em suas múltiplas formas e dimensões, constitui uma evidência de como os mecanismos de regulação operados hodiernamente e à distância podem produzir novas e mais sofisticadas relações de poder e controle, pelo Estado, de programas e ações de que decorrem as políticas de avaliação educacional. Para Barroso (2011, p. 108), os mecanismos de regulação em curso, em vários países, produzem, no seu conjunto, “um modo de governo; um padrão para a investigação; um referencial para o trabalho docente; um programa de melhoria; um dispositivo de avaliação; um sistema de prestação de contas”.

Conforme demonstram estudos sobre o tema (AFONSO, 2005, 2009; BALL, 1994, 2001; BARROSO, 2005, 2011), os modelos de regulação da educação utilizados, em período recente e em vários países, integram tecnologias complexas às funções do Estado. Essas tecnologias não substituem modelos centralizados e hierárquicos de administração das ações públicas, predominantes até o início dos anos 1980 em vários países da Europa Ocidental e nos Estados Unidos, mas atuam no fortalecimento de referenciais ajustados às leis de mercado para as quais a implantação de políticas de avaliação educacional constitui “um dos eixos estruturantes” (SCHNEIDER; ROSTIROLA, 2015, p. 497) do modelo em curso. Não se trata, pois, de afastamento do Estado na regulação e administração dos serviços, mas do redirecionamento de seu foco para um modelo de Estado Avaliador (YANNOULAS; SOUZA; ASSIS, 2009).

A repercussão do modelo gerencial de administração pública gerou mudanças profundas nas formas de conceber e operar a regulação e o controle sobre os processos educativos. Em sintonia com uma tendência que se espalhava globalmente, o apelo discursivo à necessidade de melhoria da qualidade educacional, utilizado por diversos aparelhos privados e também do Estado, serviu de mote para a implementação de políticas públicas cujo marco referencial passou a ser a avaliação periódica das instituições públicas e privadas de ensino.

A nosso ver, a configuração da política nacional de avaliação constitui estratégia importante de governança, posto que, na lógica gerencial, a atuação do Estado sobre a educação passa a ocorrer prioritariamente com base em resultados de exames e testes aplicados periodicamente aos estudantes. Especificamente sobre a educação básica, o monitoramento dos resultados por escola e sistema de ensino tem favorecido a desconcentração de responsabilidades em relação à administração dos processos educativos, o descompromisso do Estado com a oferta educacional e a expansão de um importante mercado ocupado por grandes conglomerados, redes e instituições privadas de ensino. Em vista desse processo, acende-se um importante sinal de alerta sobre as condições de oferta de uma educação pública,

gratuita, laica e de qualidade, bandeira de muitos educadores desde o consagrado *Manifesto dos Pioneiros da Educação Nova*.¹

Considerando o exposto, o presente trabalho discute alguns aspectos relacionados aos contornos do Estado Avaliador no Brasil analisando a evolução histórica do Sistema de Avaliação da Educação Básica (Saeb).² Para isso, revisita a trajetória histórica do Saeb, destacando alterações legais produzidas ao longo dos quase trinta anos de sua existência. No plano teórico-conceitual, evidenciam-se relações entre as reconfigurações globais do papel do Estado, os modos de regulação da administração pública que emergem em consonância a essas reconfigurações, o conceito de governança e a constituição de um determinado modelo de Estado Avaliador. O estudo assume uma perspectiva explanatória fundamentada na análise da política educacional e de seus mecanismos fundamentais (MAINARDES, 2018), para o que se vale de levantamento e análise de portarias e de um decreto promulgados pelos órgãos de regulação do Saeb no período de 1994 – ano de institucionalização formal do Saeb – a 2020 – ano em que foram instituídas normas complementares ao cumprimento do Decreto n. 9.432 (BRASIL, 2018), que regulamentou a política nacional de avaliação da educação básica.

Tendo em conta os procedimentos metodológicos referidos, o texto toma como ponto de partida a discussão em torno das reconfigurações do Estado, no plano globalizatório, suas características enquanto Estado Avaliador e os modos de governança operados na regulação dos serviços públicos, mormente do setor educacional. No plano prático, recupera a trajetória do Saeb, enquanto sistema nacional de avaliação da educação básica, buscando ampliar o espectro de análise em direção aos contornos do Estado Avaliador, a partir dos dois últimos anos (2018 e 2020), com a regulamentação da Política Nacional de Avaliação da Educação Básica.

RECONFIGURAÇÕES DO ESTADO E AS ESTRATÉGIAS DE GOVERNANÇA EDUCACIONAL DO ESTADO AVALIADOR

Consoante Yannoulas, Souza e Assis (2009, p. 65), “o processo de transformação do Estado em Avaliador traduz-se como um momento histórico fundamental à

1 Amplamente difundido pela literatura da área, o *Manifesto dos Pioneiros da Educação Nova* tornou-se marco inaugural do projeto de renovação educacional do país. Redigido por Fernando de Azevedo, em 1932, o documento consolidou a visão educacional de um segmento importante da elite intelectual brasileira. A principal bandeira de luta dos signatários do documento era por uma escola única, pública, laica, obrigatória e gratuita.

2 É importante sublinhar que, apesar de o Saeb ser referência nacional em avaliação, a expressão nacional nunca foi incluída nas legislações pertinentes ao tema. O pacto nacional firmado entre a União, governos municipais e estaduais, a partir da publicação do Decreto n. 6.094 (BRASIL, 2007), que instituiu o Plano de Metas Compromisso Todos pela Educação, tornou obrigatório o cumprimento das metas pactuadas pelos agentes públicos de todas as esferas federativas. Somente então o Saeb tornou-se uma política nacional de cumprimento obrigatório para todo território nacional. Considerando a forma como a denominação é tratada nos documentos regulatórios, neste texto optamos por manter sua designação oficial.

compreensão da configuração atual das políticas educacionais no que concerne tanto à sua formulação quanto à sua implementação”. O momento histórico a que se referem os autores está relacionado ao impacto do modelo gerencial implantado na administração pública a partir de meados dos anos 1970 em virtude do esfacelamento do tipo de Estado predominante até então.

Sobre esse aspecto, Abrucio (1997, p. 6) destaca que a crise econômica, social e administrativa que se instalou especialmente nos países capitalistas, na década de 1970, gerou um “amplo consenso social a respeito do papel do Estado”. A redução dos gastos públicos na área social e a redefinição das funções e modos de atuação do Estado foram, segundo o autor, saídas “introduzidas, em larga escala”, para responder à crise econômica e social gerada pelo modelo burocrático weberiano de administração pública.

Posicionando a discussão no âmbito teórico-conceitual, Antunes (2005, p. 5) chama atenção aos papéis emergentes do Estado que, conforme defende, não indicam necessariamente a supressão da centralidade da sua ação ou a diminuição de seu poder de governar. Para a autora, as formas mais recentes de sua atuação revelam a reconstituição e reorientação do Estado em relação à “distribuição do poder entre as escalas supra e subnacional” e aos “movimentos e aspirações sociais e reacções e respostas de *baixo para cima*, no sentido de ampliar a participação democrática e a transparência na administração e no governo” (ANTUNES, 2005, p. 5, *itálicos do original*).

Em atenção aos papéis que assume frente às áreas de sua atuação, a autora define três configurações emergentes do Estado, quais sejam: Estado de competição, Estado em rede e Estado articulador. Para Antunes (2005, p. 39), “embora de modo imperfeito e insatisfatório, [essas configurações] pretendem dar conta de formas específicas de ação estatal que podem surgir singularizadas ou combinadas em determinadas áreas da vida social”, mas, que, no seu conjunto, reforçam o protagonismo do Estado na “gestão de novas formas de regulação social”.

A primeira configuração – Estado de competição – orienta-se, segundo a autora, “para a actuação em instâncias supranacionais e para a intervenção no nível nacional de modo a promover a competitividade da sua economia e a expandir as oportunidades de acumulação” (ANTUNES, 2011, p. 9). A segunda configuração – Estado em rede – tem a ver com as ações do Estado na articulação de determinados segmentos, tendo em vista sua intervenção em áreas da vida social cujo controle escapa às fronteiras de soberania nacional. Essa forma de atuação conformaria, segundo a autora, uma “reestruturação emergente dos papéis do Estado” (ANTUNES, 2011, p. 39), sendo responsável por fazer surgir formas combinadas de administração do Estado em algumas áreas da vida social. Por fim, a terceira configuração – Estado articulador – indica que o Estado se volta à “criação de condições de

mediação dos interesses sociais, sob novas fórmulas e arranjos institucionais em que não é o único nem o principal envolvido” (ANTUNES, 2011, p. 9). Para a autora, é prioritariamente por meio dessas três configurações que o Estado tem reforçado seu protagonismo na produção de novas formas de regulação social.

Barroso (2005) evidencia que, na educação, as medidas políticas e administrativas engendradas pelas novas configurações do Estado ocorrem majoritariamente em três direções. A primeira delas, baseada na racionalidade técnica (neotecnicismo), é justificada pela necessidade de modernização, desburocratização e combate à suposta ineficiência do Estado. Nessa configuração, o Estado incorpora princípios da Nova Gestão Pública (NGP), promovendo uma gestão educacional com base na meritocracia gerencialista, da qual emergem as políticas baseadas em *standards*.

A referência a uma regulação baseada no modelo da NGP tem em mira, segundo o autor, fabricar a imagem de um Estado “menos prescritivo e regulamentador” (BARROSO, 2005, p. 732). Em que pese a ideia de modernização da gestão, o modelo de regulação baseado na NGP toma como referência os resultados educacionais. Trata-se de um controle indireto, que tem como estratégia prioritária uma regulação educacional amparada em resultados de aplicação de exames de largo espectro.

A segunda direção das medidas político-administrativas empreendidas pelo Estado é justificada por imperativos de natureza política e está assentada em projetos neoliberais e neoconservadores. Nessa direção, propõe-se a privatização do sistema público de educação como forma de, supostamente, “libertar a sociedade civil’ do controlo do Estado” (BARROSO, 2005, p. 726, destaque do original). Embora repouse no discurso de melhoria da oferta dos serviços educacionais, as medidas assentadas nesses imperativos inauguram a possibilidade de o setor público ser administrado pela iniciativa privada, matizando, segundo Freitas (2012, p. 386), as relações entre público e privado com a construção da ideia de um “público não estatal”. Trata-se de uma medida que, a nosso ver, aumenta ainda mais a dualidade educacional por focalizar não o bem comum, mas apenas o de alguns.

A terceira direção apontada por Barroso (2005, p. 726) obedece a imperativos “de natureza filosófica e cultural (promover a participação comunitária, adaptar ao local) e de natureza pedagógica (centrar o ensino nos alunos e suas características específicas”. Essas medidas apostam, segundo o autor, em “múltiplas regulações” (BARROSO, 2005, p. 734), nem sempre convergentes. É nesse contexto que, de acordo com Barroso (2005), pais, professores, alunos e sociedade em geral são chamados a contribuir com a escola na melhoria dos resultados educacionais traduzidos como indicadores da qualidade da escola.

Em que pese a distinção de cada uma das medidas político-administrativas empreendidas pelo Estado, em contexto global, é possível verificar algumas

convergências que podem ser explicadas por meio da tese de Dale (2004) sobre a existência de uma Agenda Globalmente Estruturada para a Educação (AGEE), que determina, segundo o autor, políticas comuns em âmbito global.

Para Dale (2004, p. 441), os quadros regulatórios dos Estados-nação são “moldados e delimitados por forças supranacionais, assim como por forças político-econômicas nacionais”. Essas forças, sustenta, operam por meio de diferentes organizações internacionais que, apesar de atuarem em áreas e setores específicos da sociedade e se posicionarem diferentemente em relação a algumas prioridades, fazem-no sempre com o interesse de assegurar a reprodução das relações sociais capitalistas. Imposta pelo processo de globalização, essa agenda condiciona a relação dos Estados-nação com os processos econômicos, políticos e culturais globais, produzindo como consequência uma complexa interação entre os contextos local e global.

A tese da AGEE também pode contribuir para explicar o conceito de governança que emerge em associação com os princípios da NGP. Em sentido amplo, o debate da governança “refere-se aos diferentes mecanismos empregados para conferir ordem à população de atores, por adaptação, negociação, ordem e obediência” (AMOS, 2010, p. 25). Já em seu sentido estrito, refere-se “às diferentes formas de ação resoluta voltada a preocupações coletivas” (AMOS, 2010, p. 25). Em se tratando da educação, a ideia de governança ajuda a explicar o conjunto de medidas operadas pelo Estado nos instrumentos e na organização dos serviços educacionais com vistas a construir determinados consensos sobre o que é importante fazer, em termos de políticas públicas, para garantir certos padrões educacionais de qualidade nas escolas e nos sistemas de ensino.

Comum na ciência política e nas ciências sociais, para Jacobi (2012, p. 335), o conceito de governança relaciona-se “com a implementação socialmente aceitável de políticas públicas”. Trata-se, segundo o autor, de um termo “mais inclusivo que governo, por abranger a relação sociedade, Estado, mercado, direito, instituições, políticas e ações governamentais”, implicando “o estabelecimento de um sistema de regras, normas e condutas que reflitam os valores e visões de mundo daqueles indivíduos sujeitos a esse marco normativo”. Inclui desde leis, regulamentos e instituições, até políticas e ações de governo, bem como iniciativas locais, redes de influência, tais como os mercados internacionais, o setor privado e a sociedade civil.

É principalmente pelo discurso da gestão democrática, fundada na ideia de autonomia, descentralização de responsabilidades e participação da comunidade nas decisões educacionais, bem como da necessidade de melhoria da qualidade educacional, que o conceito de governança educacional se afirma. Pautados nesse discurso, vemos surgir novos aparelhos privados de hegemonia (CASIMIRO, 2018), representados especialmente por grupos de intelectuais de diferentes

segmentos da sociedade civil. Esses segmentos espelham, em alguma medida, alguns dos contornos atuais das parcerias público-privadas, especialmente no Brasil.

Por orientar-se por uma matriz de organização societal neoliberal, liderada pelas elites políticas e econômicas, Antunes (2011) defende que a vaga dos novos padrões de governança produz arranjos contraditórios interpelados e alimentados por aspirações de autonomia e autorregulação. Esses arranjos levam à coexistência de espaços e arenas sociais e políticas cujos processos de decisão e ação são instáveis e irregularmente coordenados. Assim, ainda que evoque os pressupostos de uma gestão democrática, com apelo sedutor e retórico à ampla participação social, ao modelo de governança operado por essa lógica subjaz a ideia de que a comunidade deve “atuar na escola com os olhos do Estado” (SCHNEIDER, 2019, p. 485), fiscalizando as ações de gestão e ajudando a formar consensos em torno de certas decisões políticas. É nessa lógica que o conceito de governança aquiesce aos preceitos político-sociais e econômicos do Estado Avaliador na fase atual de seu desenvolvimento, visto que ambos estão orientados pelos preceitos da NGP, responsável pela introdução de formas de gestão pública pautadas nos resultados.

Em associação com o conceito de governança, a ideia de um Estado Avaliador tem como premissa a produção de *standards* por meio dos quais são definidas as regras de aplicação da meritocracia gerencial. Esse modelo de Estado ganhou força nos anos de 1980, em países centrais, tais como Inglaterra e Estados Unidos, e um dos pioneiros do uso do conceito foi o cientista social britânico Guy Richard Neave.³

Na educação, o estabelecimento de padrões de desempenho, por meio de exames nacionais, transmite a ideia de um Estado preocupado com a qualidade dos serviços educacionais, questão apontada por Afonso (2009, p. 117) como “a principal preocupação do mercado”. Nesse contexto, a avaliação é considerada uma estratégia fundamental para melhorar a gestão das escolas e controlar as despesas públicas, e, por isso, peça central da governança educacional. Desse marco prioritário emerge a necessidade de mobilização de uma diversidade de ferramentas de *accountability* escolar⁴ associadas aos mecanismos de avaliação, com vista a melhorar a eficiência e eficácia dos sistemas de ensino.

3 Segundo Yannoulas, Souza e Assis (2009, p. 59), Neave teria utilizado a expressão “*the rise of the evaluative state*” pela primeira vez em 1988, num artigo publicado pelo *European Journal of Education* com o título *Research, Development and Policy*.

4 Sem tradução exata para o português (CAMPOS, 1990; PINHO; SACRAMENTO, 2009), o conceito de *accountability* está associado à ideia de transparência, prestação de contas e responsabilização. Implica, pois, a mobilização de uma diversidade de ferramentas e de arranjos institucionais para sua implementação. Na educação, as políticas de *accountability* podem assumir formas muito diferentes, de acordo com a “natureza das ferramentas empregadas e da concepção institucional da *accountability* que incorporam” (MAROY; VOISIN, 2013, p. 883).

Na consecução do que consideram, nesse modelo, boas práticas de governança educacional, a eficácia do Estado Avaliador está condicionada ao *quantum* de *accountability* utilizado no controle dos resultados. Procedimentos para avaliar os estabelecimentos de ensino e, direta ou indiretamente, o trabalho dos gestores e dos professores são acompanhados de mecanismos “mais ou menos instantâneos” de *accountability*, supostamente utilizados para “garantir a gestão de todo o sistema educativo, orientar para certas prioridades e controlar a qualidade” (MAROY; VOISIN, 2013, p. 882). Desse modelo, fazem parte também uma diversidade de arranjos institucionais e de bases político-normativas que respondem por uma concepção de escola não mais como instituição, mas como “um sistema de produção escolar” (MAROY; MANGEZ, 2011⁵ *apud* MAROY; VOISIN, 2013, p. 883).

No Estado Avaliador, em vez de o que seria desejável e definido em termos de valores que deveriam ser praticados pela escola, são priorizadas as ações que garantem a eficiência e a eficácia esperadas no desempenho do sistema. Trata-se, conforme Maroy e Voisin (2013, p. 884, destaque do original), “de melhorar a qualidade, a organização, a coordenação, as ‘competências’ dos agentes, todos eles termos que podem ser utilizados em qualquer tipo de organização de produção”.

Apesar das críticas à consistência do Estado Avaliador nas condições de melhoria educacional, o que se sabe é que, em muitos países, as ações voltadas ao robustecimento das ferramentas de avaliação continuam a ser, segundo Afonso (2013, p. 271), “um dos eixos estruturantes das reformas da administração pública”, ou, dito de outra forma, de uma governação por números, indicadores e medidas quantificadas.

ASCENSÃO E CONSOLIDAÇÃO DO MODELO DE ESTADO AVALIADOR NO BRASIL

No Brasil, os primeiros indícios de um Estado Avaliador são evidenciados entre o final dos anos 1980 e início dos anos 1990, na transição do modelo ditatorial para um governo democrático. O contexto de ações que começavam a imperar em âmbito global requeria um trabalhador com outros atributos, competente e polivalente, com capacidade de adequar-se rapidamente a diferentes atividades para manter-se no mercado de trabalho. A formação desse trabalhador ficou a cargo, mais uma vez, da escola, especialmente a de educação básica, que deveria dar conta de preparar os estudantes para as novas demandas do mercado global.

5 MAROY, C.; MANGEZ, C. La construction des politiques d'évaluation et de pilotage du système scolaire en Belgique francophone: nouveau paradigme politique et médiation des experts. *In*: FELOUZIS, G.; HANHART, S. (org.). *Gouverner l'éducation par des nombres?: usages, débats et controverses*. Bruxelles: De Boeck, 2011. p. 53-76.

A guinada nos currículos escolares e na atuação dos docentes para a formação de um trabalhador com os novos atributos requeridos pelo mercado foram consideradas fundamentais. Contudo, essas medidas não foram consideradas suficientes, posto não garantir, na visão dos produtores de política (BALL, 1994), mudanças substanciais nos processos educativos. Foi nesse contexto e alimentado pelo movimento globalizatório que, a essa altura, já se encontrava em estágio avançado a onda avaliacional no Brasil.

Segundo Freitas (2007), as ações desencadeadas na década de 1990, em nosso país, para medir, avaliar e informar tiveram por objetivo central controlar o que estava sendo ensinado nas escolas. Transformada em mecanismo de informação da qualidade educacional das instituições de ensino, a avaliação, em articulação com a pesquisa, a norma legal e o planejamento, fortaleceu e solidificou uma determinada estratégia de atuação do Estado na formulação e execução de políticas públicas para a educação brasileira.

Foi, pois, a partir da criação do Saeb, no ano de 1990, que as medidas impostas pelo Estado Avaliador começaram a adquirir consistência na educação básica. Concebido com a finalidade de fornecer informações sobre o desempenho dos sistemas educativos para gestores educacionais, famílias e sociedade, o Saeb consolidou-se, ao longo dos últimos 30 anos, como uma política nacional de avaliação, ainda que a obrigatoriedade dos entes federativos não estivesse declarada na sua concepção.

Inicialmente, a aplicação do Saeb era por amostra probabilística das escolas públicas urbanas. As provas eram aplicadas para quatro séries do ensino fundamental (1^a, 3^a, 5^a e 7^a), contemplando as disciplinas de Língua Portuguesa, Matemática e Ciências, incluindo ainda Redação para estudantes de 5^a e 7^a séries (WERLE, 2010). Essa dinâmica era assegurada por medidas provisórias, que requeriam regulamentação para converter-se em uma lei ordinária definitiva.

No ano de 1994, a Portaria n. 1.795, de 27 de dezembro (BRASIL, 1994), formalizou a criação do Saeb. Diferentemente da avaliação da educação superior que, desde 2004, conta com aprovação em lei própria,⁶ as reformulações a que foi submetido o Saeb ao longo de seu percurso histórico sempre ocorreram por meio de fundamentos legais provisórios, ou seja, por portarias exaradas pelo Ministério da Educação. Mas isso não representou um impedimento para sua consolidação como iniciativa prática demandada pela administração federal.

No ano de 1995, um ano após a formalização de sua criação, importantes mudanças foram incorporadas de modo a conferir ao Saeb um caráter modernizante

6 Referimo-nos à Lei n. 10.861 (BRASIL, 2004), que institui o Sistema Nacional de Avaliação da Educação Superior (Sinaes).

em relação ao modelo de avaliação praticado no Brasil. As provas realizadas pelos estudantes a cada dois anos passaram a contar com um sistema sofisticado de correção, denominado Teoria da Resposta ao Item (TRI). Trata-se de uma metodologia que qualifica o item a ser avaliado de acordo com algumas características que estimam a habilidade do candidato avaliado a partir de um conjunto de itens, bem como a possibilidade de comparar esse conjunto a outro na mesma escala.

Em 1997, foram elaboradas matrizes curriculares de referência para a constituição das questões e produzidas escalas de referência para mensuração dos resultados. Outra mudança importante naquele ano foi em relação ao público-alvo das provas, que passou a ser estudantes da 4ª série (5ª série, atualmente) e 8ª série (9ª série, atualmente) do ensino fundamental, e 3ª série do ensino médio. As disciplinas avaliadas passaram a ser Língua Portuguesa, Matemática, Ciências Naturais, Física, Química e Biologia. A seleção das disciplinas sofreria alterações na avaliação seguinte, realizada em 1999. Naquele ano, além das avaliadas em 1997, houve acréscimo de história e geografia ao rol das disciplinas avaliadas pelo Saeb.

As medidas introduzidas representaram um avanço importante em direção à consolidação de um sistema nacional de avaliação da educação básica, visto terem sido padronizados os procedimentos avaliativos para todas as escolas participantes, incluindo, a partir de 1997, também uma amostra das escolas da rede privada. A questão das disciplinas avaliadas a cada biênio parecia, contudo, um aspecto não resolvido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), órgão responsável pelo processo de avaliação da educação básica. Desde sua criação, em 1990, até 2001, quando então passaram a ser avaliadas apenas as disciplinas de Língua Portuguesa e Matemática,⁷ a definição das áreas de conhecimento a se avaliar foi um dos aspectos mais contestados na política do Saeb.

A grande virada do Saeb iria acontecer em 2005, quando este deixou de ser apenas amostral e passou a contar com avaliações censitárias. Para Haddad (2008, p. 12), essa reviravolta aumentou “significativamente a responsabilização da comunidade de pais, professores, dirigentes e da classe política com o aprendizado”. Esperava-se que, de posse dos resultados, a comunidade escolar passasse a realizar um acompanhamento mais efetivo das aprendizagens dos estudantes.

A Portaria Ministerial n. 931, de 21 de março de 2005, desdobrou o Saeb em dois processos avaliativos: a Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc), mais conhecida por Prova Brasil. Os objetivos da Anresc foram alargados em relação ao Saeb, passando a ser:

7 Importa lembrar que essas disciplinas avaliadas coincidem com as do Programa Internacional de Avaliação de Estudantes (Pisa). Além de Língua Portuguesa e Matemática, o Pisa avalia também o domínio dos estudantes em Ciências.

- a) avaliar a qualidade do ensino ministrado nas escolas, de forma que cada unidade escolar receba o resultado global;
- b) contribuir para o desenvolvimento, em todos os níveis educativos, de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação brasileira e adequados controles sociais de seus resultados;
- c) concorrer para a melhoria da qualidade de ensino, redução das desigualdades e a democratização da gestão do ensino público nos estabelecimentos oficiais, em consonância com as metas e políticas estabelecidas pelas diretrizes da educação nacional;
- d) oportunizar informações sistemáticas sobre as unidades escolares. Tais informações serão úteis para a escolha dos gestores da rede a qual pertençam. (BRASIL, 2005, p. 1)

A alteração da configuração e de alguns de seus objetivos, viabilizada pela referida portaria, foi extremamente importante na medida em que tornou possível a produção de dados por escola, elemento fundamental à implementação de ferramentas de *accountability* alinhadas aos resultados das avaliações do Saeb. Com o desenvolvimento de provas censitárias, abriu-se a possibilidade de todas as escolas públicas e privadas de educação básica terem resultados do Saeb e poderem ser incluídas no Índice de Desenvolvimento da Educação Básica (Ideb), índice criado dois anos depois, em 2007, com o intuito de permitir a regulação dos sistemas com base em *standards* educacionais a partir da inclusão em avaliações censitárias.

Inaugurado com o objetivo anunciado de promover políticas para a melhoria educacional, o Ideb afere o cumprimento de metas fixadas no Plano de Metas Compromisso Todos pela Educação, instituído pelo Decreto n. 6.094 (BRASIL, 2007). A definição de metas intermediárias por escola, unidades da federação e sistemas de ensino, a serem alcançadas a cada dois anos nas avaliações do Saeb, constitui outra peça importante na consolidação do modelo de Estado Avaliador para a educação básica brasileira.

Do ponto de vista do potencial do Ideb de pressionar as escolas a trabalhar pelo atendimento às metas, Fernandes e Gremaud (2009, p. 11), responsáveis pela criação do índice, afirmam que ele compõe um modelo de “*accountability* fraca” ou, conforme Bonamino e Sousa (2012, p. 380), uma “responsabilização branda”, uma vez que “a Prova e o uso de seus resultados para composição do Ideb [...] se limitam a traçar metas e a divulgar os resultados dos alunos por escola e rede de ensino, sem atrelar prêmios ou sanções a esses resultados”.

As denominações dadas por esses autores consideram que uma *accountability* fraca, em oposição a uma *accountability* forte, tem em conta a decorrência ou não de consequências pelos resultados logrados em avaliações tais como a do Saeb. No

caso do Ideb, a produção de uma *accountability* fraca remeteria à inexistência de ferramentas com altas consequências para as escolas pelo não alcance das metas, tais como sanções condicionadas ao acesso a níveis superiores de educação aos alunos, publicação de resultados aos pais com objetivo de proporcionar a livre escolha da escola e, até mesmo, ameaça de substituição da equipe de gestão em caso de repetidos maus desempenhos e fechamento das escolas (MAROY; VOISIN, 2013). De igual forma, não há previsão de premiação às escolas na forma de incentivos financeiros individuais ou coletivos associados aos resultados no Ideb.

Apesar de a regulação por resultados do Ideb não produzir efeitos como os mencionados, algumas condições orquestradas em escala federal, estadual e municipal vêm contribuindo para a produção de um sistema de *accountability* com grandes desafios para as escolas brasileiras. Constituem exemplos de políticas que atuam nessa direção a definição de normas curriculares nacionais, com a implantação da Base Nacional Comum Curricular (BNCC), o alinhamento das matrizes de referência e da formação inicial e continuada de professores a essas normas e a criação de sistemas estaduais e municipais de avaliação com utilização de ferramentas de *accountability* de alto impacto.

Sousa e Koslinski (2017) identificam treze iniciativas de políticas desenvolvidas por estados da federação que premiam profissionais da educação, alunos ou escolas pelo alcance de metas traçadas no âmbito das redes públicas estaduais de ensino. Essas iniciativas localizam-se, conforme os autores, nas seguintes regiões e estados brasileiros: Norte (Amazonas, Roraima e Tocantins); Nordeste (Ceará, Paraíba, Pernambuco e Sergipe); Centro-Oeste (Goiás e Mato Grosso do Sul); Sudeste (todos os estados da região, ou seja, Espírito Santo, Minas Gerais, Rio de Janeiro e São Paulo). Nos estados da região Sul, não foram identificadas práticas de responsabilização de alto impacto (SOUSA; KOSLINSKI, 2017).

A conformação do Estado Avaliador, na educação básica brasileira, assume novos contornos a partir de 2018, com a promulgação do Decreto presidencial n. 9.432 (BRASIL, 2018). Esse decreto definiu a Política Nacional de Avaliação e Exames da Educação Básica, reposicionando e alterando sua composição, ao incluir, na política nacional do Saeb, o Exame Nacional do Ensino Médio (Enem) e o Exame de Certificação da Competências de Jovens e Adultos (Encceja). Apesar de praticados desde 1998 e 2002, respectivamente, estes exames não faziam parte do Saeb. Outras avaliações realizadas na educação básica, tais como a Avaliação Nacional da Alfabetização (ANA), criada em 2013, também não faziam parte do Saeb.

No decreto, o Saeb é referido como “um conjunto de instrumentos que permite a produção e a disseminação de evidências, estatísticas, avaliações e estudos a respeito da qualidade das etapas que compõem a educação básica” (BRASIL, 2018, Art. 5º). Essa concepção pode tornar o sistema mais abrangente na medida em que

a composição dos resultados pode contar com dados que não sejam unicamente os obtidos em provas e exames. Contudo, por enquanto, trata-se de uma promessa, porquanto não são declaradas as fontes de dados de produção dessas informações. O documento informa apenas que a aferição do Saeb contará “com a coleta de dados junto aos sistemas de ensino e às escolas públicas e privadas brasileiras” (BRASIL, 2018, art. 5º, § único). Apesar de, até 2018, serem incluídos questionários a diretores e professores das escolas participantes dos exames, eles não eram considerados na pontuação do Saeb.

Enquanto um dos atos do presidente interino Michel Temer (2016-2018), o decreto estabeleceu um novo estatuto ao Saeb, o de uma peça da política nacional de avaliação, e não mais a política em si. Passa a compreender, a partir de então, as avaliações que compõem as etapas da educação básica, quais sejam, educação infantil, ensino fundamental e ensino médio.

A inclusão da educação infantil na política nacional de avaliação pareceu almejar resolver uma celeuma sobre essa etapa da educação básica, posto que o tema tem sido, nos últimos doze anos, um dos pontos centrais nas discussões de especialistas sobre o aprendizado das crianças. Contudo, deixa em aberto um ponto importante dessa discussão, uma vez que, apesar de especulações sobre como será o desenho da avaliação da educação infantil, no documento nada é mencionado a respeito.

Tão importante quanto o Decreto n. 9.432 (BRASIL, 2018) é a Portaria n. 458 (BRASIL, 2020), exarada em 5 de maio de 2020, na vigência do segundo ano de mandato do presidente Jair Messias Bolsonaro. Essa portaria instituiu normas complementares consideradas *necessárias* ao cumprimento da Política Nacional de Avaliação da Educação Básica. Nela, reafirma-se o compromisso que os sistemas de ensino de todos os entes federativos devem ter com o processo avaliacional e é admitida a possibilidade de servidores públicos e integrantes dos sistemas de ensino participarem da coleta de dados do Saeb, “desde que o ônus da respectiva remuneração seja do órgão ou entidade cedente” (BRASIL, 2020, art. 3º, § único).

Ao que parece, trata-se de uma tentativa de o Estado salvaguardar a não remuneração de funcionários públicos caso eles possam ser chamados a fornecer informações sobre os sistemas de ensino ou escolas de educação básica, considerando, especialmente, as medidas necessárias à implementação da política nacional de avaliação. A redução de despesas na administração pública é uma das atribuições do Estado Avaliador.

A portaria regulamenta, ainda, a realização dos exames que compõem o Saeb, que passam a ser anuais, de caráter censitário e com o objetivo de “aferir o domínio das competências e das habilidades esperadas ao longo da educação básica, de acordo com a BNCC e as correspondentes diretrizes curriculares nacionais”

(BRASIL, 2020, art. 8º). Até 2019, as provas que compreendiam a Saeb (Aneb e Anresc) eram bianuais.

Embora, à primeira vista, possa parecer que a promulgação do Decreto n. 9.432 (BRASIL, 2018) e da respectiva portaria (BRASIL, 2020) pouco alterou em relação à política nacional de avaliação praticada até então, as mudanças operadas revelam importante avanço no modelo de Estado Avaliador em nosso país. O alinhamento da política nacional de avaliação com a BNCC cumpre importante medida nessa direção.

Implantada em 2017, e em processo de implementação pelos sistemas de ensino e escolas públicas de educação básica, a BNCC definiu um conjunto de competências gerais para toda a educação básica, competências específicas para cada área de conhecimento e componentes curriculares e, por fim, habilidades relacionadas aos objetos de conhecimento (conteúdos, conceitos e procedimentos), organizados em unidades temáticas (BRASIL, 2017).

A definição da concepção pedagógica das avaliações, incumbência atribuída ao Estado a partir de 2018, é conferida à BNCC, que, como sabemos, resultou de um trabalho articulado de uma rede nacional de governança⁸ que teve por objetivo dar legitimidade à base. A ênfase em aprendizagens essenciais remonta à ideia de currículos mínimos, requeridos pelos setores produtivos.

A padronização curricular imposta pela BNCC aos sistemas de ensino e às escolas de educação básica exige do Ministério da Educação (MEC) e do Inep, instituições incumbidas de implantar os procedimentos atinentes à política nacional de avaliação, uma recomposição do modelo curricular praticado na Política Nacional de Avaliação. Desde 1997, foram utilizadas matrizes de referência nos exames e provas do Saeb, com a descrição de competências e habilidades que os estudantes deveriam dominar em cada série coberta pelas avaliações. A construção dessas matrizes deu-se a partir do levantamento dos conteúdos praticados nas escolas públicas de ensino fundamental e médio. O alinhamento dessa política à BNCC impõe a revisão desse modelo e dos objetos de sua composição.

As reformulações operadas ao longo da história do Saeb e, especialmente, a partir da regulamentação da política nacional, evidenciam um conjunto desarticulado de ações em torno da constituição da Política Nacional de Avaliação da Educação Básica. Contudo, esse processo parece não ter afetado a edificação e o fortalecimento do Estado Avaliador no Brasil que, como foi possível constatar, assegura ao Estado maior poder de controle da educação, define metas para as escolas, descentraliza responsabilidades, determinando um novo estatuto de governança nacional.

8 Participaram da elaboração da BNCC a Fundação Lemann, o Instituto Inspirare, o Instituto Península, a Fundação Roberto Marinho, o Instituto Ayrton Senna, o Instituto Natura, o Instituto Unibanco, personalidades políticas, gestores educacionais, entre outros sujeitos individuais e coletivos.

CONSIDERAÇÕES FINAIS

O processo de transformação do Estado em Avaliador, no Brasil, configura-se em um momento histórico essencial à compreensão da atual conjuntura de formulação e implementação das políticas educacionais. A implantação do Saeb, em 1990, e a do Ideb, em 2007, marcam, de um lado, o surgimento desse modelo de Estado e, de outro, o seu fortalecimento, respectivamente. Um Estado que, aludindo à descentralização de responsabilidades administrativas, promete auxílio técnico e financeiro aos entes federativos por meio de ações compartilhadas; estabelece, arbitrariamente e verticalmente, metas educacionais para as escolas e redes de ensino; e difunde um discurso de participação e democracia ao mesmo tempo em impõe o desenvolvimento de arranjos institucionais locais para o cumprimento das metas projetadas.

Na consecução desse modelo, há o robustecimento de ferramentas de *accountability* escolar, utilizadas não apenas para avaliar as instituições de ensino, mas principalmente o trabalho de seus gestores e professores. O uso dessas ferramentas passa a orientar as prioridades e controlar a qualidade da educação, definida por *standards* de desempenho.

De forma conclusiva, reiteramos que os estudos sobre os contornos do Estado Avaliador na educação básica brasileira têm procurado demonstrar os efeitos e as consequências das avaliações considerando o viés meritocrático e gerencialista das reformas educacionais. Contudo, há que se pôr em evidência as razões, especialmente num momento em que, marcado pelo isolamento social imposto pela pandemia do novo coronavírus (Covid-19), até mesmo a necessidade da educação escolar é colocada à prova.

O estudo realizado reforça a convicção acerca da importância de aprofundamento do debate acadêmico sobre o Estado, considerando a redefinição de suas funções mediante a formação de parcerias público-privadas sob um outro estatuto de governança.

REFERÊNCIAS

ABRUCIO, Fernando Luiz. *O impacto do modelo gerencial na Administração Pública: um breve estudo sobre a experiência internacional recente*. Brasília, DF: Escola Nacional de Administração Pública, 1997. (Cadernos Enap, n. 10).

AFONSO, Almerindo Janela. *Avaliação educacional: regulação e emancipação*. São Paulo: Cortez, 2005.

AFONSO, Almerindo Janela. Nem tudo o que conta em educação é mensurável ou comparável: crítica à *accountability* baseada em testes estandardizados e rankings escolares. *Revista Lusófona de Educação*, Lisboa, n. 13, p. 13-29, 2009.

AFONSO, Almerindo Janela. Mudanças no Estado-avaliador: comparativismo internacional e teoria da modernização revisitada. *Revista Brasileira de Educação*, Rio de Janeiro, v. 18, n. 53, p. 267-284, abr./jun. 2013.

AMOS, Karin. Governança e governamentalidade: relação e relevância de dois conceitos científico-sociais proeminentes na educação comparada. *Educação e Pesquisa*, São Paulo, v. 36, n. especial, p. 23-38, abr. 2010.

ANTUNES, Fátima. Reconfigurações do Estado e da Educação: novas instituições e processos educativos. *Revista Lusófona de Educação*, Lisboa, v. 5, p. 37-62, 2005.

ANTUNES, Fátima. Governança, reformas do Estado e políticas de educação de adultos em Portugal: pressões globais e especificidades nacionais, tensões e ambivalências. *Revista Crítica de Ciências Sociais*, Coimbra, v. 92, p. 3-29, mar. 2011.

BALL, Stephen J. *Education reform: a critical and post-structural approach*. Buckingham: Open University Press, 1994.

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.

BARROSO, João. O Estado, a educação e a regulação das políticas públicas. *Educação & Sociedade*, Campinas, SP, v. 26, n. 92 (n. especial), p. 725-751, out. 2005.

BARROSO, João. Da política baseada no conhecimento às práticas baseadas em evidência – consequências para a regulação do trabalho docente. In: DUARTE, Adriana; OLIVEIRA, Dalila Andrade de (org.). *Políticas públicas e educação: regulação e conhecimento*. Belo Horizonte: Fino Traço, 2011. p. 91-116.

BONAMINO, Alicia; SOUSA, Sandra Zakia de. 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, abr./jun. 2012.

BRASIL. Ministério da Educação. Portaria n. 1.795, de 27 de dezembro de 1994. Cria o Sistema Nacional da Educação Básica. *Diário Oficial da União*, Brasília, DF, 28 dez. 1994. Seção 1, p. 91-92. Disponível em: <http://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=91&data=28/12/1994>. Acesso em: 23 jan. 2020.

BRASIL. Lei n. 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior (Sinaes). *Diário Oficial da União*, Brasília, DF, 15 abr. 2004. Seção 1, p. 3-4. Disponível em: https://abmes.org.br/arquivos/legislacoes/Lei_10861_140404.pdf. Acesso em: 27 jan. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Portaria n. 931, de 21 de março de 2005. Institui o Sistema de Avaliação da Educação Básica – SAEB, que será composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica – Aneb, e a Avaliação Nacional do Rendimento Escolar – Anresc. *Diário Oficial da União*, Brasília, DF, 22 mar. 2005. Seção 1, p. 17. Disponível em: <http://provabrazil.inep.gov.br/legislacao>. Acesso em: 20 out. 2019.

BRASIL. Decreto n. 6.094, de 24 de abril de 2007, dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação. *Diário Oficial da União*, Brasília, DF, 25 abr. 2007. Seção 1, p. 5-6. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6094.htm. Acesso em: 12 jan. 2020.

BRASIL. Ministério da Educação. *Base Nacional Comum Curricular*. Brasília: Ministério da Educação, 2017. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_20dez_site.pdf. Acesso em: 27 dez. 2019.

BRASIL. Decreto n. 9.432, de 29 de junho de 2018. Regulamenta a Política Nacional de Avaliação e Exames da Educação Básica. *Diário Oficial da União*, Brasília, DF, 2 jun. 2018. Seção 1, p. 1. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Decreto/D9432.htm. Acesso em: 27 nov. 2019.

BRASIL. Portaria n. 458, de 5 de maio de 2020. Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica. *Diário Oficial da União*, Brasília, DF, 6 maio 2020. Seção 1, p. 57. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Decreto/D9432.htm. Acesso em: 27 nov. 2019.

CAMPOS, Anna Maria. Accountability: quando poderemos traduzi-la para o português? *Revista de Administração Pública*, Rio de Janeiro, v. 24, n. 2, p. 30-50, fev./abr. 1990.

CASIMIRO, Flávio Henrique Calheiros. *A nova direita: aparelhos de ação política e ideológica no Brasil contemporâneo*. São Paulo: Expressão Popular, 2018.

DALE, Roger. Globalização e educação: demonstrando a existência de uma “Cultura Educacional Mundial Comum” ou localizando uma “Agenda Globalmente Estruturada para a Educação”? *Educação & Sociedade*, Campinas, SP, v. 25, n. 87, p. 423-460, 2004.

FERNANDES, Reynaldo; GREMAUD, Amauri Patrick. *Qualidade da educação: avaliação, indicadores e metas*. São Paulo: Fundação Santillana; Inep, 2009.

FREITAS, Dirce Nei Teixeira de. *A avaliação da educação básica no Brasil: dimensão normativa, pedagógica e educativa*. Campinas: Autores Associados, 2007.

FREITAS, Luiz Carlos de. Os reformadores empresariais da educação: da desmoralização do magistério à destruição do sistema público de educação. *Educação & Sociedade*, Campinas, SP, v. 33, n. 119, p. 379-404, abr./jun. 2012.

HADDAD, Fernando. *O Plano de Desenvolvimento da Educação: razões, princípios e programas*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2008. (Série Documental: Textos para Discussão, v. 30).

JACOBI, Pedro R. Governança ambiental, participação social e educação para a sustentabilidade. In: PHILIPPI JR., Arlindo et al. (org.). *Gestão de natureza pública e sustentabilidade*. Barueri: Manole, 2012. p. 343-361.

MAINARDES, Jeferson. Reflexões sobre o objeto de estudo da Política Educacional. *Laplage em Revista*, Sorocaba, SP, v. 4, n. 1, p. 186-201, jan./abr. 2018.

MAROY, Christian; VOISIN, Annelise. As transformações recentes das políticas de *accountability* na educação: desafios e incidências das ferramentas de ação pública. *Educação & Sociedade*, Campinas, SP, v. 34, n. 124, p. 881-901, jul./set. 2013.

PINHO, José Antonio Gomes de; SACRAMENTO, Ana Rita Silva. *Accountability*: já podemos traduzi-la para o português? *Revista de Administração Pública*, Rio de Janeiro, v. 43, n. 6, p. 1343-1368, nov./dez. 2009.

SCHNEIDER, Marilda Pasqual. Dispositivos de *accountability* na reforma da educação básica brasileira: tendências em curso. *Revista Diálogo Educacional*, Curitiba, v. 19, n. 60, p. 469-493, jan./mar. 2019.

SCHNEIDER, Marilda Pasqual; ROSTIROLA, Camila Regina. Estado-Avaliador: reflexões sobre sua evolução no Brasil. *Revista Brasileira de Política e Administração da Educação*, Brasília, v. 31, n. 3, p. 493-510, set./dez. 2015.

SOUSA, Sandra Zákia; KOSLINSKI, Mariane. Avaliação em larga escala, índices e premiação: iniciativas de estados brasileiros e seus efeitos. In: SETTON, Maria da Graça Jacintho *et al.* (org.). *Mérito, desigualdade e diferença: cenários da (in)justiça escolar Brasil e Portugal*. Alfenas: Editora Universidade Federal de Alfenas, 2017.

WERLE, Flávia Obino Corrêa. Papel dos pais na busca de um padrão alto de democratização, acesso e permanência na escola. In: WERLE, Flávia Obino Corrêa (org.). *Avaliação em larga escala: foco na escola*. São Leopoldo: Oikos; Brasília: Liber Livros, 2010. p. 188-202.

YANNOULAS, Sílvia Cristina; SOUZA, Camila Rosa Fernandes de; ASSIS, Samuel Gabriel. Políticas educacionais e o estado avaliador: uma relação conflitante. *Sociedade em Debate*, Pelotas, RS, v. 5, n. 2, p. 55-67, jul./dez. 2009. Disponível em: <http://www.rle.ucpel.tche.br/index.php/rsd/article/view/351>. Acesso em: 10 nov. 2019.

Recebido em: 28 JANEIRO 2020

Aprovado para publicação em: 1 OUTUBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

<https://doi.org/10.18222/ea.v31i78.7073>

UMA ANÁLISE DOS RESULTADOS EDUCACIONAIS DOS ESTADOS BRASILEIROS

 PEDRO JORGE HOLANDA ALVES^I

 JEVUKS MATHEUS ARAÚJO^{II}

 GUSTAVO SARAIVA FRIO^{III}

 LYVIA CABRAL CORDEIRO^{IV}

^I Universidade Católica de Brasília (UCB), Brasília-DF, Brasil; pedrojorge_holanda@hotmail.com

^{II} Universidade Federal da Paraíba (UFPB), João Pessoa-PB, Brasil; jevuks@gmail.com

^{III} Universidade de Brasília (UnB), Brasília-DF, Brasil; gustavo.frio@gmail.com

^{IV} Universidade Federal Rural de Pernambuco (UFRPE), Recife-PE, Brasil; lyviacordeiro@hotmail.com

RESUMO

O objetivo deste trabalho é analisar os indicadores da qualidade educacional dos estados brasileiros entre os anos de 2007 e 2017, bem como apresentar uma medida de eficiência técnica para eles. Utilizou-se o método de fronteira estocástica para os anos finais do ensino fundamental e o ensino médio, encontrando evidências que favorecem a importância da qualificação do docente e do tamanho da turma para melhores desempenhos no Índice de Desenvolvimento da Educação Básica. O estudo também mostra que a estrutura escolar obteve resultados negativos para o desempenho no ensino médio e que as escolas responsáveis pelos anos finais do ensino fundamental foram eficientes na realização do gasto, enquanto o ensino médio apresentou ineficiência dos gastos públicos estaduais em educação.

PALAVRAS-CHAVE AVALIAÇÃO DA EDUCAÇÃO • INDICADORES EDUCACIONAIS • EFICIÊNCIA • QUALIDADE DA EDUCAÇÃO.

UN ANÁLISIS DE LOS RESULTADOS EDUCACIONALES DE LOS ESTADOS BRASILEÑOS

RESUMEN

El objetivo de este trabajo es analizar los indicadores de la calidad educacional de los estados brasileños entre los años de 2007 y 2017, así como presentar una medida de eficiencia técnica para ellos. Se utilizó el método de frontera estocástica para los años finales de la educación básica y media, encontrando evidencias que favorecen la importancia de la cualificación del docente y del tamaño del grupo para mejorar el desempeño en el Índice de Desarrollo de la Educación Básica. La estructura escolar obtuvo resultados negativos para el desempeño en la educación media, y las escuelas responsables por los años finales fueron eficientes en la realización del gasto, mientras la educación media presentó ineficiencia de los gastos públicos estadales en educación.

PALABRAS CLAVES EVALUACIÓN DE LA EDUCACIÓN • INDICADORES EDUCATIVOS • EFICIENCIA • CALIDAD DE LA EDUCACIÓN.

ANALYSIS OF THE EDUCATIONAL RESULTS OF BRAZILIAN STATES

ABSTRACT

This study aims to analyze the educational quality indicators of Brazilian states between 2007 and 2017, as well as to present a measure for their technical efficiency. The stochastic frontier method was used for the final years of elementary school and high school, and showed the importance of teacher qualification and class size to achieve better performances in the Basic Education Development Index. School structure obtained negative results for high school performance. The elementary schools responsible for the final years were efficient in terms of education spending, contrary to high schools that proved inefficient regarding the states' public education spending.

KEYWORDS EDUCATION ASSESSMENT • EDUCATIONAL INDICATORS • EFFICIENCY • QUALITY OF EDUCATION.

INTRODUÇÃO

A educação é fundamental para uma sociedade, tanto na questão de externalidades positivas como na questão do capital humano. Segundo Schultz (1961), o “capital educacional” é considerado importante para o aumento da produtividade, que por sua vez é um dos fatores chave para o crescimento econômico. Assim, a demanda por educação é tratada como uma forma de investimento, visto que a teoria do capital humano a vê como um insumo de produtividade.

Segundo Barros, Henriques e Mendonça (2001), a educação também entra como fator importante no combate da desigualdade de renda no Brasil. Para os autores, eliminar os diferenciais de renda por nível educacional poderia originar uma redução entre metade e um terço da desigualdade de renda. De acordo com Fernandes e Narita (2001), essa influência acontece por conta da desigualdade educacional entre os trabalhadores e a sensibilidade dos salários com base no nível educacional.

Dessa forma, visando a reduzir a desigualdade e seguindo os pressupostos de atribuições distributiva e locativa do Estado (MUSGRAVE, 1956), o Brasil, desde a Constituição de 1988, garante como direito fundamental a proteção aos desamparados, a educação, a saúde, entre outros, e ainda garante prioridade ao atendimento do ensino básico, com acesso universal e equitativo. Contudo, Hanushek e Raymond (2005) alertam que aplicações apenas em insumos podem resultar em perdas de eficiência das escolas, pois seriam recompensadas pelos seus processos sem obter resultado significativo em relação ao seu desempenho. De acordo com Mandl, Dierx e Ilzkovitz (2008), compreender até que ponto os governos podem aumentar seu desempenho mantendo gastos constantes e simplesmente aumentando a eficiência dos gastos poderia ajudar os formuladores de políticas a alcançar disciplina fiscal sustentada.

Assim, não basta que o gasto com educação seja elevado, é preciso que seja aliado a medidas administrativas eficientes (tais como diminuir o tamanho das turmas e melhorar a qualificação do docente). Para que o resultado seja positivo, é necessário planejamento, organização, formação continuada e avaliação para que se gere elementos construtivos para a sociedade e se atinja seus objetivos (DIAZ, 2007).

A literatura relacionada à avaliação de eficiência do gasto do governo tem sido geralmente obtida por meio de fronteiras de eficiência, tanto na aplicação paramétrica, com a Análise de Fronteira Estocástica (SFA), como na aplicação não paramétrica, com o Free Disposal Hull (FDH) e Data Envelopment Analysis (DEA). A eficiência do gasto público foi avaliada por autores como Gupta, Verhoeven e Tiongson (2002), Afonso, Schuknecht e Tanzi (2003, 2005), Herrera e Pang (2005), Afonso e Aubyn (2005), Sutherland *et al.* (2007), Afonso e Fernandes (2008), Aubyn, Garcia e Pais (2009), Afonso, Romero e Monsalve (2013), Qutb (2016), Almeida e Aguiar (2017) e Alves e Araújo (2018), para vários países.

Coleman *et al.* (1966) buscaram analisar os diferentes resultados dos exames nas escolas dos Estados Unidos com base nas diferenças de receita de cada escola e nas características socioeconômicas dos alunos. Além de provar que não se pode apenas utilizar as notas dos alunos como medida, conseguiram mensurar a qualidade das escolas por meio dos insumos escolares. Em seu relatório, concluíram que o desempenho de cada aluno está relacionado às características familiares, não se podendo atribuir a proficiência dos alunos apenas à escola ou à gestão.

O objetivo deste trabalho é estimar a eficiência das Unidades Federativas (UFs) brasileiras na provisão de educação pública para os anos finais do ensino fundamental e para o ensino médio. O Índice de Desenvolvimento da Educação Básica (Ideb) foi escolhido como variável dependente; assim o fizeram diversos autores por ser um bom indicador para mensurar a eficiência e por ser um indutor das políticas educacionais (SILVA; SILVA; SANTOS, 2019) e uma política de Estado capaz de impulsionar as práticas de ensino, a capacitação e a habilitação dos professores e a elaboração de políticas que fomentam o treino (ALVES; FIALHO; LIMA, 2018).

O presente trabalho mostra que os resultados educacionais (avaliados pelo Ideb) não acompanharam o crescimento do gasto público no ensino médio, tendo havido um deslocamento negativo na fronteira de produção. Esse resultado nos leva a inferir a má alocação dos recursos aplicados na educação. Também observamos que o estado de São Paulo (que saiu de sexto para quarto lugar na classificação do Ideb do ensino médio entre 2007 e 2017) é o 25º no *ranking* de eficiência do ensino médio, ficando à frente apenas do Rio Grande do Norte e da Bahia, indicando que, apesar de evoluir e ter um dos melhores Ideb do país, o investimento em educação precisa ser mais bem alocado para alcançar melhores resultados.

A estrutura deste trabalho é formada, além de por esta introdução, por mais quatro partes. Na segunda seção, é apresentado o modelo de fronteira estocástica, com sua estrutura e hipóteses. Na terceira seção, o cenário educacional brasileiro é comparado ao de diversos países do mundo, com sua evolução ao longo do tempo. Os resultados estatísticos e econométricos são apresentados na quarta seção. Por fim, na quinta seção, os principais achados do trabalho são apresentados, seguidos pelas considerações finais.

METODOLOGIA

Análise da fronteira estocástica

Partindo da necessidade de se estimar a eficiência, começou-se a realizar estudos para mensurar a fronteira. Segundo Kumbhakar e Lovell (2000) e Varian (1999), para analisar as fronteiras de produção e de custo é necessário considerar os componentes de eficiência técnica e alocativa.

Marinho e Ataliba (2000) concluíram que o objetivo principal da fronteira é estimar uma função de produção, cuja finalidade é obter o máximo de eficiência do seu produto. Contudo, a existência de ineficiência técnica (gerando um resultado abaixo da fronteira máxima de produção) não garante o resultado eficiente. Assim, surgiu o modelo de produção estocástica.

Pereira e Moreira (2007) afirmam que a metodologia da Fronteira de Produção Estocástica é caracterizada por adicionar um termo de erro à função. O objetivo é medir a ineficiência técnica e a ineficiência dos efeitos aleatórios (KUMBHAKAR; LOVELL, 2000).

A mensuração de eficiência pode ser feita a partir de dados de cortes transversal ou em painel. Segundo Kumbhakar e Lovell (2000), sua especificação pode ser a mesma, mas com estimações distintas, podendo representar uma distribuição normal, normal truncada (STEVENSON, 1980; BATTESE; COELLI, 1988, 1992; WANG, 2002), gama (GREENE, 1990) e exponencial (MEEUSEN; VAN DEN BROECK, 1977; AIGNER; LOVELL; SCHMIDT, 1977). Os autores afirmam que a diferença na escala da distribuição dos impactos é pequena nas medidas de eficiência. Por isso, apesar de as distribuições normal e normal truncada receberem algumas críticas, são utilizadas por possuírem modelos mais simples, tornando-as as mais comuns distribuições na literatura técnica.

A partir do modelo proposto por Battese e Coelli (1988, 1992, 1995), foi possível gerar o modelo de fronteira estocástica para captar os efeitos da ineficiência, tentando captá-los pelo modelo invariante (presume-se que o termo de ineficiência tenha uma distribuição normal truncada) no tempo e variante (presume-se que o termo de ineficiência tenha uma variável aleatória normal truncada multiplicada por uma função específica variante do tempo) no tempo. Ainda que seja possível a estimação por mínimos quadrados ordinários (MQO) e mínimos quadrados generalizados (MQG), os estimadores utilizados no modelo foram ajustados e obtidos pelo estimador de Máxima Verossimilhança. Com isso, o modelo é apresentado como:

$$Y_{it} = \exp(X_{it}\beta + v_{it} - u_{it}) \quad (1)$$

em que o modelo exponencial pode ser definido da seguinte forma: Y_{it} é produto do i -ésimo estado (*output*); β são os vetores ($k \times 1$) de insumos a serem estimados; i é o indivíduo, cada um dos estados do Brasil; t é a variação no tempo; $v_{it} \sim N(0, \sigma_v^2)$ é o componente do erro aleatório; $u_{it} \sim N^+(\mu, \sigma^2)$ é o termo de erro aleatório não observável em relação ao termo de ineficiência técnica. O termo de erro é composto pela captação dos aspectos relacionados à ineficiência técnica e aos resíduos estatísticos.

$$\varepsilon \equiv v_{it} - u_{it} \quad (2),$$

De acordo com Battese e Coelli (1995), a fronteira estocástica pode ser especificada da mesma forma representada em (2). Contudo, o termo de ineficiência é definido como variável aleatória, expressado da seguinte forma:

$$u_{it} = \delta Z_{it} + W_{it} \quad (3),$$

em que W_{it} representa uma variável aleatória com média zero e variância constante; Z_{it} representa um vetor ($p \times 1$) de variáveis que afetam a ineficiência; δ representa um vetor ($1 \times p$) de estimação dos parâmetros.

A eficiência técnica (ET) foi definida nos estudos de Battese e Coelli (1992, 1995) da seguinte forma:

$$e^{-u_i} = ET_{it} = \frac{E[y_{it}|x_{it}, u_i]}{E[y_{it}|x_{it}, u_i = 0]} = \frac{Y_i}{Y_i^*} \quad (4),$$

em que a eficiência técnica (ou o termo de ineficiência) é mensurada pela razão entre a produção observada e a estimada (da mesma forma que na fronteira determinística).

Como foi explicitado na equação (2), o termo de erro é composto por dois componentes. O primeiro é representado pelos desvios relacionados ao componente de choque aleatório v_{it} , que é independentemente distribuído sobre as observações da amostra, enquanto u_{it} é representado pelos desvios em relação aos componentes de eficiência. Seguem as hipóteses de validade:

- a) Resíduo com distribuição normal com média zero e variância constante.
- b) Coeficiente de ineficiência com distribuição normal truncada com média e variância constante.
- c) Independência entre o termo de ineficiência e o resíduo.
- d) Independência do termo de ineficiência e do resíduo em relação às variáveis explicativas.

Além dos modelos de estimação, a modelagem de eficiência considera duas formas funcionais. Uma representada pela função Cobb-Douglas, e a outra, no modelo da função Translog, como descrito a seguir.

Cobb-Douglas:

$$\ln(y_{it}) = \beta_0 + \sum_{n=1}^N \beta_n \ln(x_{it}) + (v_{it} - u_{it}) \quad (5),$$

Translog:

$$\ln(y_{it}) = \beta_0 + \sum_{n=1}^N \beta_n \ln(x_{nit}) + \frac{1}{2} \sum_{n=1}^N \sum_{j=1}^N \beta_{nj} \ln(x_{it}) * \ln(x_{it}) + (v_{it} - u_{it}) \quad (6),$$

A melhor forma escolhida é encontrada pelo teste da razão de Máxima Verossimilhança (*likelihood-ratio test*). A estatística do teste (LR) é dada por:

$$LR = -2[\ln LLH_0 - \ln LLH_1] \quad (7),$$

em que H_0 representa o LL Cobb-Douglas, e H_1 representa o LL Translog. Deve-se rejeitar H_0 se $LR > TKP$ (KODDE; PALM, 1986).

PANORAMA EDUCACIONAL BRASILEIRO

O Brasil tem buscado políticas públicas que visem a melhorar o sistema e o desempenho público, principalmente após a Constituição de 1988. Uma das medidas para a melhoria do sistema foi a Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), que busca regulamentar o sistema educacional (público ou privado) do Brasil, da educação básica ao ensino superior. O objetivo é estabelecer os princípios educacionais e os deveres do Estado, definindo as responsabilidades em regime de colaboração entre as esferas federativas.

Como forma de incentivar os estados e municípios a atingirem tais metas, a união oferece ganhos de transferências de verbas vinculadas diretamente à educação pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb) e Fundo Nacional de Desenvolvimento da Educação (FNDE) – que estipulam que os estados e municípios devem gastar no mínimo 25% da receita corrente líquida em educação – e pela Lei de Diretrizes Básicas – que determina que os municípios devem gastar 60% da receita do Fundeb em remuneração dos profissionais da educação infantil e ensino fundamental.

Além de estabelecer responsabilidades de acordo com o que foi determinado pelo FNDE, Fundeb e Lei de Responsabilidade Fiscal (LRF), o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) criou, em 2007, um indicador para mensurar o desempenho dos bens oferecidos pelos entes federados: o Ideb. Além de servir como meta para o FNDE, o indicador surge como o principal indicador da qualidade da educação básica no Brasil, em que seu resultado varia entre 0 e 10 para avaliar o desempenho das escolas. Para a construção do indicador, foram utilizadas variáveis de fluxo (representadas pela taxa de aprovação dos alunos) e aprendizado (representado pelo resultado dos estudantes no Sistema de Avaliação da Educação Básica – Saeb).

A combinação de ambos tem o mérito de equilibrar as duas dimensões: o fluxo irá pesar negativamente se um sistema de ensino reter seus alunos para obter resultados de melhor qualidade no Saeb ou Prova Brasil. Se o contrário acontecer (ou seja, aumentar o número de aprovações sem qualidade), o resultado das avaliações indicará igualmente a necessidade de melhoria do sistema.

Dessa forma, o objetivo da seguinte seção é analisar o desempenho educacional brasileiro, principalmente das competências dos estados (anos finais do ensino fundamental e ensino médio), comparando também os resultados brasileiros com resultados de outros países.

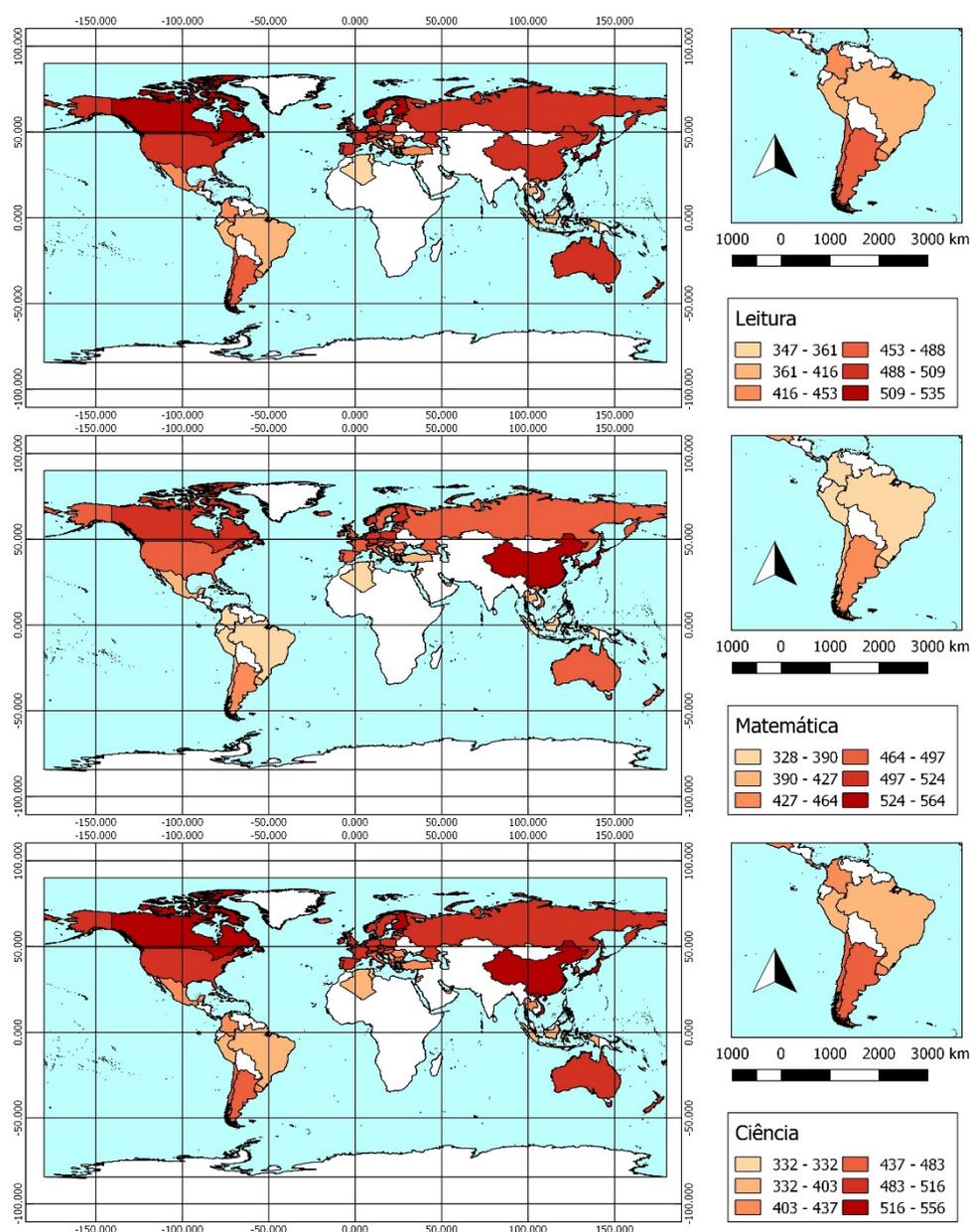
Análise geral

Apesar do aumento dos gastos nos períodos recentes (ver Quadro 1 e Gráfico 3), os dados do Inep e da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) mostram que, ainda que se tenha notado progresso, o Brasil apresenta resultados muito abaixo de países que se destacam. Desde os anos 2000, o Brasil participa da prova do Programa Internacional de Avaliação de Alunos (Pisa) (no original, Programme for International Student Assessment), aplicado e avaliado pela OCDE. O teste busca analisar o conhecimento dos alunos de 15 anos de idade.

De acordo com tais resultados (ver Figura 1), em 2015, o Brasil estava no nível de países com baixo índice de desenvolvimento e se manteve entre as piores posições, visto que os resultados para o país foram negativos em todas as modalidades – matemática, ciências e leitura. Cabe destacar que países vizinhos, como Argentina e Uruguai, alcançaram melhor desempenho comparados ao Brasil.

No geral, os melhores resultados do Pisa estão no Canadá, China e países europeus. Vale destacar também que a maioria dos países que não participa do programa está situada no continente africano. Fica evidente também que os países apresentam notas menores na prova de matemática, seguida de ciências. As provas de matemática tiveram tanto as menores notas como os piores resultados entre as maiores notas.

FIGURA 1 - Notas do Pisa - 2015



Fonte: Elaboração própria a partir dos dados do Pisa (2015)¹ do software QGIS.

Conforme Mincer (1974), ganhos de produtividade fazem com que a educação forneça retornos positivos para o mercado de trabalho. Portanto, um aumento do gasto público de um país com educação tende a ter como consequência indireta um aumento na produtividade da sociedade no futuro. Assim, é de se esperar que investimentos na educação básica gerem maior empregabilidade no futuro.

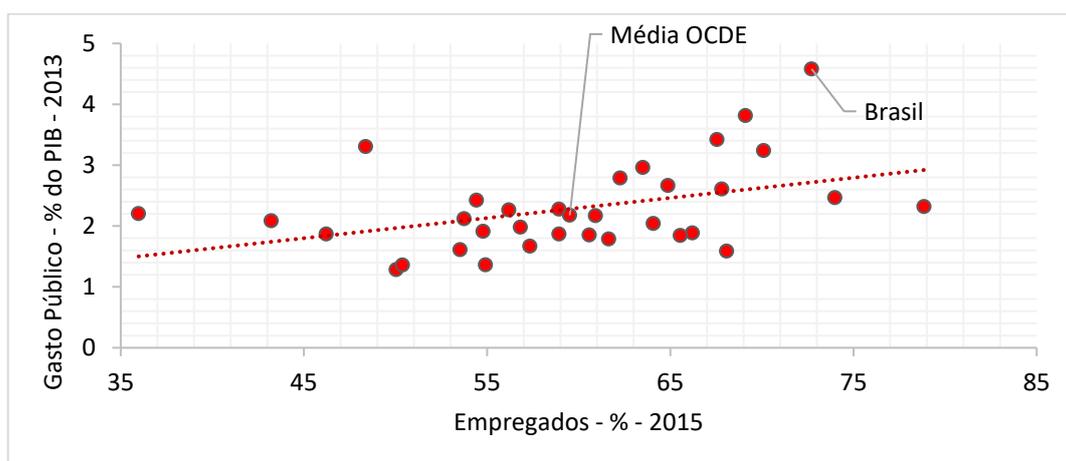
De acordo com Heiniger e Imdorf (2018), fatores relacionados à oferta educacional tendem a gerar melhores desempenhos educacionais e, como consequência, geram impacto positivo no mercado de trabalho. Ainda de acordo com os autores,

¹ Disponível em: <https://www.oecd.org/pisa/PISA-2015-Brazil-PRT.pdf>. Acesso em: dez. 2019.

países que investem mais em fatores educacionais relacionados à transmissão para o mercado de trabalho tendem a ser mais exitosos no número de indivíduos no mercado de trabalho.

Apesar de não especificar precisamente quais investimentos são realizados nessa transição, de acordo com o Gráfico 1, a relação entre gasto público no ensino fundamental – em % do Produto Interno Bruto (PIB) – e o número de empregados que possuem ensino fundamental é crescente. Com base nessa linha de tendência da OCDE, percebemos que o Brasil apresenta alto nível de indivíduos empregados que concluíram o ensino fundamental. Porém, o preço pago para atingir esse nível foi muito superior à tendência (média) observada. Ainda vale salientar que o Brasil possui baixa escolaridade média da população, que os indicadores educacionais, quando comparados com o restante da América Latina, só superam os de países mais pobres, e que ainda há diferenças significativas na educação entre estados e regiões do país (SILVA; HASENBALG, 2000).

GRÁFICO 1 – Relação entre gasto público no ensino fundamental como porcentagem do PIB (2013) e porcentagem do número de empregados com ensino fundamental (2015)



Fonte: Elaboração própria com dados da OCDE (2017).²

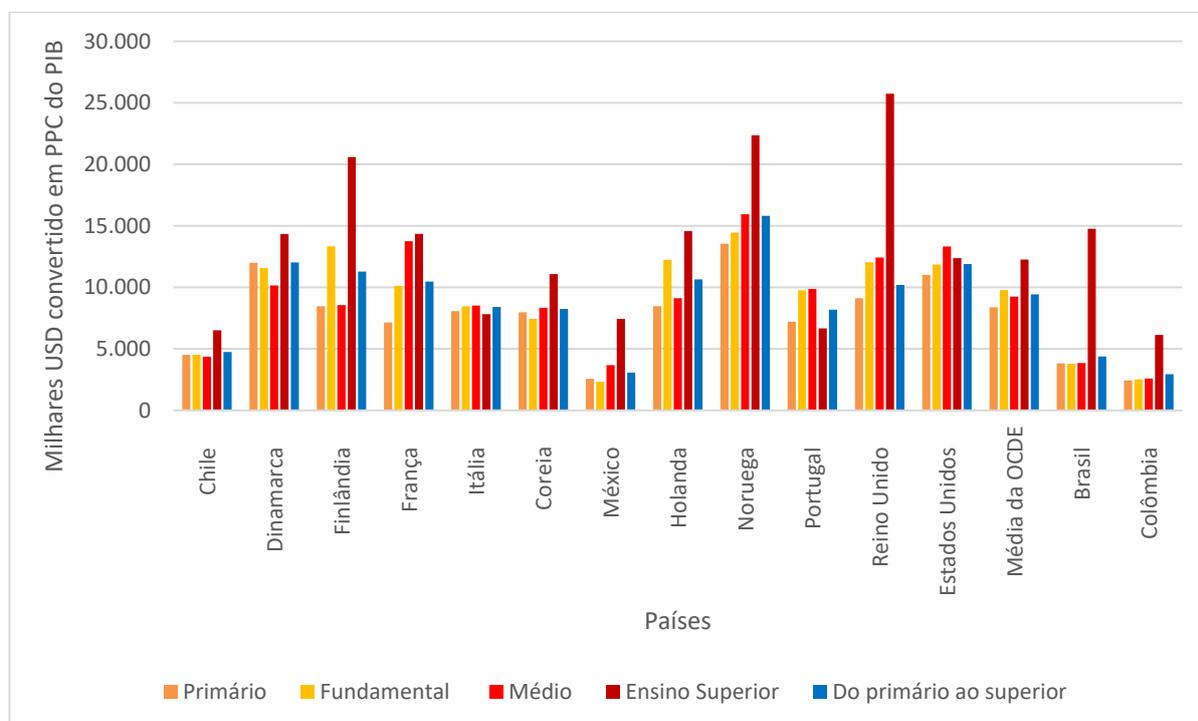
Em relação ao gasto por aluno, o Brasil apresenta valores inferiores ao de outros países e à média da OCDE (ver Gráfico 2). Quando se compara países próximos ao Brasil, como Chile e Colômbia, os níveis de gasto são semelhantes na educação básica, mas com diferenças nos resultados do Pisa (ver Figura 1). Discrepante ao restante dos níveis escolares, os resultados do Gráfico 2 mostram que, no Brasil, o gasto com ensino superior é maior que o com o restante dos níveis.

Para os casos dos Estados Unidos e do Reino Unido, o gasto com ensino superior é muito elevado por conta do elevado número de estrangeiros que iniciam

² Disponível em: https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS. Acesso em: dez. 2019.

sua formação nesses países, mas, para os países em desenvolvimento, percebe-se que não existe esse tipo de demanda. Mesmo assim, os gastos no ensino superior continuam em patamares elevados em relação ao restante dos níveis educacionais.

GRÁFICO 2 - Despesas anuais por aluno de instituições públicas de educação, segundo o país e o nível de ensino em 2013



Fonte: Elaboração própria com dados da OCDE (2017).³

Nota: Dólar americano equivalente convertido utilizando Paridade do Poder de Compra (PPC) para o PIB, por nível de educação, com base em equivalentes de tempo integral.

De acordo com os resultados do Inep, entre 1980 e 2015, houve grande crescimento no número de matrículas de ensino fundamental nos primeiros dez anos, seguido de uma leve queda e crescimento leve nos anos subsequentes, de forma que, em 2015, tenha atingido 96,5% da taxa de cobertura de matrícula.

Já o ensino médio apresentou baixas proporções no início dos anos 1980 (cerca de 14,3%), e foi crescendo consideravelmente ao longo dos anos (chegando a 56,9% em 2015). Dessa forma, os resultados mostram efeitos positivos das políticas de conscientização pública para a educação da população de manter os alunos na escola na idade padrão estabelecida pela rede de ensino.

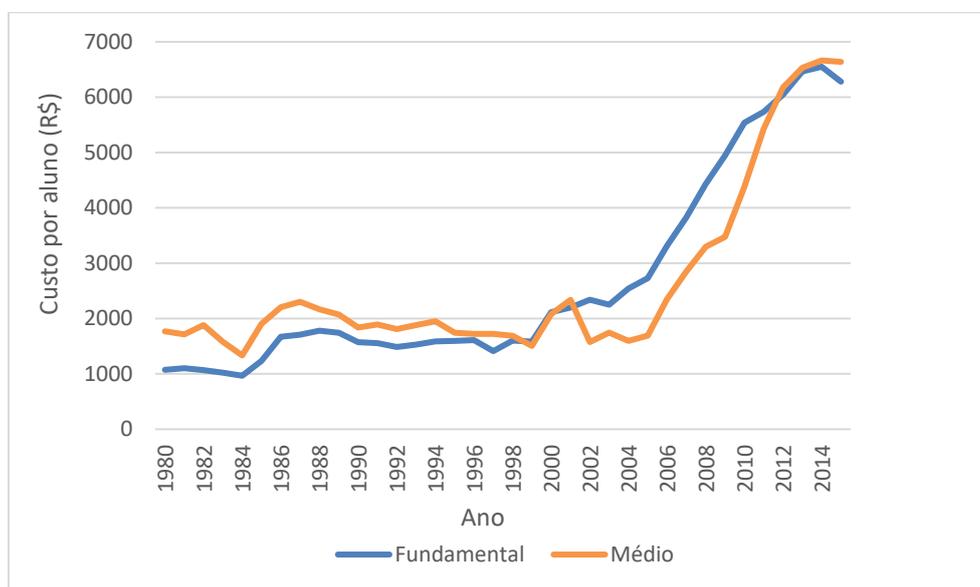
O Gráfico 3 apresenta os resultados do custo por aluno do ensino fundamental e médio apenas para o Brasil. Fica evidente que, por volta de 1984 a 1988, houve um aumento de todas as redes de ensino. Nos anos seguintes, o custo por aluno

³ Disponível em: https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS. Acesso em: dez. 2019.

para o ensino fundamental e médio manteve-se pouco variável até 2002, quando cresceu exponencialmente. Uma explicação lógica de tal comportamento consiste no grande investimento educacional durante o governo Lula em educação, em razão do *boom* das *commodities* brasileiras no exterior.

Apesar do crescimento dos gastos por aluno no ensino fundamental e médio (Gráfico 3), os dados do Inep mostram que o ensino superior se mantém com a maior proporção de gastos educacionais do governo (Gráfico 2), consumindo cerca de três vezes mais que o ensino básico. Heckman (2008) reforça o que já foi encontrado na literatura, afirmando que os anos básicos são os que proporcionam maior rentabilidade.

GRÁFICO 3 - Custo por aluno no Brasil, por nível de ensino (1980-2015)



Fonte: Elaboração própria com base em dados de Maduro Júnior (2007) de 1980 até 1999 e do Inep (BRASIL, 2018) de 2000 até 2015.

Análise estadual

As tabelas a seguir apresentam os resultados de âmbito estadual no período de 2007 a 2017, bem como uma breve análise. A Tabela 1 mostra que a média do número de alunos por turma diminuiu no período estudado na maioria dos estados, enquanto apenas o Mato Grosso do Sul (para os anos finais) e Minas Gerais (para os anos finais e ensino médio) apresentaram crescimento.

Uma redução no número de matrículas por turma pode ter efeitos negativos do ponto de vista de eficiência, uma vez que recursos como sala de aula e professor são destinados para menos pessoas, em média. Porém, do ponto de vista de resultados escolares, menos alunos por turma pode gerar um resultado positivo – pois professores teriam menor dificuldade, em turmas menores, de compartilhar conhecimento (DYNARSKI; HYMAN; SCHANZENBACH, 2013).

TABELA 1 – Média do número de alunos por turma nas escolas estaduais do Brasil

UNIDADE FEDERATIVA	ANOS FINAIS DO ENSINO FUNDAMENTAL						ENSINO MÉDIO					
	2007	2009	2011	2013	2015	2017	2007	2009	2011	2013	2015	2017
Acre	28.4	28.2	27.9	26.1	25.9	26.8	32.9	31.1	29.6	29.6	28.9	28.4
Alagoas	35.8	36.0	35.0	32.8	32.6	31.7	41.2	41.9	40.2	38.2	37.1	35.3
Amapá	29.1	29.0	30.1	27.3	25.3	25.7	33.1	32.8	32.6	29.1	29.1	29.1
Amazonas	34.1	35.3	34.6	34.4	32.9	32.8	33.8	33.0	31.5	30.9	29.4	29.8
Bahia	33.4	32.2	30.7	29.5	29.3	30.4	37.1	34.8	31.5	29.6	28.2	30.1
Brasil	31.2	30.7	29.8	28.8	27.9	28.0	34.2	32.9	31.5	30.5	29.8	29.8
Ceará	34.2	35.1	34.6	33.3	31.3	30.4	38.8	39.1	37.0	36.1	34.2	35.8
Distrito Federal	35.7	35.5	34.1	31.5	30.4	29.6	38.0	38.5	36.9	34.9	34.4	33.9
Espírito Santo	29.3	28.8	27.8	26.9	27.0	28.3	31.8	31.8	30.4	29.3	29.4	31.4
Goiás	31.3	29.5	28.2	28.0	28.8	28.0	33.8	31.5	28.7	28.5	28.7	27.9
Maranhão	33.4	32.3	31.8	31.9	30.3	29.6	37.9	35.5	34.0	34.1	33.4	32.7
Mato Grosso	31.1	30.3	29.7	29.9	30.1	23.6	33.9	33.5	33.0	32.5	33.0	24.7
Mato Grosso do Sul	28.3	29.3	29.1	29.6	28.9	29.8	29.3	30.0	29.5	29.5	28.8	29.2
Minas Gerais	27.5	26.7	26.5	25.5	24.7	29.1	29.8	28.1	28.9	28.1	26.8	32.6
Pará	39.4	33.9	32.5	31.1	30.0	30.3	39.8	34.6	32.8	31.8	31.6	32.0
Paraíba	31.3	29.4	28.4	27.0	25.1	26.1	35.1	32.1	30.8	28.5	26.3	27.9
Paraná	30.7	30.2	28.5	26.5	26.6	26.7	32.0	30.5	27.8	27.3	28.7	28.7
Pernambuco	38.2	36.2	35.1	34.6	33.0	34.3	40.1	35.0	35.7	35.7	34.5	35.9
Piauí	27.3	25.8	25.8	26.6	24.9	25.2	37.3	33.3	31.0	30.2	29.1	29.3
Rio de Janeiro	33.8	33.8	32.0	29.7	27.4	29.5	36.9	35.7	32.2	30.2	28.5	28.9
Rio Grande do Norte	29.8	29.9	28.5	28.8	28.8	28.6	37.5	36.1	34.7	33.5	32.6	33.2
Rio Grande do Sul	24.1	25.3	23.9	22.7	22.1	22.6	28.2	29.2	27.6	26.8	25.9	25.5
Rondônia	29.9	29.8	28.4	26.3	26.6	27.5	29.6	29.1	27.5	25.8	26.2	26.0
Roraima	22.5	23.0	23.5	22.0	21.8	21.4	23.3	23.8	24.3	22.4	22.1	21.8
Santa Catarina	26.6	26.8	25.5	25.4	24.5	24.9	28.8	28.8	25.8	26.5	26.4	25.7
São Paulo	35.4	34.3	33.7	32.0	30.2	30.3	36.6	35.5	35.4	34.3	32.7	33.4
Sergipe	33.3	32.3	31.1	28.8	27.1	25.4	37.4	34.7	33.6	33.0	32.5	30.8
Tocantins	28.8	30.1	27.8	28.9	28.2	26.3	29.5	28.8	28.5	28.1	26.5	25.0

Fonte: Elaboração própria com base nos dados fornecidos pelo Inep (BRASIL, 2018).

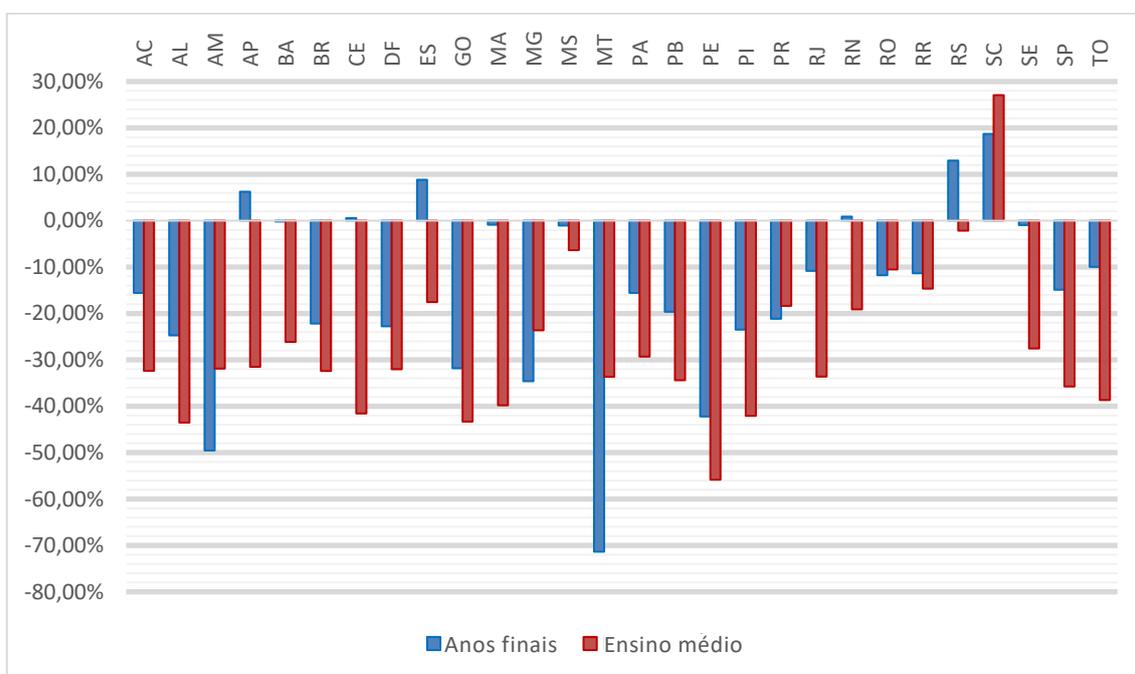
Souza, Oliveira e Annegues (2018) investigaram a relação entre as características das famílias e o desempenho escolar dos alunos de 5º e 9º ano do ensino fundamental das escolas públicas e privadas. Utilizando um modelo não paramétrico, o *generalized additive model* (GAM), um estimador em dois estágios, e partindo dos dados do Saeb de 2011 e do Censo Escolar, os autores encontraram evidências de que variáveis ligadas ao *background* familiar exercem grande influência no desempenho escolar dos alunos em comparação com as variáveis ligadas à escola, aos docentes e aos diretores, sobretudo para os alunos de escola pública.

O Gráfico 4 mostra que, no período de 2007 a 2017, houve uma queda geral na distorção idade-série nos anos finais do ensino fundamental, com destaque para o

Mato Grosso, que reduziu mais de 70% da distorção. Dentre os estados com aumento na proporção de alunos distorcidos, Ceará e Rio Grande do Norte tiveram uma variação menor que 1%; o crescimento no Amapá e Espírito Santo foi de 6,25% e 8,82%, respectivamente. Os estados que se destacam com acréscimo na distorção idade-série foram Rio Grande do Sul (12,94%) e Santa Catarina (18,69%).

Para o ensino médio, a maior parte dos estados apresentou decréscimo na distorção idade-série, com exceção de Santa Catarina, que registrou aumento de 27,05%. Cinco unidades federativas reduziram em pelo menos 40% a distorção no ensino médio entre 2007 e 2017: Alagoas, Ceará, Goiás, Pernambuco e Piauí, com destaque para Pernambuco, com uma queda de 55,82%.

GRÁFICO 4 - Taxa de crescimento da taxa de distorção idade-série, por nível de ensino (2007-2017)



Fonte: Elaboração própria, por meio dos dados do Inep (BRASIL, 2018).

Quando retratamos a situação do percentual de professores com ensino superior (Tabela 2), tanto nos anos finais da educação fundamental como no ensino médio, percebe-se que, entre 2007 e 2017, houve grande crescimento (com exceção de Ceará, Goiás, Santa Catarina e São Paulo), sendo a Bahia o estado que apresentou o maior crescimento na qualificação dos docentes para o ensino médio.

De forma geral, pode-se observar que o número de professores mais qualificados aumentou ao longo do tempo. Esse crescimento pode ser explicado pelo fato de que os governos em geral estão investindo mais em educação por meio da contratação de professores mais qualificados.

Esse resultado indica que, nesse aspecto, o país está no caminho certo; porém, é preciso reconhecer que ainda há uma grande concentração de alunos por professor em comparação a outros países, situação que se agrava quando considerada a inadequada estrutura da maioria das escolas públicas e que gera impactos negativos sobre o aprendizado e, conseqüentemente, sobre o resultado educacional.

TABELA 2 - Porcentagem de docentes com ensino superior das escolas estaduais do Brasil

UNIDADE FEDERATIVA	ANOS FINAIS DO ENSINO FUNDAMENTAL						ENSINO MÉDIO					
	2007	2009	2011	2013	2015	2017	2007	2009	2011	2013	2015	2017
Acre	45,0	51,2	69,0	73,6	72,3	72,7	92,8	89,6	87,8	89,8	86,9	87,1
Alagoas	62,2	67,0	83,5	79,8	85,1	91,9	89,2	86,4	84,1	77,2	84,8	91,6
Amapá	69,8	77,4	80,4	86,7	88,8	90,1	93,2	95,1	94,1	97,8	96,2	97,5
Amazonas	63,8	73,5	94,8	97,8	97,5	97,3	88,9	92,5	95,1	98,3	98,0	98,5
Bahia	39,3	49,7	73,8	83,9	90,4	94,7	72,9	74,7	71,0	79,1	85,0	91,3
Brasil	78,7	83,2	92,7	92,7	93,3	94,0	93,4	93,6	93,1	93,3	93,6	94,6
Ceará	78,7	79,7	91,1	84,5	85,1	85,1	95,5	91,0	90,0	87,2	87,8	89,9
Distrito Federal	96,9	95,5	99,0	99,1	98,4	98,0	94,4	95,1	99,6	99,5	99,0	99,2
Espírito Santo	78,9	90,4	94,9	98,0	97,7	98,3	88,6	94,1	94,0	98,1	97,5	97,8
Goiás	85,3	88,8	90,5	88,0	85,4	83,0	90,1	91,2	92,1	90,6	88,1	86,3
Maranhão	45,3	50,0	77,4	78,7	77,8	74,5	87,4	84,6	87,7	95,1	93,8	95,6
Mato Grosso	80,7	83,7	92,6	94,1	95,4	95,3	91,1	91,1	94,0	96,2	96,9	96,7
Mato Grosso do Sul	88,3	93,4	99,4	97,6	97,0	96,4	91,3	96,5	99,0	97,4	96,4	95,7
Minas Gerais	86,3	90,7	95,7	94,0	93,6	94,3	93,1	94,8	95,4	94,2	93,0	94,0
Pará	48,5	54,8	94,1	97,3	97,8	98,6	89,8	94,5	95,9	98,1	98,3	98,7
Paraíba	75,9	75,5	77,1	82,9	85,7	90,0	88,9	86,4	80,9	86,7	89,2	91,4
Paraná	96,5	97,2	96,2	95,8	97,8	98,9	98,3	98,2	96,1	96,0	97,2	98,2
Pernambuco	78,5	81,1	92,1	93,3	91,8	89,8	96,4	94,9	94,6	95,7	95,5	95,0
Piauí	67,6	67,7	81,4	85,0	90,6	88,7	93,4	88,2	85,1	86,3	91,3	88,1
Rio de Janeiro	92,6	93,9	95,3	97,9	97,1	97,6	98,2	97,5	96,9	99,1	97,9	97,9
Rio Grande do Norte	76,2	76,3	86,5	93,7	94,7	95,1	85,9	83,9	90,7	96,2	95,8	95,4
Rio Grande do Sul	86,7	91,3	93,1	92,8	92,6	91,9	92,8	95,3	94,1	94,1	93,7	93,6
Rondônia	75,1	85,2	92,5	94,0	93,9	94,8	94,2	95,6	94,7	96,4	96,6	97,6
Roraima	37,6	58,2	57,0	60,4	62,0	62,4	66,0	79,2	71,0	75,3	75,2	72,5
Santa Catarina	88,2	91,4	87,3	83,1	80,5	88,0	92,7	94,1	88,7	84,4	81,6	88,1
São Paulo	98,1	98,0	97,7	94,3	95,2	95,7	98,8	98,5	97,7	94,6	95,0	95,5
Sergipe	80,4	80,8	96,3	98,0	98,4	98,4	95,3	91,5	97,3	98,7	98,6	98,7
Tocantins	71,9	79,5	91,8	93,3	91,7	91,7	95,1	96,4	96,3	96,8	95,1	94,4

Fonte: Elaboração própria com base nos dados fornecidos pelo Inep (BRASIL, 2018).

Para Bertê, Borges e Brunet (2008), os gastos públicos com educação serão eficientes apenas se seu aumento estiver acompanhado de mais critério e qualidade, visto que melhorias na educação geram investimentos a longo prazo. Além disso,

os gastos devem proporcionar eficiência e equidade nos sistemas de ensino, focados no desempenho e no desenvolvimento dos alunos.

Segundo Heckman (2008), os gastos públicos com educação nada mais são que grandes investimentos que o Estado realiza para o desenvolvimento intelectual da sociedade. Quanto mais cedo se começa a investir na educação (principalmente entre as crianças oriundas de família de baixa renda), maior será o benefício. Guryan (2001) comprovou em seu estudo que os aumentos dos gastos públicos por aluno resultaram em uma significativa melhoria dos resultados educacionais.

Com relação aos estados brasileiros, de acordo com o Quadro 1, quando se compara o gasto público em educação ao longo do período de estudo (de 2007 a 2017), percebe-se que alguns estados aumentaram o gasto real em educação no período de dez anos, enquanto outros diminuíram – como é o caso do Rio de Janeiro. Salienta-se que, no período analisado, houve uma redução nas matrículas na rede estadual de ensino de todas as UF's do Brasil, com exceção do Acre, que aumentou em 11,7% (DAVIES; ALCÂNTARA, 2020).

QUADRO 1 – Gasto estadual total *per capita* com educação

ESTADO	2007	2017
Acre	290,19	1425,54
Alagoas	1201,36	209,63
Amapá	525,40	1125,34
Amazonas	281,75	672,58
Bahia	427,15	320,41
Ceará	879,33	330,23
Distrito Federal	464,68	1423,25
Espírito Santo	482,56	295,10
Goiás	238,10	737,31
Maranhão	2516,75	367,04
Mato Grosso	513,16	3122,75
Mato Grosso do Sul	81,22	686,79
Minas Gerais	266,62	122,31
Pará	343,96	383,57
Paraíba	687,68	512,93
Paraná	263,10	960,21
Pernambuco	343,24	311,09
Piauí	640,51	490,80
Rio de Janeiro	480,43	361,25
Rio Grande do Norte	332,81	386,66
Rio Grande do Sul	559,28	351,19
Rondônia	1488,52	621,59
Roraima	403,20	1185,52
Santa Catarina	804,66	434,86

(continua)

(continuação)

ESTADO	2007	2017
São Paulo	475,43	721,11
Sergipe	661,74	390,90
Tocantins	290,19	743,69

Fonte: Elaboração própria com dados da Secretaria do Tesouro Nacional (BRASIL, 2017).

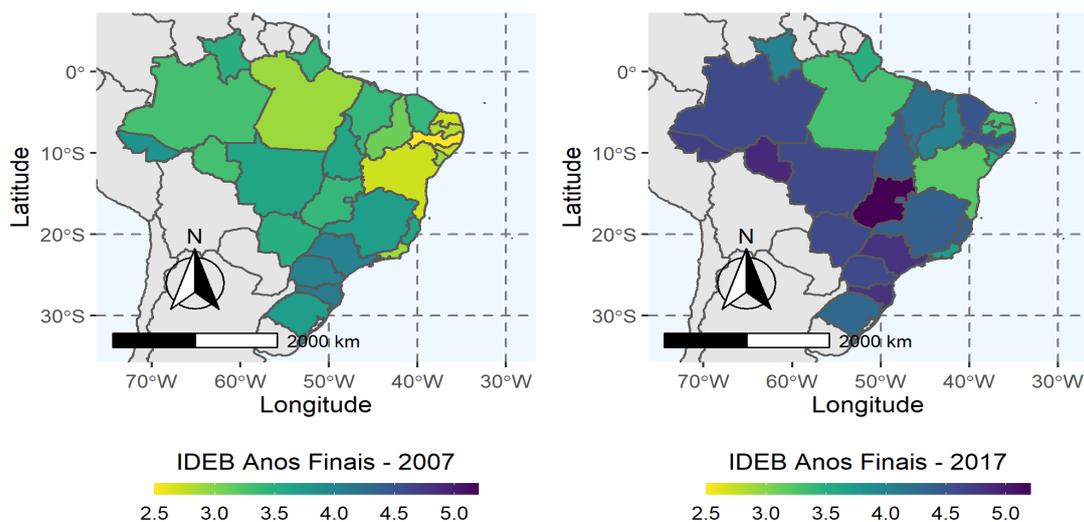
Nota: Em valores de 2017.

Sousa *et al.* (2018) realizaram seu estudo voltado para os 56 municípios do sertão pernambucano, encontrando que cidades que apresentaram menos gastos médios com alunos conseguiram obter êxito nas metas do Ideb. Ferreira *et al.* (2018) procuraram analisar como os gestores apontam o Ideb nas escolas públicas da cidade de Salgueiro, em Pernambuco, encontrando divergência por meio dos resultados do indicador e de entrevista com os gestores, e que estes veem o índice como insuficiente para representar e subsidiar as práticas das instituições, mas que é um importante instrumento de monitoramento e avaliação da rede de ensino.

A Figura 2 revela que em 2007 pelo menos seis estados nordestinos apresentaram Ideb inferior ou igual a 3,0, enquanto dois estados do Sul e um estado do Sudeste apresentavam Ideb de pelo menos 4,0 – superando inclusive a média nacional do ensino fundamental de 3,5. Já em 2017, houve um grande aumento do Ideb em todos os estados brasileiros, com destaque para as regiões Norte e Nordeste. Na Região Sudeste, apenas o Rio de Janeiro apresentou pequeno crescimento no Ideb. Apesar do resultado positivo, observa-se que, para os anos finais do ensino fundamental, os resultados ainda precisam melhorar, já que o estado com melhor resultado obteve pontuação de 5,2, mostrando que as UFs estão longe da meta número 7⁴ definida pelo Plano Nacional de Educação (PNE).

4 De acordo com a Meta 7 do PNE (BRASIL, 2001) é importante que seja fornecida educação básica de qualidade em todas as etapas e modalidades, com melhoria do fluxo escolar e da aprendizagem, de modo a atingir as seguintes médias nacionais para o Ideb: 6,0 nos anos iniciais do ensino fundamental; 5,5 nos anos finais do ensino fundamental; 5,2 no ensino médio.

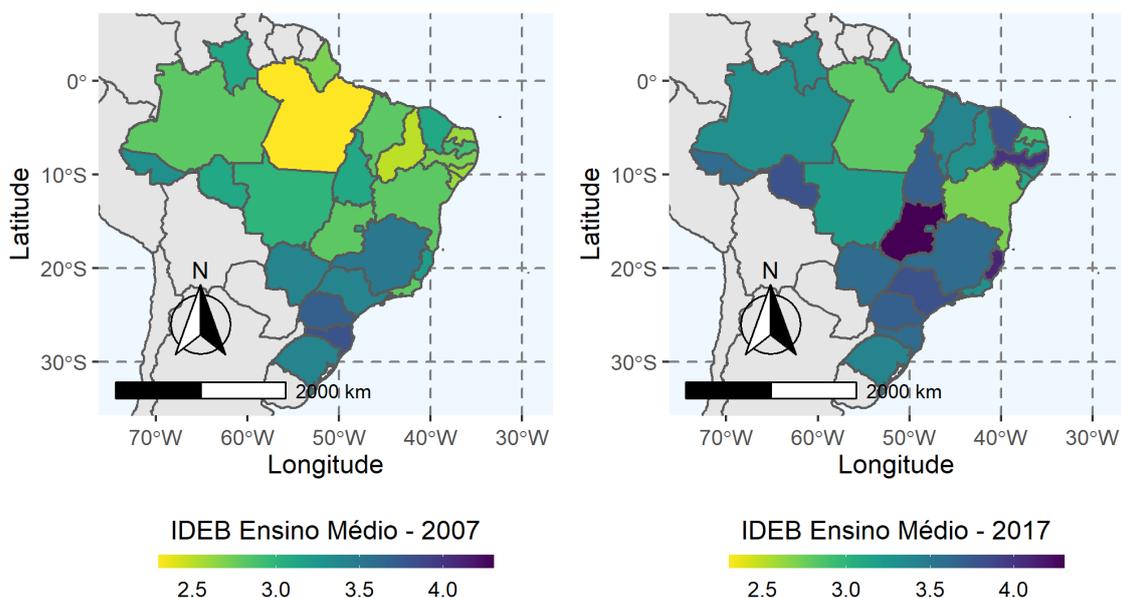
FIGURA 2 - Mapas do Ideb (anos finais do ensino fundamental) - 2007 e 2017



Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2018).

Em relação ao ensino médio, a Figura 3 mostra um aumento significativo no indicador Ideb para o ensino médio; todavia nenhum estado conseguiu atingir o valor 5 em 2017. Estados das regiões Norte e Nordeste, entre eles Ceará e Pernambuco, apresentaram crescimento, e a Bahia apresentou queda. Assim como acontece nos anos finais do ensino fundamental, o estado de Goiás apresentou o maior Ideb para o ensino médio: 4,3. Os demais estados com nota 4 são Espírito Santo (4,1) e Pernambuco (4,0).

FIGURA 3 - Mapas do Ideb (ensino médio) - 2007 e 2017

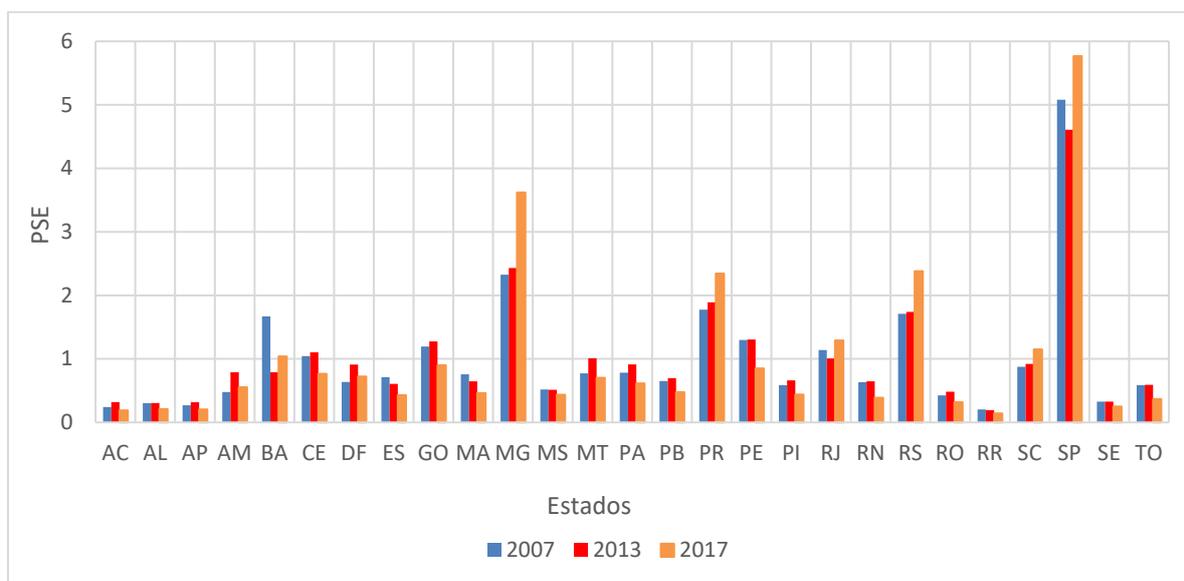


Fonte: Elaboração própria com base nos dados do Inep (BRASIL, 2018).

Card e Krueger (1996), Deon Filmer e Lant Pritchett (1999) apontam que a infraestrutura do sistema educacional é um requisito importante para o desenvolvimento do sistema e que investimentos em equipamentos geram retornos significativos para a educação. Além disso, Libâneo (2008) afirma que a infraestrutura deve ser adequada, de forma que assegure o desenvolvimento do sistema educacional. De mesmo pensamento, Marzocchi e Oliveira (2009) indicam que equipamentos e recursos educacionais melhores podem gerar, em média, resultados de desempenho 6,9% maiores para os alunos participantes do Pisa.

O Gráfico 5 mostra os resultados obtidos a partir do indicador de estrutura criado como objeto de pesquisa para a infraestrutura das escolas, determinado para 2007, 2013 e 2017 para os estados brasileiros. Os resultados mostram sentido diferente do que foi encontrado até o presente momento no quesito desempenho. Destacam-se os estados de São Paulo, seguido de Minas Gerais, Paraná e Rio Grande do Sul, entre os estados com melhores desempenhos em estrutura escolar dentre os comparados.

GRÁFICO 5 - Indicador de estrutura (PSE) para os estados brasileiros



Fonte: Elaboração própria, por meio dos dados do Inep (BRASIL, 2018).

ANÁLISE EMPÍRICA

O modelo

A partir dos dados do Censo Escolar do Inep, é possível analisar algumas informações educacionais básicas do Brasil. A variável dependente é o Ideb. Como variáveis explicativas, também oriundas do Inep, tem-se: a média do número de alunos por turma, a distorção idade-série, taxa de reprovação, de abandono, total

de matrículas, a meta do Ideb estipulada e a estrutura escolar – esta é formada pela média entre as variáveis ponderadas do número de computadores para alunos, número de salas e o número de escolas com quadra e biblioteca. A ponderação é criada com base em Afonso, Schuknecht e Tanzi (2005), que pondera as variáveis pela média geral de cada ano.

Uma variável de gasto foi inserida com informações da Secretaria do Tesouro Nacional (BRASIL, 2017) e do Inep (BRASIL, 2018). Com base em dados coligidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2017) foram inseridas variáveis de renda *per capita*, população, taxa de homicídio e expectativa de vida. A região de pertencimento da UF também foi adicionada como controle por meio de variáveis binárias (*dummies*).

A estratégia empírica parte de um modelo linear para a fronteira estocástica. Com o objetivo de mensurar a eficiência dos estados brasileiros mediante seus desempenhos na nota do Ideb, o modelo pode ser descrito da seguinte forma:

$$IDEB_{it} = X_{ti}\beta + Y_{ti}\alpha + v_{it} - u_{it} \quad (8),$$

em que Ideb é produto (*output*); β representa os vetores ($k \times 1$) das variáveis de interesse a serem estimados; e α representa os vetores ($k \times 1$) das variáveis socioeconômicas, ambos para os estados i no tempo t ; $v_{it} \sim N(0, \sigma_v^2)$ representa o componente do erro aleatório; e $u_{it} \sim N(\mu, \sigma^2)$ representa o termo de erro aleatório não observável em relação ao termo de ineficiência técnica.

De acordo com a Tabela 3, em média, os estados mostram alguns resultados negativos, como, por exemplo, o de que o Ideb está distante de atingir a meta número 7 definida pelo PNE (SOARES; XAVIER, 2013). Em média, pode-se observar também que o resultado do Ideb está atrás da meta esperada para o período de estudo – a qual foi definida a partir do Ideb escolar de 2005, e inclui também a meta de convergência das notas em 9,9 para 2096 e o esforço necessário para se alcançar tal meta (FERNANDES, 2007) –, e precisa ser considerado com outros fatores contextuais (MATOS; RODRIGUES, 2016). Comparando as redes de ensino, percebe-se que a taxa de alunos reprovados é maior no ensino fundamental, enquanto o ensino médio tem maior taxa de abandono, maior taxa de distorção idade-série e maior média do número de alunos por turma. Por outro lado, o ensino médio possui maior número de professores com ensino superior, enquanto o Ideb dos anos finais do ensino fundamental é maior.

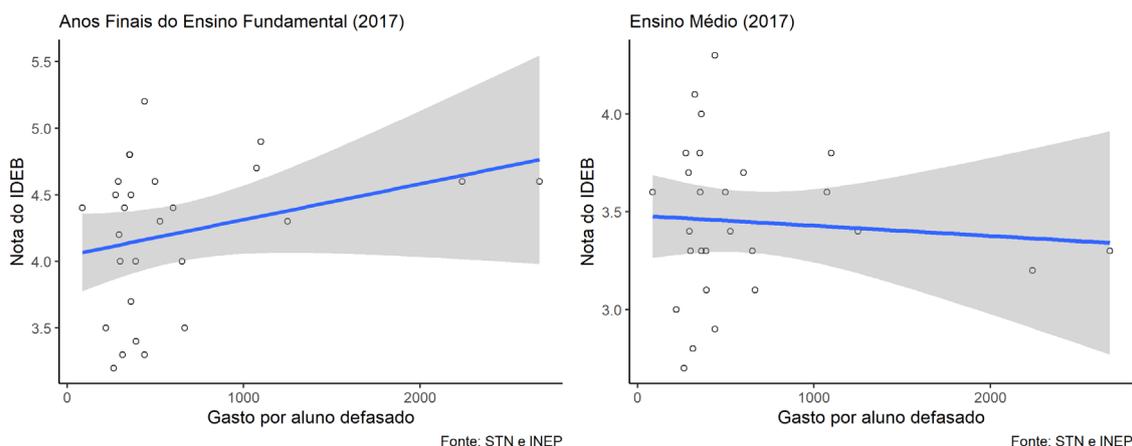
TABELA 3 – Estatística descritiva das variáveis utilizadas no modelo

VARIÁVEIS	MÉDIA	DESVIO PADRÃO	MÍNIMO	MÁXIMO
ANOS FINAIS DO ENSINO FUNDAMENTAL				
Ideb	3,72	0,58	2,50	5,20
Meta Ideb	3,80	0,69	2,40	5,50
Média de alunos por turma	29,4	3,58	21,4	39,4
Taxa de distorção idade-série	34,92	10,35	6,5	57,4
Reprovação	13,5	5,88	1,2	28,4
Abandono	5,55	3,67	1,00	18,9
Docentes com ensino superior	79,00	15,00	38,00	99,00
Matrícula	235.348	341.137	25.132	1.897.636
ENSINO MÉDIO				
Ideb	3,24	0,38	2,30	4,30
Meta Ideb	3,42	0,57	2,30	4,80
Média de alunos por turma	31,47	4,07	21,8	41,9
Taxa de distorção idade-série	40,26	12,47	14,7	75,8
Reprovação	12,69	4,17	5,1	25,0
Abandono	12,25	5,21	1,5	27,4
Docentes com ensino superior	91,65	6,9	66,01	99,6
Matrícula	343.164	450.027	14.530	3.159.247
CONTROLE				
Gasto em educação defasado	2.430.000.000	4.290.000.000	204.000.000	30.400.000.000
Taxa de homicídio	32,36	12,78	10,4	71,4
População	7.323.833	8.600.084	421.499	45.100.000
Renda <i>per capita</i>	767,28	395,5	1,01	2441,0
Expectativa de vida	73,34	2,61	68,10	79,40
Estrutura	1,00	1,00	0,14	5,77

Fonte: Elaboração própria com base nos dados da pesquisa.

Dentre as variáveis que afetam os resultados educacionais, o gasto público aparece como uma variável importante para detectar possíveis ganhos do desempenho, já que boas medidas administrativas e a adoção de políticas eficientes resultam em melhoria do sistema público do país (HANUSHEK, 1995; AFONSO; SCHUKNECHT; TANZI, 2010). Na Figura 4, é possível observar a relação entre o gasto defasado dos estados em três anos e as notas do Ideb para os anos finais e o ensino médio em 2017, indicando que os anos finais do ensino fundamental apresentam tendência positiva, enquanto o ensino médio apresenta uma tendência pouco inclinada, mas negativa. Esse resultado, contudo, precisa de maiores evidências empíricas, uma vez que outras variáveis também podem estar afetando os desempenhos educacionais, e a inclusão delas é necessária para uma análise mais robusta.

FIGURA 4 - Relação do gasto público defasado e a nota do Ideb (2017)



Fonte: Elaboração própria por meio dos dados da pesquisa.

Resultados econométricos

Neste trabalho, foram realizadas estimações para as 27 UFs nos anos de 2007, 2009, 2011, 2013, 2015 e 2017. Em virtude da composição dos dados, os autores preferiram adotar um modelo de estimação de dados em painel, e assim verificar o impacto das variáveis explicativas sobre o Ideb (variável dependente).

Na Tabela 4, tem-se o resultado do teste de razão de verossimilhança. Tanto para o ensino médio quanto para o fundamental não é possível rejeitar a hipótese nula; logo, o modelo mais adequado é o de Cobb-Douglas.

TABELA 4 - Resultados do teste de razão de verossimilhança para as formas funcionais

TESTE DE HIPÓTESE	ENSINO FUNDAMENTAL	ENSINO MÉDIO
H0: Cobb-Douglas H1: Translog	Lr = 8,88 P = 0,1802	Lr = 7,62 P = 0,1788

Fonte: Elaboração própria por meio dos dados da pesquisa.

A Tabela 5 apresenta os resultados da aplicação dos modelos para o ensino fundamental e ensino médio. Cabe salientar que, para o ensino fundamental, não tem o modelo variante no tempo, pois não converge no *software*. Pelos resultados obtidos, é possível afirmar que o modelo de ineficiência é adequado, haja vista o valor de gama (que mede a participação do componente da ineficiência) ser bastante expressivo. A ausência de ineficiência também pode ser rejeitada, uma vez que, em todos os modelos, o valor log verossimilhança é maior do que o valor tabelado. Os modelos também são válidos, uma vez que o teste de Wald indica a significância geral dos parâmetros estimados em 1%.

TABELA 5 – Resultado das estimações

TEMPO	ENSINO FUNDAMENTAL		ENSINO MÉDIO
	INVARIANTE	VARIANTE	INVARIANTE
Ln (gastos defasados)	0,2557***	-0,1003	-0,1368**
Meta Ideb	0,0198	0,0536	0,0936
Alunos por turma	-0,017*	-0,0282***	-0,0186**
Taxa de distorção idade-série	-0,0055	-0,0082***	-0,009***
Reprovação	-0,0367***	-0,0063	-0,0086
Abandono	-0,0578***	-0,0287***	-0,0273***
Docentes com ensino superior	0,0396	0,0084**	0,008**
Ln (Mat)	-0,0095	0,0771***	0,074***
Estrutura	-0,0212	-0,0076*	-0,0075
Taxa de homicídio	0,0036	0,0023	0,0021
Ln (população)	-0,1703	0,0038	-0,0139
Ln (renda <i>per capita</i>)	-0,0206	-0,0408***	-0,0366***
Anos esperados	0,0176	0,0319***	0,0158
Constante	1,263	3,0540***	4,7928***
<i>Dummies</i> de região	Sim	Sim	Sim
Lnsigma2	-2,5572***	-2,4438***	-2,8848***
Gama	0,1378	0,488	0,1271
Eta		-0,1106**	
Log Verossimilhança	-13,8136	8,5303	4,3862
Chi ²	524,53	176,60	205,81
Prob > chi2	0,0000	0,0000	0,0000

Fonte: Elaboração própria por meio dos dados da pesquisa com o uso do *software* Stata versão 15.

Nota: Nível de significância do coeficiente: * 10%, ** 5% e *** 1%.

Para o ensino fundamental, é possível perceber que os gastos defasados aumentam a eficiência das UFs brasileiras, a reprovação e o abandono geram gastos que acabam não dando retorno. De acordo com Bissoli (2010), um aluno que abandona a escola antes de concluir uma série ou determinado nível resulta em oneração dos recursos destinados ao investimento educacional, reduzindo sua eficiência. Macgregor (2007) afirma que as principais causas da evasão são a falta de qualidade das escolas, motivos de dificuldade financeira dos alunos e a baixa escolaridade e renda de suas respectivas famílias.

Os resultados da Tabela 5 mostram que, quanto maior for o número de alunos por turma, menor será a eficiência do gasto. Ainda que possa ser estratégia de governo turmas maiores, com o intuito de reduzir o gasto em educação, tal política mostra-se ineficiente, de acordo com os resultados aqui apresentados. A variável tamanho de turma, porém, é controversa, uma vez que parte da literatura aponta que turmas menores são importantes para melhorar a proficiência do escolar (JEPSEN; RIVKIN, 2009), mas também mostra que tal efeito é nulo ou até negativo sobre as notas dos alunos (HOXBY, 2000; LEUVEN; LØKKEN, 2018; SOUZA; OLIVEIRA; ANNEGUES, 2018).

Para o ensino médio, também de acordo com a Tabela 5, conforme supracitado sobre o valor de gama, o modelo mais adequado é o modelo variante no tempo, por isso a análise será focada na coluna do modelo variante. Novamente, o número de alunos por turma afeta negativamente a eficiência, assim como a distorção idade-série e o abandono escolar. A estrutura escolar mostra que as UFs que investem muito em estrutura têm menor eficiência, assim como as UFs com maior renda *per capita* – tal resultado também é encontrado por Frio *et al.* (2018), que mostra que escolas com maior infraestrutura possuem, em média, menos eficiência; porém, o nível socioeconômico não afeta a eficiência. Por outro lado, as UFs com maior expectativa de vida possuem maior eficiência, e a eficiência do gasto também é aumentada de acordo com a nota de matemática e com a porcentagem de docentes com ensino superior.

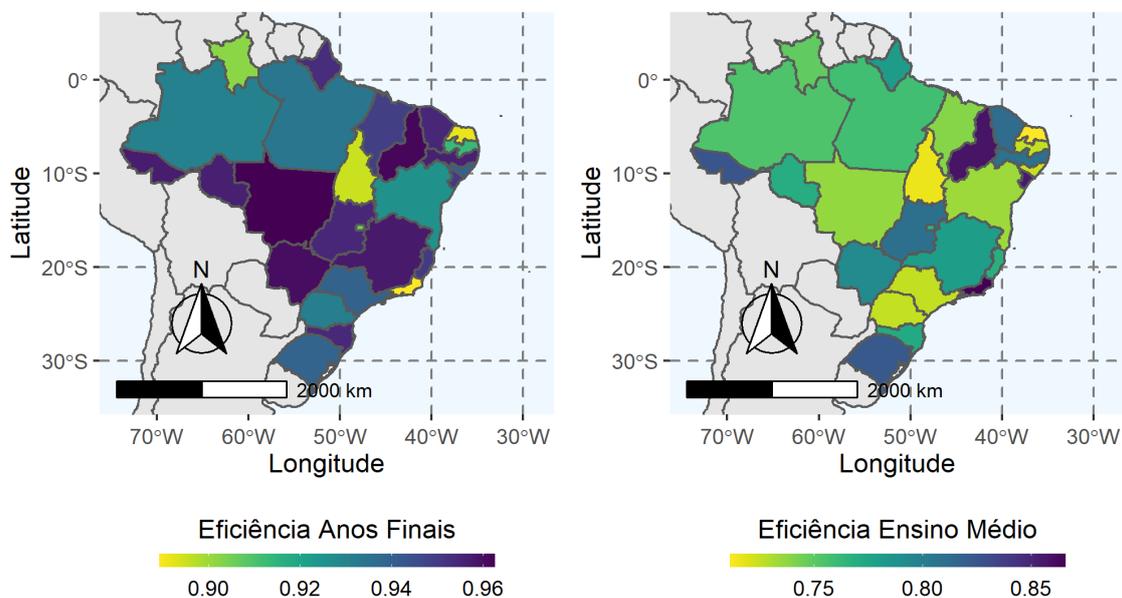
Jacob *et al.* (2018) mostram que a seleção de professores tem papel relevante nas notas dos alunos. No processo de seleção, a maior formação dos professores afeta positivamente as notas; ou seja, a educação do professor é um bom preditor de sua qualidade. O resultado que encontramos tem o mesmo sentido, mostrando que ter mais professores com ensino superior aumenta a eficiência do gasto em educação no ensino médio. Cowan e Goldhaber (2016) mostram que o programa de certificação de professores no estado de Washington, nos Estados Unidos, é eficiente no aumento das notas dos alunos, tanto em leitura como em matemática.

Ainda que nos resultados aqui apresentados não haja efeito no tamanho da população, resultado que corrobora os achados de Gonçalves e França (2013) para os municípios do Brasil, segundo Scarpin *et al.* (2012), as cidades catarinenses mais populosas são, em média, menos eficientes, resultado que também é encontrado por Macêdo *et al.* (2012) para os municípios paranaenses.

A taxa de homicídios não apresenta efeito sobre a eficiência, ainda que se encontrem na literatura trabalhos que mostrem o efeito da violência sobre os resultados escolares. De acordo com o trabalho de Monteiro e Rocha (2017) para o Rio de Janeiro, alunos que foram expostos a tiroteios ou guerras entre facções obtêm pior desempenho em matemática na Prova Brasil.

A Figura 5 mostra que existem dois cenários bastante distintos na eficiência: enquanto os estados possuem uma eficiência relativa considerada alta nos anos finais do ensino fundamental (o valor mínimo é de 0,889, em uma escala que varia de 0 a 1), no ensino médio o termo de ineficiência do modelo invariante no tempo é bastante elevado, em que a maior eficiência registrada é de 0,865.

FIGURA 5 - Mapas da fronteira de eficiência (anos finais do ensino fundamental e ensino médio) - invariante no tempo



Fonte: Elaboração própria.

A Tabela 6 apresenta a eficiência técnica estimada pela aplicação do modelo de fronteira de produção estocástica, estimado na equação variante no tempo para o ensino médio da tabela anterior. A Tabela 6 apresenta os valores da eficiência de cada estado (2007 e 2017) e seu posicionamento no *ranking* de 2017, nela é possível perceber que os estados se mantiveram no *ranking* de eficiência da provisão de educação para o ensino médio, mas que houve uma queda expressiva das eficiências relativas ao longo do tempo.

TABELA 6 - Eficiência técnica para o ensino médio dos estados brasileiros - 2007 e 2017

UF	EFICIÊNCIA 2007	EFICIÊNCIA 2017	RANKING	VARIAÇÃO EFICIÊNCIA (2007/2017)
Acre	0.956	0.875	4º	-8.45%
Alagoas	0.894	0.715	19º	-19.99%
Amapá	0.933	0.814	9º	-12.77%
Amazonas	0.887	0.698	21º	-21.30%
Bahia	0.822	0.555	27º	-32.49%
Ceará	0.946	0.847	6º	-10.41%
Distrito Federal	0.911	0.759	13º	-16.76%
Espírito Santo	0.933	0.815	8º	-12.69%
Goiás	0.965	0.900	3º	-6.72%
Maranhão	0.882	0.687	22º	-22.08%
Mato Grosso	0.891	0.708	20º	-20.51%
Mato Grosso do Sul	0.936	0.823	7º	-12.15%
Minas Gerais	0.910	0.755	14º	-17.06%

(continua)

(continuação)

UF	EFICIÊNCIA 2007	EFICIÊNCIA 2017	RANKING	VARIAÇÃO EFICIÊNCIA (2007/2017)
Pará	0.896	0.721	17º	-19.54%
Paraíba	0.854	0.624	24º	-26.96%
Paraná	0.857	0.629	23º	-26.57%
Pernambuco	0.932	0.812	10º	-12.89%
Piauí	0.978	0.935	1º	-4.36%
Rio de Janeiro	0.951	0.861	5º	-9.46%
Rio Grande do Norte	0.836	0.585	26º	-30.08%
Rio Grande do Sul	0.925	0.794	11º	-14.23%
Rondônia	0.917	0.772	12º	-15.75%
Roraima	0.903	0.738	15º	-18.30%
Santa Catarina	0.901	0.732	16º	-18.75%
São Paulo	0.844	0.602	25º	-28.69%
Sergipe	0.974	0.923	2º	-5.16%
Tocantins	0.896	0.721	18º	-19.57%

Fonte: Elaboração própria.

Com base na Tabela 6, ainda é possível analisar por região do Brasil. Enquanto no Nordeste estão os dois estados mais eficientes, Piauí e Sergipe, também estão os estados que mais desperdiçam, o Rio Grande do Norte e a Bahia, mostrando uma disparidade regional bem acentuada. No Norte, o destaque positivo é o Acre, que ocupa a quarta colocação, porém o Amazonas ocupa a 21ª posição. No Centro-Oeste, Goiás está em terceiro, e Mato Grosso do Sul, em sétimo, mas destaca-se negativamente a vigésima colocação, ocupada por Mato Grosso. A disparidade regional do Sudeste se dá por São Paulo e Rio de Janeiro: enquanto aquele ocupa a 25ª posição, este é o 5º colocado. Por fim, no Sul estão o Rio Grande do Sul e Santa Catarina como estados de eficiência média – 11º e 16º, respectivamente –, enquanto o Paraná ocupa a 23ª colocação.

Ou seja, o cenário educacional brasileiro sofreu modificações nos anos analisados, porém ainda apresenta notórias disparidades nas regiões e estados. Essas mudanças representam apenas o início da grande evolução desejada pelos formuladores de políticas e pela sociedade e de que a educação brasileira tanto precisa.

CONSIDERAÇÕES FINAIS

Ao longo do período analisado, o sistema educacional brasileiro sofreu mudanças em suas políticas e em sua atuação. Vários investimentos foram feitos, percebidos por meio do aumento no número de professores qualificados, combate à evasão e melhora da estrutura escolar. Porém, há de se considerar que nosso sistema educacional ainda possui muitas metas a conquistar para chegar ao nível educacional adequado.

Em relação aos indicadores de qualidade de educação, pode-se constatar, pela análise estatística, que a média de alunos por turma e a distorção idade-série têm diminuído ao longo do tempo para a maioria dos estados, enquanto o percentual de professores com ensino superior aumentou na maioria dos estados, bem como o gasto público em educação. Apesar de os estados das regiões Norte e Nordeste apresentarem as maiores melhoras entre 2007 e 2017, seus resultados ainda se mostram críticos.

Destacamos que o Ideb cumpre um papel importante no monitoramento da qualidade da educação e como guia para o direcionamento da política educacional do país. Entretanto, mostramos que o simples monitoramento do Ideb não é suficiente para mensurar a eficiência dos resultados da política educacional. Como evidenciado pelos resultados obtidos pelo estado de São Paulo (quarto do *ranking* do Ideb e 25º no *ranking* de eficiência), a simples observação do desempenho pode mascarar falhas significativas na política educacional, como, por exemplo, a má alocação de recursos. Dessa forma, acreditamos que a política educacional deve ser avaliada de forma multidimensional, com critérios e indicadores que vão além da sinalização de desempenho dada pelo Ideb.

Com base nesses primeiros resultados, é possível perceber o investimento que está sendo feito no sistema educacional pelo governo, ao passo que ainda se identificam falhas em suas políticas educacionais. Com base nos resultados econométricos, é possível realizar algumas conjecturas. Dentre elas:

- a) Apesar do forte crescimento do gasto público estadual estar associado a melhorias no desempenho educacional dos anos finais do ensino fundamental, os resultados mostram evidências de desperdício (ou ineficiência) na aplicação de recursos no ensino médio.
- b) Os resultados também mostram o relevante papel da qualificação dos professores, do combate à evasão escolar e do tamanho da turma. Esse resultado é importante para balizar o direcionamento da aplicação dos recursos.
- c) Os resultados também mostram a necessidade de uma análise de longo prazo, observando as mudanças estruturais na educação. Fatores como a mudança no *background* escolar impactam fortemente a educação e só podem ser observados de forma consistente num longo espaço de tempo.

O tema em questão é de extrema importância, já que a educação é considerada uma das bases sobre as quais a política, a economia e a sociedade se assentam. Conforme apontado neste trabalho, o sistema educacional brasileiro ainda tem um longo trajeto rumo à qualidade, proporcionando educação pública e eficiente tal como a de países desenvolvidos. Assim, o presente trabalho tem o papel de auxiliar futuras pesquisas e formuladores de política em geral sobre quais políticas

educacionais devem ser adotadas nos estados e regiões, além de revelar quais indicadores interferem no desenvolvimento da educação. Sugere-se como avanços para a literatura estudos causais de como o aumento do gasto em educação e o investimento em infraestrutura e corpo docente afetam a eficácia do gasto, uma vez que se decidiu recentemente pela promulgação da Proposta de Emenda Constitucional 26 de 2020, que define o Fundeb como permanente (AGÊNCIA SENADO, 2020), trazendo consigo a necessidade de avaliações periódicas da eficiência do gasto público em educação para que tal investimento seja alocado da melhor maneira possível.

Ainda é possível compreender que a ação do poder público, sozinha, é insuficiente para resolver um problema de tal envergadura e com um passivo histórico de tão grandes proporções. Como bem aponta o movimento “Todos pela Educação” (BRASIL, 2006), somente o envolvimento e a participação de diversos segmentos da sociedade – engajados na obtenção das mesmas metas e alinhados com as diretrizes das políticas públicas educacionais – poderão encontrar as melhores soluções e as condições efetivas para que sejam implementadas. Logo, o envolvimento e o compromisso de toda a sociedade são fundamentais para promover o salto de qualidade de que a educação básica brasileira necessita.

REFERÊNCIAS

- AFONSO, António; AUBYN, Miguel St. Non-parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics*, v. 8, n. 2, p. 227-246, Nov. 2005.
- AFONSO, António; FERNANDES, Sónia. Assessing and explaining the relative efficiency of local government. *The Journal of Socio-Economics*, v. 37, n. 5, p. 1946-1979, Oct. 2008.
- AFONSO, António; ROMERO, Alma; MONSALVE, Emma. *Public sector efficiency: evidence for Latin America*. [S.l.]: Banco Interamericano de Desenvolvimento, 2013.
- AFONSO, António; SCHUKNECHT, Ludger; TANZI, Vito. *Public sector efficiency: an international comparison*. Frankfurt: European Central Bank, 2003.
- AFONSO, António; SCHUKNECHT, Ludger; TANZI, Vito. Public sector efficiency: an international comparison. *Public Choice*, v. 123, n. 3-4, p. 321-347, June 2005.
- AFONSO, António; SCHUKNECHT, Ludger; TANZI, Vito. Public sector efficiency: evidence for new EU member states and emerging markets. *Applied Economics*, v. 42, n. 17, p. 2147-2164, Feb. 2010.
- AGÊNCIA SENADO. Emenda Constitucional que garante a permanência do Fundeb é promulgada. *Senado Notícias*, 26 ago. 2020. Disponível em: <https://www12.senado.leg.br/noticias/materias/2020/08/26/emenda-constitucional-que-garante-a-permanencia-do-fundeb-e-promulgada>. Acesso em: 4 set. 2020.
- AIGNER, Dennis; LOVELL, C. A. Knox; SCHMIDT, Peter. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, v. 6, n. 1, p. 21-37, July 1977.

ALMEIDA, Aléssio Tony Cavalcanti de; AGUIAR, Margarida Noélia Cunha de. Eficiência dos gastos públicos em educação básica dos municípios nordestinos: 2007 a 2013. *Revista Econômica do Nordeste*, Fortaleza, v. 48, n. 4, p. 55-71, out./dez. 2017.

ALVES, Francione Charapa; FIALHO, Lia Machado Fiuza; LIMA, Maria Socorro Lucena. Formação em pesquisa para professores da educação básica. *Revista Tempos e Espaços em Educação*, São Cristóvão, SE, v. 11, n. 27, p. 285-300, set. 2018.

ALVES, Márcia da Conceição Pereira; BARROS, Renata Chrystina Bianchi de; CARROZZA, Guilherme. O Ideb e seus efeitos de sentido na Educação Básica do Brasil. *Revista Interfaces*, Guarapuava, PR, v. 9, n. 2, p. 29-40, jul./set. 2018.

ALVES, Pedro Jorge Holanda; ARAÚJO, Jevuks Matheus de. A study on the educational results obtained by municipalities of Paraíba in the years 2011, 2013 and 2015. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 26, n. 100, p. 1038-1057, jul./set. 2018.

AUBYN, Miguel St.; GARCIA, Filomena; PAIS, Joana. *Study on the efficiency and effectiveness of public spending on tertiary education*. Bruxelas: Directorate General of Economic and Financial Affairs, European Commission, 2009.

BARROS, Ricardo Paes de; HENRIQUES, Ricardo; MENDONÇA, Rosane Silva Pinto de. *A estabilidade inaceitável: desigualdade e pobreza no Brasil*. Brasília: Ipea, 2001. (Textos para Discussão, 800).

BATTESE, George Edward; COELLI, Tim J. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, v. 38, n. 3, p. 387-399, July 1988.

BATTESE, George Edward; COELLI, Tim J. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, v. 3, n. 1, p. 153-169, 1992.

BATTESE, George Edward; COELLI, Tim J. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, v. 20, n. 2, p. 325-332, June 1995.

BERTÊ, Ana Maria de Aveline; BORGES, Clayton Brito; BRUNET, Júlio Francisco Gregory. Qualidade do gasto público em educação nas redes públicas estaduais e municipais. *XIII Prêmio Tesouro Nacional*, 2008.

BISSOLI, Ana Cristina da Silva. *Evasão escolar: o caso do Colégio Estadual Antonio Francisco Lisboa*, 2010. Disponível em: http://www.repositorio.seap.pr.gov.br/arquivos/File/artigos/educacao/evasao_escolar.pdf. Acesso em: 20 maio 2017.

BRASIL. Constituição da República Federativa do Brasil, de 5 de outubro de 1988. *Diário Oficial da União*, Brasília, DF, 5 out. 1988. p. 1. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 10 dez. 2020.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep. *Sinopse Estatística da Educação Básica 2017*. Brasília, 2018. Disponível em: <http://portal.inep.gov.br/sinopses-estatisticas-da-educacao-basica>. Acesso em: 20 jan. 2017.

BRASIL. Lei n. 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. *Diário Oficial da União*, Brasília, DF, 23 dez. 1996. p. 27833.

BRASIL. Movimento Todos pela Educação. 2006. Disponível em: <http://www.todospelaeducacao.org.br/institucional/quem-somos>. Acesso em: 16 fev. 2017.

BRASIL. *Plano Nacional de Educação*. 2001. Disponível em: <http://portal.mec.gov.br/arquivos/pdf/pne.pdf>. Acesso em: 10 jan 2017.

BRASIL. Secretaria do Tesouro Nacional – STN. Finbra 2017. Brasília, 2017. Disponível em: <https://www.gov.br/tesouronacional/pt-br/estados-e-municipios/dados-consolidados/finbra-financas-municipais>. Acesso em: 11 fev. 2017.

CARD, David; KRUEGER, Alan B. *Labor market effects of school quality: theory and evidence*. Cambridge, MA: National Bureau of Economic Research, 1996.

COLEMAN, James S. *et al.* Equality of educational opportunity. In: KEEVES, John P.; WATANABE, Ryo (org.). *International handbook of educational research in the Asia-Pacific region*. Dordrecht: Kluwer Academic Publishers, 1966. p. 143-154.

COWAN, James; GOLDBERGER, Dan. National Board certification and teacher effectiveness: evidence from Washington State. *Journal of Research on Educational Effectiveness*, v. 9, n. 3, p. 233-258, May 2016.

DAVIES, Nicholas; ALCÂNTARA, Alzira Batalha. A evolução das matrículas na educação básica no Brasil. *Revista HISTEDBR On-line*, Campinas, SP, v. 20, e020016, 2020.

DIAZ, Maria Dolores Montoya. Qualidade do gasto público em educação no Brasil. Relatório Parcial. Instituto Brasileiro de Ética Concorrencial–ETCO, p. 47, nov. 2007.

DYNARSKI, Susan; HYMAN, Joshua; SCHANZENBACH, Diane Whitmore. Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, v. 32, n. 4, p. 692-717, Fall 2013.

FERNANDES, Reynaldo. *Índice de Desenvolvimento da Educação Básica (IDEB): metas intermediárias para a sua trajetória no Brasil, estados, municípios e escolas*. Brasília: MEC/Inep, 2007.

FERNANDES, Reynaldo; NARITA, Renata del Tedesco. Instrução superior e mercado de trabalho no Brasil. *Economia Aplicada*, São Paulo, v. 5, n. 1, p. 7-32, jan./mar. 2001.

FERREIRA, Nádia dos Santos; FERREIRA, Stefanny Bárbara de Jesus; SANTOS, Marcelo Henrique Pereira dos; ARAUJO, Edilson Pinheiro. Avaliação educacional em larga escala: uma análise sobre os resultados do Ideb e a visão dos gestores das escolas públicas do município de Salgueiro/PE. *Revista Opara*, Paulo Afonso, BA, v. 8, n. 2, p. 26-41, 2018.

FILMER, Deon; PRITCHETT, Lant. The effect of household wealth on educational attainment: evidence from 35 countries. *Population and Development Review*, v. 25, n. 1, p. 85-120, Mar. 1999.

FRIO, Gustavo Saraiva; FOCHEZATTO, Adelar; TRIACA, Livia Madeira; FINN, Eduardo Schirmer; BRAATZ, Jacó. Eficiência na educação: uma análise por escola no Rio Grande do Sul utilizando o método DEA em dois estágios. *Revista Brasileira de Estudos Regionais e Urbanos*, Curitiba, v. 12, n. 1, p. 74-89, jul. 2018.

GONÇALVES, Flávio de Oliveira; FRANÇA, Marco Túlio Aniceto. Eficiência na provisão de educação pública municipal: uma análise em três estágios dos municípios brasileiros. *Estudos Econômicos*, São Paulo, v. 43, n. 2, p. 271-299, abr./jun. 2013.

GREENE, William H. A gamma-distributed stochastic frontier model. *Journal of Econometrics*, v. 46, n. 1-2, p. 141-163, Oct./Nov. 1990.

GUPTA, Sanjeev; VERHOEVEN, Marijn; TIONGSON, Erwin. R. The effectiveness of government spending on education and health care in developing and transition economies. *European Journal of Political Economy*, v. 18, n. 4, p. 717-737, Nov. 2002.

GURYAN, Jonathan. *Does money matter?* Regression-discontinuity estimates from education finance reform in Massachusetts. Cambridge, MA: National Bureau of Economic Research, 2001.

HANUSHEK, Eric A. Interpreting recent research on schooling in developing countries. *The World Bank Research Observer*, v. 10, n. 2, p. 227-246, Aug. 1995.

HANUSHEK, Eric A.; RAYMOND, Margaret E. Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, v. 24, n. 2, p. 297-327, Spring 2005.

HECKMAN, James J. The case for investing in disadvantaged young children. *CESifo DICE Report*, v. 6, n. 2, p. 3-8, Summer 2008.

HEINIGER, Melina; IMDORF, Christian. The role of vocational education in the transmission of gender segregation from education to employment: Switzerland and Bulgaria compared. *Journal for Labour Market Research*, v. 52, n. 1, p. 1-21, Dec. 2018.

HERRERA, Santiago; PANG, Gaobo. *Efficiency of public spending in developing countries: an efficiency frontier approach*. Washington, DC: The World Bank, 2005.

HOXBY, Caroline M. The effects of class size on student achievement: new evidence from population variation. *The Quarterly Journal of Economics*, v. 115, n. 4, p. 1239-1285, Nov. 2000.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Banco de dados do Sidra*. Disponível em: <https://sidra.ibge.gov.br/home/ipp/brasil>. Acesso em: 11 fev. 2017.

JACOB, Brian; ROCKOFF, Jonah E.; TAYLOR, Eric S.; LINDY, Benjamin; ROSEN, Rachel. Teacher applicant hiring and teacher performance: evidence from DC public schools. *Journal of Public Economics*, v. 166, p. 81-97, Oct. 2018.

JEPSEN, Christopher; RIVKIN, Steven. Class size reduction and student achievement the potential tradeoff between teacher quality and class size. *Journal of Human Resources*, v. 44, n. 1, p. 223-250, Winter 2009.

KODDE, David A.; PALM, Franz C. Wald criteria for jointly testing equality and inequality restrictions. *Econometrica*, v. 54, n. 5, p. 1243-1248, Sept. 1986.

KUMBHAKAR, Subal C.; LOVELL, C. A. Knox. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press, 2000.

LEUVEN, Edwin; LØKKEN, Sturla A. Long-term impacts of class size in compulsory school. *Journal of Human Resources*, v. 55, n. 1, p. 309-348, Aug. 2018.

LIBÂNEO, José Carlos. Didática e epistemologia: para além do embate entre a didática e as didáticas específicas. In: VEIGA, Ilma P. A.; DÁVILA, Cristina Maria. *Profissão docente: novos sentidos, novas perspectivas*. 2. ed. Campinas, SP: Papyrus, 2008.

MACÊDO, Francisca Francivânia Rodrigues Ribeiro; KLOEPEL, Nilton Roberto; RODRIGUES JUNIOR, Moacir Manoel; SCARPIN, Jorge Eduardo. Análise da eficiência dos recursos públicos direcionados à educação: estudo nos municípios do Estado do Paraná. In: CONGRESSO BRASILEIRO DE CUSTOS, 19, 12 a 14 de novembro de 2012, Bento Gonçalves, RS. *Anais [...]*, Bento Gonçalves, RS: Associação Brasileira de Custos, 2012. p. 1-16.

MACGREGOR, Karen. South Africa: student drop-out rates alarming. *University World News*, 28 Oct. 2007. Disponível em: <https://www.universityworldnews.com/post.php?story=20071025102245380>. Acesso em: dez. 2018.

MADURO JÚNIOR, Paulo Rogério Rodrigues. *Taxas de matrícula e gastos em educação no Brasil*. 2007. Dissertação (Mestrado em Economia) – Fundação Getúlio Vargas, Rio de Janeiro, 2007.

MANDL, Ulrike; DIERX, Adriaan; ILZKOVITZ, Fabienne. *The effectiveness and efficiency of public spending*. Bruxelas: Directorate General of Economic and Financial Affairs, European Commission, 2008.

MARINHO, Emerson; ATALIBA, Flávio. Avaliação do crescimento da produtividade e progresso tecnológico dos estados do Nordeste com o uso da fronteira de produção estocástica. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 30, n. 3, p. 427-452, dez. 2000.

MARZOCCHI, Maura; OLIVEIRA, Adailda Gomes. Interlocação entre autonomia escolar e desempenho: quais as analogias possíveis. In: CONGRESSO BRASILEIRO DE SOCIOLOGIA, 14., 2009, Rio de Janeiro. *Anais [...]*. Rio de Janeiro: SBS, 2009.

MATOS, Daniel Abud Seabra; RODRIGUES, Erica Castilho. Indicadores educacionais e contexto escolar: uma análise das metas do Ideb. *Estudos em Avaliação Educacional*, São Paulo, v. 27, n. 66, p. 662-688, set./dez. 2016.

MEEUSEN, Wim; VAN DEN BROECK, Julien. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, v. 18, n. 2, p. 435-444, jun. 1977.

MINCER, Jacob. *Schooling, experience, and earnings*. New York: Columbia University Press, 1974.

MONTEIRO, Joana; ROCHA, Rudi. Drug battles and school achievement: evidence from Rio de Janeiro's favelas. *Review of Economics and Statistics*, v. 99, n. 2, p. 213-228, May 2017.

MUSGRAVE, Richard A. A multiple theory of budget determination. *FinanzArchiv*, v. 17, n. 3, p. 333-343, 1956.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OCDE. OECD database. 2017. Disponível em: https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS. Acesso em: 15 mar. 2017.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OCDE. *Pisa 2009 results: overcoming social background*. Paris: OCDE, 2010.

PEREIRA, Manuel Coutinho; MOREIRA, Sara. Eficiência das escolas secundárias portuguesas: uma análise de fronteira de produção estocástica. *Boletim Económico do Banco de Portugal*, v. 13, n. 1, p. 101-118, primavera 2007.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO – PNUD. Disponível em: <http://www.br.undp.org/>. Acesso em: 25 jan. 2017.

QUTB, Rasha Muhammad. Analyzing the external and internal efficiency considerations in public subsidization of education in Egypt. *Journal of Economics and Sustainable Development*, v. 7, n. 12, p. 164-172, 2016.

SCARPIN, Jorge Eduardo; MACÊDO, Francisca Francivânia Rodrigues Ribeiro; STAROSKY FILHO, Loriberto; RODRIGUES JÚNIOR, Moacir Manoel. Análise da eficiência dos recursos públicos direcionados à educação: estudo nos municípios do estado de Santa Catarina. *Gestão Pública: Práticas e Desafios*, Recife, v. 3, n. 2, p. 27-48, dez. 2012.

SCHULTZ, Theodore W. Investment in human capital. *The American Economic Review*, v. 51, n. 1, p. 1-17, Mar. 1961.

SILVA, Nelson do Valle; HASENBALG, Carlos. Tendências da desigualdade educacional no Brasil. *Dados: Revista de Ciências Sociais*, Rio de Janeiro, v. 43, n. 3, p. 423-445, 2000.

SILVA, Givanildo da; SILVA, Alex Vieira da; SANTOS, Inalda Maria dos. O Ideb e as políticas públicas educacionais: estratégias, efeitos e consequências. *Revista Exitus*, Santarém, PA, v. 9, n. 1, p. 258-285, jan./mar. 2019. Disponível em: <https://doi.org/10.24065/2237-9460.2019v9n1ID723>. Acesso em: 11 out. 2019.

SOARES, José Francisco; XAVIER, Flávia Pereira. Pressupostos educacionais e estatísticos do Ideb. *Educação & Sociedade*, Campinas, SP, v. 34, n. 124, p. 903-923, set. 2013.

SOUSA, Uelitania Dantas de; SANTOS, Josaias Santana dos; CALADO, João Eudes de Souza; CAVALCANTE JÚNIOR, Florisvaldo Cunha; SOUSA, Wellington Dantas de. Gastos públicos com educação e meta do Ideb no ensino fundamental: uma análise dos municípios do sertão de Pernambuco. *Brazilian Journal of Development*, Curitiba, v. 4, n. 5, Edição Especial, p. 2578-2599, ago. 2018.

SOUZA, Wallace Patrick Santos de Farias; OLIVEIRA, Victor Rodrigues de; ANNEGUES, Ana Cláudia. *Background* familiar e desempenho escolar: uma abordagem não paramétrica. *Pesquisa e Planejamento Econômico*, v. 48, n. 2, p. 133-162, ago. 2018.

STEVENSON, Rodney E. Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, v. 13, n. 1, p. 57-66, May 1980.

SUTHERLAND, Douglas; PRICE, Robert; JOURMARD, Isabelle; NICQ, Chantal. *Performance indicators for public spending efficiency in primary and secondary education*. [S.l.]: OCDE, 2007. (Economics Department Working Papers, n. 546).

VARIAN, Hal R. *Markets for information goods*. Tóquio: Institute for Monetary and Economic Studies, Bank of Japan, 1999.

WANG, Hung-Jen. Heteroscedasticity and non-monotonic efficiency effects of a stochastic frontier model. *Journal of Productivity Analysis*, v. 18, n. 3, p. 241-253, Nov. 2002.

NOTA: Os autores contribuíram para a elaboração do artigo da seguinte maneira: Pedro Jorge Holanda Alves – elaboração dos gráficos, figuras, estimação dos resultados e escrita de parte dos resultados; Jevuks Matheus Araújo – elaboração da introdução e revisão do texto; Gustavo Saraiva Frio – elaboração da introdução, revisão da literatura, parte dos resultados e considerações finais; Lyvia Cabral Cordeiro – construção do objetivo de pesquisa e elaboração inicial do texto.

Recebido em: 16 JANEIRO 2020

Aprovado para publicação em: 27 NOVEMBRO 2020



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

AGRADECIMENTOS AGRADECIMIENTOS ACKNOWLEDGEMENTS

<https://doi.org/10.18222/eae.v31i78.8427>

Aos pareceristas que contribuíram, durante o ano de 2020, avaliando e selecionando artigos para publicação na revista, a fim de garantir a qualidade de seu padrão científico.

Adelaide Alves Dias

(Universidade Federal da Paraíba – UFPB, João Pessoa-PB, Brasil)

Adelina Novaes

(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Adenize Costa Acioli

(Secretaria de Estado de Educação e de Esporte de Alagoas – SEEE/AL, Maceió-AL, Brasil)

Adilson Dalben

(Centro Universitário Salesiano de São Paulo – Unisal, Campinas-SP, Brasil)

Adolfo Samuel de Oliveira

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Adriana Justin Cerveira Kampff

(Pontifícia Universidade Católica do Rio Grande do Sul – PUC-RS, Porto Alegre-RS, Brasil)

Adriano Moro

(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Adriano Souza Senkevics

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Alexandre José de Souza Peres

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Aline Cristina da Cruz

(Universidade Federal de São João Del Rei – UFSJ, São João Del Rei-MG, Brasil)

Almir Cléydison Joaquim da Silva

(Universidade Federal da Paraíba – UFPB, João Pessoa-PB, Brasil)

Amélia Cristina Abreu Artes

(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Ana Lúcia Manrique

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Ana Maria da Costa Santos

(Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp, Presidente Prudente-SP, Brasil)

Ana Maria Gimenes Corrêa Calil

(Universidade de Taubaté – Unitau, Taubaté-SP, Brasil)

Ana Maria Klein

(Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp, São José do Rio Preto-SP, Brasil)

Andrea Baptista Almeida

(Centro de Instrução e Adestramento Almirante Newton Braga – Cianb, Rio de Janeiro-RJ, Brasil)

Angela Magda Rodrigues Virgolim

(Universidade de Brasília – UnB, Brasília-DF, Brasil)

Angela Maria Martins

(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Angelica Caetano da Silva

(Colégio Pedro II, Rio de Janeiro-RJ, Brasil)

Anna Carolina Salgado Jardim

(Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – IFSP, São Roque-SP, Brasil)

Carla Viana Coscarelli

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Carlos Eduardo Mathias Motta

(Universidade Federal Fluminense – UFF, Niterói-RJ, Brasil)

Cenilza Pereira dos Santos

(Universidade do Estado da Bahia – Uneb, Salvador-BA, Brasil)

Cláudia Maria Bogus

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Claudia Regina Baukat Silveira Moreira

(Universidade Federal do Paraná – UFPR, Curitiba-PR, Brasil)

Cristiane Fuzer

(Universidade Federal de Santa Maria – UFSM, Santa Maria-RS, Brasil)

Cristiane Machado

(Universidade Estadual de Campinas – Unicamp, Campinas-SP, Brasil)

Cristina Botti de Souza Rossetto

(Universidade Federal do Paraná – UFPR, Curitiba-PR, Brasil)

Daniel Bartholomeu

(Centro Universitário Fieo – Unifieo, São Paulo-SP, Brasil)

Darlan Christiano Kroth

(Universidade Federal da Fronteira Sul – UFFS, Chapecó-SC, Brasil)

Denise Cristina Correa da Rocha

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Diana Gomes da Silva Cerdeira

(Universidade do Estado do Rio de Janeiro – UERJ, Rio de Janeiro-RJ, Brasil)

Dirce Aparecida Foletto de Moraes

(Universidade Estadual de Londrina – UEL, Londrina-PR, Brasil)

Domingos Manuel Barros Fernandes

(Instituto Universitário de Lisboa – ISCTE, Lisboa, Portugal)

Donaldo Bello de Souza

(Universidade do Estado do Rio de Janeiro – UERJ, Rio de Janeiro-RJ, Brasil)

Elieuzza Aparecida de Lima

(Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp, Marília-SP, Brasil)

Elisangela da Silva Bernardo

(Universidade Federal do Estado do Rio de Janeiro – UNIRIO, Rio de Janeiro-RJ, Brasil)

Eliza Bartolozzi Ferreira

(Universidade Federal do Espírito Santo – UFES, Vitória-ES, Brasil)

Elizabeth Balbachevsky

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Eloisa Maia Vidal

(Universidade Estadual do Ceará – UECE, Fortaleza-CE, Brasil)

Érica Maria Toledo Catalani

(Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária – Cenpec, São Paulo-SP, Brasil)

Fábio Domingues Waltenberg

(Universidade Federal Fluminense – UFF, Niterói-RJ, Brasil)

Fernanda Cristina Barbosa Pereira Queiroz

(Universidade Federal do Rio Grande do Norte – UFRN, Natal-RN, Brasil)

Fernando Augusto Silva

(Serviço Nacional de Aprendizagem Industrial – Senai – Escola Senai Luiz Simon, São Paulo-SP, Brasil)

Fernando de Souza Coelho

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Filipe Ferreira Ghidetti

(Universidade Federal do Espírito Santo – UFES, Vitória-ES, Brasil)

Flavia Pereira Xavier

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Flávia Vitor Longo

(Universidade Estadual de Campinas – Unicamp, Campinas-SP, Brasil)

Francisca Izabel Pereira Maciel

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Gabriela Schneider

(Universidade Federal do Paraná – UFPR, Curitiba-PR, Brasil)

Gerusa de Fátima Tomé Sabino

(Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, Diamantina-MG, Brasil)

Géssica Priscila Ramos

(Universidade Federal de São Carlos – UFSCar, São Carlos-SP, Brasil)

Gilberto José Miranda

(Universidade Federal de Uberlândia – UFU, Uberlândia-MG, Brasil)

Girlene Ribeiro de Jesus

(Universidade de Brasília – UnB, Brasília-DF, Brasil)

Gisela Lobo B. P. Tartuce

(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Gizele de Souza

(Universidade Federal do Paraná – UFPR, Curitiba-PR, Brasil)

Gleiber Couto Santos

(Universidade Federal de Goiás – UFG, Catalão-GO, Brasil)

Hélio Radke Bittencourt

*(Pontifícia Universidade Católica do Rio Grande do Sul – PUC-RS,
Porto Alegre-RS, Brasil)*

Hustana Maria Vargas

(Universidade Federal Fluminense – UFF, Niterói-RJ, Brasil)

Iracema Campos Cusati

(Universidade de Pernambuco – UPE, Petrolina-PE, Brasil)

Itamar Mendes da Silva

(Universidade Federal do Espírito Santo – UFES, Vitória-ES, Brasil)

Jamerson Viegas Queiroz

(Universidade Federal do Rio Grande do Norte – UFRN, Caicó-RN, Brasil)

João Luiz Horta Neto

*(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep,
Brasília-DF, Brasil)*

José Carlos Rothen

(Universidade Federal de São Carlos – UFSCar, São Carlos-SP, Brasil)

Josiane Souza de Paula

(Universidade Federal de Uberlândia – UFU, Uberlândia-MG, Brasil)

Julia Polessa Maçaíra

(Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro-RJ, Brasil)

Julio Cesar Torres

*(Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp,
São José do Rio Preto-SP, Brasil)*

Karina Alves Biasoli Stanich

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Lara Elena Ramos Simielli

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Laurizete Ferragut Passos

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Lenice Medeiros

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Leonardo Rosa

(Stanford University – Palo Alto, Califórnia, Estados Unidos)

Ligia Martha Coimbra da Costa Coelho

(Universidade Federal do Estado do Rio de Janeiro – Unirio, Rio de Janeiro-RJ, Brasil)

Lílian Ghiuro Passarelli

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Lourdinha Maria de Lourdes Lima Bragion

(Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas – IFSULDEMINAS, Belo Horizonte-MG, Brasil)

Lucia Alves Faria Mattos

(Faculdade de Estudos Administrativos de Minas Gerais – FEAD, Belo Horizonte-MG, Brasil)

Luciano Silva

(Universidade Federal de Ouro Preto – UFOP, Ouro Preto-MG, Brasil)

Lucilia Regina de Souza Machado

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Luiz Guilherme Dácar da Silva Scorzafave

(Universidade de São Paulo – USP, Ribeirão Preto-SP, Brasil)

Luiza Helena da Silva Christov

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Lys Maria Vinhaes Dantas

(Universidade Federal do Recôncavo da Bahia – UFRB, Cachoeira-BA, Brasil)

Marco Túlio Ancieto França

*(Pontifícia Universidade Católica do Rio Grande do Sul – PUC-RS,
Porto Alegre-RS, Brasil)*

Marcos Vinicius Francisco

(Universidade do Oeste Paulista – Unoeste, Presidente Prudente-SP, Brasil)

Maria Aparecida de Oliveira Freitas

(Universidade Federal de São Paulo – Unifesp, São Paulo-SP, Brasil)

Maria Carolina Tomás

*(Pontifícia Universidade Católica de Minas Gerais – PUC-MG,
Belo Horizonte-MG, Brasil)*

Maria Clara Di Pierro

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Maria de Fátima Barbosa Abdalla

(Universidade Católica de Santos – Unisantos, Santos-SP, Brasil)

Maria Dyrce Dias Meira

(Centro Universitário Adventista de São Paulo – Unasp, São Paulo-SP, Brasil)

Maria Fernanda Rezende Nunes

(Universidade Federal do Estado do Rio de Janeiro – Unirio, Rio de Janeiro-RJ, Brasil)

Maria José Oliveira Duboc

(Universidade Estadual de Feira de Santana – UEFS, Feira de Santana-BA, Brasil)

Maria José Viana Marinho de Mattos

*(Pontifícia Universidade Católica de Minas Gerais – PUC-Minas,
Belo Horizonte-MG, Brasil)*

Maria Teresa Gonzaga Alves

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Mariane Campelo Koslinski

(Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro-RJ, Brasil)

Mariangela Abrão

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Marisa Assunção Teixeira

(Secretaria de Educação da Prefeitura do Município de São Bernardo do Campo, São Bernardo do Campo-SP, Brasil)

Marisa Brandão Rocha

(Centro Federal de Educação Tecnológica Celso Suckow da Fonseca – CEFET, Rio de Janeiro-RJ, Brasil)

Mary Ângela Brandalise

(Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, Brasília-DF, Brasil)

Mauren Porciúncula Moreira da Silva

(Universidade Federal do Rio Grande – FURG, Rio Grande, Rio Grande do Sul-RS, Brasil)

Melissa Caldeira Brant de Souza Lima

(International Institute for Applied Systems Analysis – IIASA, Luxemburgo, Áustria)

Naira da Costa Muylaert Lima

(Universidade Federal de Juiz de Fora – UFJF, Juiz de Fora-MG, Brasil)

Nathalia Cassettari

(Universidade de Brasília – UnB, Brasília-DF, Brasil)

Nayara Abreu Julião

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Nilce Maria da Silva Campos Costa

(Universidade Federal de Goiás – UFG, Goiânia-GO, Brasil)

Ocimar Munhoz Alavarse
(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Paulo de Martino Jannuzzi
(Instituto Brasileiro de Geografia e Estatística – IBGE, Rio de Janeiro-RJ, Brasil)

Paulo Henrique Arcas
(Instituto Internacional de Planeamiento de la Educación – IIPE Unesco,
Buenos Aires, Argentina)

Paulo Vitor Souza de Souza
(Universidade Federal do Pará – UFPA, Belém-PA, Brasil)

Raquel Alvares
(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Raquel Aparecida Marra da Madeira Freitas
(Pontifícia Universidade Católica de Goiás – PUC-Goiás, Goiânia-GO, Brasil)

Raquel da Cunha Valle
(Fundação Carlos Chagas – FCC, São Paulo-SP, Brasil)

Raul Aragão Martins
(Universidade Estadual Paulista "Júlio de Mesquita Filho" – Unesp,
São José do Rio Preto-SP, Brasil)

Regina Madalozzo
(Insper, São Paulo-SP, Brasil)

Renato de Sousa Porto Gilioli
(Câmara dos Deputados, Consultoria Legislativa, Brasília-DF, Brasil)

Rodnei Pereira
(Universidade Paulista – UNIP, São Paulo-SP, Brasil)

Rodrigo Travitzki
(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Ronaldo Castro D'Ávila
(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Rosana Rodrigues Heringer

(Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro-RJ, Brasil)

Rosangela Fritsch

(Universidade do Vale do Rio dos Sinos – Unisinos, São Leopoldo-RS, Brasil)

Saddo Ag Almouloud

(Universidade Federal do Pará – UFPA, Belém-PA, Brasil)

Sandra Maria Zákia Lian Sousa

(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Sergio Martinic

(Universidad de Aysén – Uaysen, Coyhaique, Chile)

Sergio Vasconcelos de Luna

(Pontifícia Universidade Católica de São Paulo – PUC-SP, São Paulo-SP, Brasil)

Silke Weber

(Universidade Federal de Pernambuco – UFPE, Recife-PE, Brasil)

Silvia Vieira Cruz

(Universidade Federal do Ceará – UFC, Fortaleza-CE, Brasil)

Silvina Julia Fernández

(Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro-RJ, Brasil)

Solange Maria Santos

(Scientific Electronic Library Online – SciELO, São Paulo-SP, Brasil)

Thiago Alves

(Universidade Federal de Goiás – UFG, Goiânia-GO, Brasil)

Tiago Claudino Barbosa

(Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil)

Tufi Machado Soares

(Universidade Federal de Juiz de Fora – UFJF, Juiz de Fora-MG, Brasil)

Valéria Aparecida de Souza Siqueira
(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Valéria Virgínia Lopes
(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Victor Maia Senna Delgado
(Universidade Federal de Ouro Preto – UFOP, Mariana-MG, Brasil)

Virgínia Louzada
(Universidade do Estado do Rio de Janeiro – UERJ, Rio de Janeiro-RJ, Brasil)

Wagner Bandeira Andriola
(Universidade Federal do Ceará – UFC, Fortaleza-CE, Brasil)

Walkiria de Oliveira Rigolon
(Universidade Estadual de Campinas – Unicamp, Campinas-SP, Brasil)

Wolney Cândido de Melo
(Universidade de São Paulo – USP, São Paulo-SP, Brasil)

Zacarias Jaegger Gama
(Universidade do Estado do Rio de Janeiro – UERJ, Rio de Janeiro-RJ, Brasil)